



X CONGRESSO NAZIONALE SISMEC

Società Italiana di Statistica Medica ed Epidemiologia Clinica

Nuovi disegni nella ricerca clinica: la sfida della complessità tra etica e salute

ATTI



sismec



2019 | X CONGRESSO NAZIONALE SISMEC

Nuovi disegni nella ricerca clinica: la sfida della complessità tra etica e salute.

Segretaria di redazione: *Valentina Apostoli*

Edizione: *Marzo 2020*

A cura di: *Consiglio Direttivo SISMEC*

Copyright: *SISMEC*

ISBN: 978-88-943456-2-9

sismec

Si ringraziano

Janssen Pharmaceutical Companies of Johnson & Johnson

Bristol-Myers Squibb

Éthos

SERVIER

per il contributo non condizionato
alla realizzazione degli atti

CONSIGLIO DIRETTIVO SISMEC

PRESIDENTE
Annarita Vestri
Sapienza - Università di Roma

PRESIDENTE ELETTO
Paolo Trerotoli
Università degli Studi di Bari

SEGRETERIO
Miriam Isola
Università degli Studi di Udine

TESORIERE
Marta DI Nicola
Università di Chieti-Pescara

CONSIGLIERI
Susanna Conti
Istituto Superiore di Sanità
Referente privacy e gestione dati personali

Mirko Di Martino
Regione Lazio
Referente per EBPH

Domenica Matranga
Università degli Studi di Palermo
Referente per i social media

Giovanni Veronesi
Università degli Studi dell'Insubria
Referente per il sito

PRESIDENTE DEL CONGRESSO

Annarita Vestri
Sapienza - Università di Roma

CONSIGLIO SCIENTIFICO

Alessio Farcomeni
Sapienza - Università di Roma

Mario Di Traglia
Sapienza - Università di Roma

Viviana Egidi
Sapienza - Università di Roma

Diana Giannarelli
Unità di biostatistica IFO

Italo Nofroni
Sapienza - Università di Roma

Patrizio Pasqualetti
Afar associazione Fatebenefratelli per la ricerca

Alessandra Spagnoli
Sapienza - Università di Roma

Annarita Vestri
Sapienza - Università di Roma

COMITATO ORGANIZZATORE LOCALE

Danilo Alunni Fegatelli
Sapienza - Università di Roma

Arianna Di Rocco
Sapienza - Università di Roma

Alberto Milanese
Sapienza - Università di Roma

X CONGRESSO NAZIONALE SISMEC

Società Italiana di Statistica Medica ed Epidemiologia Clinica

Nuovi disegni nella ricerca clinica:

la sfida della complessità tra etica e salute

Il mondo della ricerca clinica è in una fase di profonda trasformazione, dovuta sia all'accelerazione delle conoscenze scientifiche degli ultimi anni, sia ad approcci metodologici che tendono a superare il modello classico di sviluppo del farmaco, che vanno dai nuovi disegni sperimentali a modelli di accesso anticipato, fino all'utilizzo dei dati di real world evidence. Le autorità regolatorie Food and Drug Administration ed European Medicines Agency hanno manifestato una disponibilità ad accettare tali nuovi approcci, al fine di garantire un rapido accesso dei pazienti ai nuovi farmaci e per contenere i costi di sviluppo.

La ricerca clinica costituisce un momento chiave nel processo di sviluppo dei farmaci e delle altre tecnologie sanitarie; oltre al valore strumentale, nell'ambito del lungo percorso che conduce la tecnologia biomedica dall'ideazione fino al letto del paziente, le sperimentazioni cliniche generano un valore in sé. Infatti, garantiscono "valore" misurabile in dimensioni diverse e sotto prospettive multiple. I pazienti, il sistema sanitario pubblico e i suoi operatori, i cittadini più in generale e la stessa industria, misurano il valore di una sperimentazione clinica valutando in modo diverso agli aspetti della sicurezza, dell'efficacia clinica e dell'impatto economico e organizzativo generato ai diversi livelli del sistema dal processo sperimentale. Il tema dei nuovi disegni di studio viene affrontato nella prima sessione plenaria di questo Convegno, sia dal punto di vista metodologico che applicativo, al fine di definire se, e come, i nuovi disegni possano contribuire ad una precoce approvazione del farmaco e quindi ad un maggiore beneficio per il paziente.

La seconda sessione plenaria affronterà il tema della normativa in ambito della ricerca clinica. La pubblicazione del nuovo Regolamento EU n.536/2014 rappresenta uno strumento fondamentale per garantire e semplificare procedure e metodologie coerenti ed omogenee in tutti gli Stati Europei; l'adozione di un unico quadro normativo si ritiene possa contribuire a migliorare l'efficienza del processo regolatorio, a sostenere la competitività dell'Europa nella ricerca clinica ed a favorire l'accesso dei pazienti a terapie innovative. L'entrata in vigore del nuovo regolamento verrà accompagnata da una riorganizzazione dei Comitati Etici, ma rimangono ancora aperti tanti problemi tra i quali la valutazione degli studi osservazionali che non rientrano in tale normativa. Su questa tipologia di studi, che ad oggi non è chiaro da quale organismo saranno valutati, si accendono spesso discussioni sulla loro definizione e sul loro utilizzo anche a fini registrativi; sulla definizione dei criteri per poter definire uno studio come finalizzato, o meno, al "miglioramento della pratica clinica"; sulla corretta classificazione, da parte dei comitati etici, di uno studio come "no-profit"; sulla complessità delle procedure e dei documenti necessari per l'approvazione e ancora sull'idoneità dei Centri che svolgono sperimentazioni.

Il congresso chiude i lavori con la tavola rotonda, dedicata al tema delle disuguaglianze di salute, argomento verso il quale negli ultimi 15 anni è cresciuta l'attenzione, non solo in Italia, ma anche in ambito europeo. Nel nostro Paese si sono osservati, nelle ultime decadi, progressivi miglioramenti delle condizioni di salute e bassi livelli di disuguaglianza tra classi sociali, anche se meno pronunciati rispetto ad altri Paesi europei. Ma sappiamo come indicatori "medi" possono mascherare l'esistenza di differenze sistematiche: gli individui più abbienti stanno meglio, si ammalano di meno e vivono più a lungo. Allo stesso modo, le regioni italiane più povere mostrano indicatori di salute meno favorevoli. Tali disuguaglianze, se non giustificate dal punto di vista biologico, possono essere definite inique perché evitabili e riconducibili, in parte, ai determinanti sociali cioè a quei fattori che determinano le differenze negli esiti della salute.

Nel tentativo di ridurre la portata e l'impatto delle disuguaglianze, nel corso degli anni, sono stati predisposti alcuni importanti atti di programmazione sanitaria, in linea con le indicazioni dell'Unione Europea. Pertanto il fine di questo incontro è accrescere il livello di consapevolezza e l'attenzione al tema delle disuguaglianze di salute, attraverso le testimonianze di alcuni tra i maggiori esponenti esperti dell'argomento, con l'auspicio di evidenziare il ruolo centrale del settore sanitario nel processo di inserimento dell'equità nella programmazione delle politiche pubbliche.

Auspico che anche la SISMEC possa collaborare, in termini di sviluppi metodologici, con gli stakeholder che si occupano di disuguaglianze di salute, al fine di pervenire a risultati significativi in termini di impatto sulle disuguaglianze.

Infine, ringrazio vivamente tutto il Consiglio Direttivo, il Consiglio Scientifico e i miei Collaboratori, che hanno contribuito alla realizzazione di questo evento.

Annarita Vestri

Epidemiologia

Evoluzione della mortalità per tutte le cause nei residenti della provincia di Pavia dal 1980 al 2015 **2**
Borrelli Paola, Trivelli Leonardo, Gnesi Marco, Cadum Ennio, Villani Simona

Geographic heterogeneity in non-communicable diseases in relation to quality of work in europe: evidence from the sharelife survey **6**
Domenica Matranga, Filippa Bono

Variation of multiple sclerosis risk in the province of Pavia: investigating pm2.5 exposure **9**
Monti Maria Cristina, Trivelli Leonardo, Gerosa Leonardo, Pisoni Enrico, Borrelli Paola, Mallucci Giulia, Bergamaschi Roberto, Montomoli Cristina

Adherence to the mediterranean diet and stomach cancer risk: results from the stomach cancer pooling project **14**
Bravi Francesca, Di Maso Matteo, Negri Eva, Palli Domenico, Lagiou Pagona, Lunet Nuno, Aragonés Nuria, Vioque Jesus, López-Carrillo Lizbeth, La Vecchia Carlo, Ferraroni Monica

Il ruolo dei determinanti Early-Life e Mid-Life sullo stato cognitivo **18**
Simeon Vittorio, Dolce Pasquale, Gallo Valentina, Signoriello Simona, Signoriello Giuseppe, Amalia Mattiello, Panico Salvatore, Chiodini Paolo

Cohort analysis of epithelial cancer mortality male to female sex ratios in the european union **22**
Malvezzi Matteo

Trattamenti farmacologici potenzialmente inappropriati e interazioni farmacologiche a rischio di eventi avversi **28**
Lallo Adele, Alagna Michela, Onder Graziano, Ricci Roberto, Chiriatti Alberto, Di Martino Mirko, Davoli Marina

Total serum bile acids concentration for the diagnosis of intrahepatic cholestasis of pregnancy: a meta-analysis using individual patient data **34**
Casazza Giovanni, Manzotti Cristina, Nikolova Dimitrinka, Colli Agostino

Utilizzo appropriato del Trastuzumab nel Lazio: scenari terapeutici e stima del risparmio per il SSN **37**
Alessandro Cesare Rosa, Felice Musicco, Paolo Marchetti, Nello Martini, Antonio Addis, Valeria Belleudi

Valutazione del percorso diagnostico terapeutico assistenziale nei pazienti con il deficit dell'ormone della crescita in età evolutiva nella regione Marche. risultati preliminari **39**
Skrami Edlira, Dajko Marianxhela, Gesuita Rosaria, Bucci Andrea, Faragalli Andrea, Ferrante Luigi, Pricci Flavia, Rotondi Daniela, Agazio Elvira, Patregnani Luigi, Brugia Marina, Cherubini Valentino, Carle Flavia

Day-time patterns of energy intake derived via principal component analysis and their relationship with body mass index and waist circumference in adolescents: an analysis of the uk national diet and nutrition survey (2008-2012) **42**
Luigi Palla, Suzana Almoosawi

Analisi e valutazione di appropriatezza nel percorso di cura del carcinoma della mammella in regione piemonte. risultati dello studio E.PIC.A **47**
Ricotti Andrea, Mistrangelo Marinella, Berchialla Paola, Dalmasso Paola, Pierini Andrea, Bono Danilo, Moirano Fulvio, Bertetto Oscar, Messori Ioli Giovanni

Diffuse and intestinal histotypes prevail respectively in young and old gastric cancer cases **52**
Verlato Giuseppe, Torroni Lorena, Bencivenga Maria, Giacopuzzi Simone, Antonucci Nadia, Vesentini Roberta, Negri Eva, La Vecchia Carlo, De Manzoni Giovanni

Un progetto lavoro di internazionalizzazione e armonizzazione della ricerca su farmaci innovativi, innovazione delle ricerche epidemiologiche, genetiche e di sviluppo globale della ricerca in Europa / Italia ai paesi orientali / Iran azerbaigian / Kazakistan / Cina 2015 2025 **54**
Franco Naccarella, Domenico Criscuolo, Giulio He, Bin He, Sophie Qi, Lei Li, Janxiun Liu, Lei Sun, Weynan Li, Youlin Qiao, Meray Abbasov, Zeynal Akparov e Siamak Saber, Pirooz Hebrahimi, Reza Saedi, Antonella Tognin, Mahmood Jamei Masjed

Metodi Biostatistici

Population size estimation with interval-censored counts and external information: prevalence of multiple sclerosis in Rome **57**
Farcomeni Alessio, Danilo Alunni Fegatelli, Arianna Di Rocco, Vestri Anna Rita

Early prediction of evidence of disease activity in multiple sclerosis patients **59**
Trivelli L., Mallucci G., Colombo E., Trojano M., Amato MP., Zaffaroni M., Iaffaldano P., Razzolini L., Saraceno L., Montomoli C., Bergamaschi R.

A bayesian mendelian randomization approach to identify disease causing proteins via pedigree data, partially observed exposures and correlated instruments **63**
Teresa Fazia, Leonardo Egidi, Burcu Ayoglu, Ashley Beecham, Jacob L. McCauley, Peter Nilsson, Carlo Berzuini, Luisa Bernardinelli

Pattern dietetici e malattie respiratorie in GEIRD **66**
Chamitava Liliya, Zanolin Maria Elisabetta, Garcia-Larsen Vanessa, Ferrari Marcello, Cerveri Isa, Pirina Pietro, Verlato Giuseppe, Cazzoletti Lucia

Potential impact fraction for a continuous risk factor: assessing the burden of oral and pharyngeal cancer according to adherence to the healthy eating index **71**
Di Maso Matteo, Tomaino Laura, Garavello Werner, Crispo Anna, Serraino Diego, Ferraroni Monica, La Vecchia Carlo, Edefonti Valeria, Bravi Francesca

A machine learning approach for inspecting relationships between parkinson disease and metal exposure **77**
Vezzoli Marika, Renzetti Stefano, Parrinello Giovanni, Calza Stefano

Silver Russell Syndrome: RBF-PSO mathematical model for the growth evaluation in patients exposed and not exposed to GH treatment **79**
Stura Ilaria, Matarazzo Patrizia, Brondino Francesca Maria, De Sanctis Luisa, Migliaretti Giuseppe

Next Generation Sequencing e analisi di correlazione genotipo fenotipo **82**
Gentilini D., Oliveri A., Fazia T., Pini A., Marelli S., Di Blasio A.M., Spada E., Bernardinelli L.

Applicazione di un indicatore di dipendenza non lineare nell'analisi statistica di tracciati elettrocardiaci **86**
Di Rocco A, Vestri A.R., Nofroni I., Di Traglia M., Centurion Aznaran C., Pandozi C.

Big data, observational research, and p-value: a recipe for false positive findings? A study of simulated and real prospective cohorts **89**
Giovanni Veronesi, Guido Grassi, Giordano Savelli, Piero Quatto, Antonella Zambon

Bivariate spline models to assess the joint effect of intensity and duration of alcohol drinking and cancer of the oral cavity: a focus on a novel approach **91**
Di Credico Gioia, Polesel Jerry, Dal Maso Luigino, La Vecchia Carlo, Pauli Francesco, Torelli Nicola, Edefonti Valeria, the INHANCE consortium investigators

Relationship among hospital volume and mortality in patients affected by lung cancer: evidence from an Italian real-world investigation **100**
Ieva Francesca, Rea Federico, Barni Sandro, Merlino Luca, Apolone Giovanni, Corrao Giovanni

Evaluating the relationship between socioeconomic indicators and premature mortality in italian municipalities: a comparison of standard and new statistical approaches **103**
Alicandro Gianfranco, Bertuccio Paola, Carlo La Vecchia, Frova Luisa

The hartung, knapp, sidik jonkman adjustment for confidence intervals in presence of high eterogenity: an application to a meta-analysis on the association between renin-angiotensin system inhibitors and risk of alzheimer's disease **107**
Scotti Lorenza, Soranna Davide, Bilo Grzegorz, Parati Gianfranco, Zambon Antonella

La validazione di un modello di predizione lineare: principali metodi e proposta di un indice alternativo **111**
Signoriello Simona, Vittorio Simeon, Chiodini Paolo

Comparison of approaches to model sequential multiple mediators **116**
Zugna Daniela, Fasanelli Francesca, Popovic Maja, Richiardi Lorenzo

Sessione Giovani

The validation of a highly skewed surrogate for a survival endpoint: new approaches **119**
Capitoli Giulia, Rotolo Federico, Rebora Paola, Valsecchi Maria Grazia, Galimberti Stefania

Bayesian model averaging to build a comorbidity score using healthcare utilization databases **121**
Bucci Andrea, Faragalli Andrea, Ferrante Luigi, Skrami Edlira, Gesuita Rosaria, Dajko Marianxhela, Carle Flavia

A meta-analytic approach to estimate the proportion of long-term survivors in immune checkpoint inhibitors clinical trials **126**
Peveri Giulia, Oriecuia Chiara, Pagan Eleonora, Conforti Fabio, Bagnardi Vincenzo, Specchia Claudia

Statistical issues arising in systematic reviews and meta-analyses of immune checkpoint inhibitors clinical trials. An application to melanoma **129**
Pagan Eleonora, Oriecuia Chiara, Peveri Giulia, Conforti Fabio, Specchia Claudia, Bagnardi Vincenzo

Text mining e inexact matching per identificare ditte con esposizione ad amianto **132**
Airoldi Chiara, Ferrante Daniela, Miligi Lucia, Migliore Enrica, Tumino Rosario, Romanelli Antonio, Mensi Carolina, Cavone Domenica, Romeo Elisa, Marinaccio Alessandro, Elisabetta Chellini, Magnani Corrado

The integration of pre-coded information with narratives did not improve injury understanding: the case of accidental falls' surveillance **134**
Lorenzoni Giulia, Rampazzo Roberta, Buratin Alessia, Gregori Dario

Association between respiratory diseases in children and air pollution particulate matter: a multi-pollutants approach with random forest **136**
Bottigliengo D., Gallo E., Bressan S., Lanera C., Lorenzoni G., Ocagli H., Berchialla P., Baldi I., Gregori D.

Handling missing data in air pollution studies: a comparison of different approaches based on multivariate time-series models **140**
Gallo E., Bottigliengo D., Lanera C., Lorenzoni G., Ocagli H., Zagolin L., Marson G., Berchialla P., Baldi I., Gregori D.

Development and validation of a novel pre-pregnancy score predictive of preterm birth in primiparous women: a population-based cohort study **142**
Merlo Ivan, Cantarutti Anna, Merlino Luca, Massi Michela Carlotta, Ieva Francesca, Corrao Giovanni

La qualità dei percorsi di cura nei pazienti con disturbi mentali gravi in Italia. Il Progetto QUADIM **146**
Monzio Compagnoni Matteo, Franchi Matteo, Merlino Luca, Saponaro Alessio, Scondotto Salvatore, Gaddini Andrea, Di Fiandra Teresa, Lora Antonio, Corrao Giovanni

Visual function and brain lesions in preterm newborns: validation of the neonatal assessment visual european grid (NAVEG) **150**
Gnesi Marco, Malerba Laura, Rossi Andrea, Strobio Caterina, Fazzi Elisa, Montomoli Cristina

Associazione tra esposizione ad antidepressivi/depressione in gravidanza ed outcomes neonatali: "umbrella review" di revisioni sistematiche e meta-analisi **154**
Biffi Annalisa, Cantarutti Anna, Corrao Giovanni

Nervecheck: a cheap and accurate diagnostic tool for diabetic peripheral neuropathy. a single-center real-life experience **160**
Pia Clara Pafundi, Raffaele Galiero, Paul Valensi

Autism spectrum disorder and young adults social cognition networks comparison **163**
Vagnetti Roberto, Pino Chiara, Mazza M, Masedu Francesco, Valenti Marco

An analysis of the food and drug administration's adverse event reporting system for detecting signals of suicidal behavior in schizophrenic patients treated with second generation antipsychotics **170**
Azzolina Danila, Lorenzoni Giulia, Bottigliengo Daniele, Gregori Dario

Application and validation of a biological aging predictor based on machine learning as a marker of public-health in the italian population **172**
Gialluisi Alessandro, Di Castelnuovo Augusto, Costanzo Simona, De Curtis Amalia, Magnacca Sara, Bonaccio Marialaura, Cerletti Chiara, Donati Maria Benedetta, De Gaetano Giovanni, Iacoviello Licia on behalf of the Moli-sani Study Investigators

Esposizione a gastroprotettori nel primo anno di vita (inibitori di pompa protonica e antagonisti del recettore h2) e il successivo sviluppo di asma nei bambini **174**
Comoretto Rosanna Irene, Barbiellini Amidei Claudio, Zanier Loris, Canova Cristina

Discontinuità dei trattamenti antipsoriasici durante la gravidanza: studio di coorte pso-mother **176**
Francesca Romana Poggi, Valeria Belleudi, Serena Perna, Luigi Naldi, Renata Bortolus, Alessandro Cesare Rosa, Ursula Kirchmayer, Marina Davoli, Antonio Addis

Impact of adherence to standards of care for early rheumatoid arthritis on population health. results from the record linkage of rheumatic disease study of the italian society for rheumatology **178**
Zanetti Anna, Scirè Carlo Alberto, Guerrini Giulio, Carrara Greta, Zambon Antonella

Increasing risk of emergency room admissions for bronchiolitis in infants exposed to air pollution **183**
Gallo Elisa, Bottigliengo Daniele, Bressan Silvia, Lanera Corrado, Geremia Sara, Zagolin L., Marson G., Gregori Dario

A weighted quantile sum (wqs) regression with double index **186**
Renzetti Stefano, McRae Nia, Gennings Chris, Wright Robert O., Calza Stefano

Application of a cart analysis to stratify prognosis of grade II gliomas **190**
Battistella Claudio, Ius Tamara, Cesselli Daniela, Pegolo Enrico, Skrap Miran, Polesel Jerry, Isola Miriam

Impact of time-dependent medication adherence on heart failure patients using a joint modelling framework **193**
Spreafico Marta, Ieva Francesca

Trajectories in cognitively healthy italian elders **198**
Ferraro Ottavia Eleonora, Guaita Antonio, Villani Simona

A comparison of different statistical approaches to deal with model misspecification and missing outcome data **201**
Sciannameo Veronica, Fadini Gian Paolo, Bottigliengo Daniele, Avogaro Angelo, DARWIN-T2D Network, Baldi Ileana, Gregori Dario, Berchialla Paola

Oversampling and replacement strategies in propensity score matching: a critical review focused on small sample size in clinical settings **207**
Bottigliengo D., Baldi I., Lanera C., Bejko J., Bottio T., Tarzia V., Carrozzini M., Gerosa G., Berchialla P., Gregori D.

Bivariate copula method for age estimation in forced migrants **211**
Faragalli Andrea, Bucci Andrea, Skrami Edlira, Cameriere Roberto, Gesuita Rosaria, Dajko Marianxhela, Carle Flavia, Ferrante Luigi

Variable selection for multiple correlated exposures in healthcare research **216**
Pesenti Nicola, Zambon Antonella, Massimo Scacchi, Colicino Elena, Quatto Piero

Socioeconomic position and gastric cancer risk in the stomach cancer pooling (stop) project **220**
Rota Matteo, Alicandro Gianfranco, Pelucchi Claudio, Bonzi Rossella, Bertuccio Paola, Lunet Nuno, Boccia Stefania, Negri Eva, Ferraroni M., La Vecchia Carlo on behalf of the StoP Project study group

Association between dietary patterns and arterial stiffness in a contemporary adult italian population: the risk of cardiovascular diseases and abdominal aortic aneurysm in varese (ROCAV) study **224**
Lasalvia Paolo, Gianfagna Francesco, Tarallo Antonino, Tozzi Matteo, Iacoviello Licia, Grandi Anna Maria, Ferrario Marco Mario, Zambon Antonella, Veronesi Giovanni

Dietary flavonoids and lung function in copd patients **227**
Mattioli Veronica, Garcia-Aymerich Judith, Zanolin Maria Elisabetta, Benet Marta, Cazzoletti Lucia, Garcia-Larsen Vanessa

The relationship between cancer and dementia: association or causation? Results from swedish population based cohort studies **231**
Damiano Cecilia, Bellocco Rino, Rizzuto Debora, Grande Giulia

Overweight in childhood, puberty or early adulthood: changing the asthma risk in future generations? **233**
Lucia Calciano, Ane Johannessen, Marianne Lønnebotn, Bryndis Benediktsdóttir, Randi Jacobsen Bertelsen, Lennart Bråbäck, Shyamali Dharmage, Karl A. Franklin, Thorarinn Gislason, Mathias Holm, Christer Janson, Deborah Jarvis, Rain Jõgi, Jeong-Lim Kim, Jorunn Kirkeleit, Caroline Lodge, Andrei Malinovschi, Jesus Martinez-Moratalla, Roy Miodini Nilsen, Antonio Pereira-Vega, Francisco Gomez Real, Vivi Schlünssen, Cecilie Svanes, Simone Accordini

L'uso dei cortisonici per aerosol nei bambini in età prescolare: un'analisi dell'appropriatezza prescrittiva nella regione Lazio **236**
Alagna Michela, Lallo Adele, Reali Laura, Cortis Elisabetta, Di Martino Mirko, Davoli Marina

Sudden cardiac death in sport: study of incidence and evaluation of the effects of the Balduzzi law **240**
Di Rocco Arianna, Farcomeni Alessio, Alunni Fegatelli Danilo, Vestri Annarita

Levelling of cancer rates in the very old? **243**
Carioli Greta, Hashim Dana, Malvezzi Matteo, Bertuccio Paola, Waxman Samuel, Negri Eva, La Vecchia Carlo, Boffetta Paolo

Rischio di ospedalizzazione per tromboembolismo venoso ed inquinamento atmosferico: un'analisi case-crossover sugli effetti a breve termine del pm2.5 **249**
Matteo Renzi, Massimo Stafoggia, Paola Michelozzi, Marina Davoli, Francesco Forastiere, Angelo G. Solimini

Quale indice di variabilità della pressione arteriosa per misurare il rischio di mortalità cardiovascolare? **251**
Soranna Davide, Parati Gianfranco, Bilo Grzegorz, Palomba Andrea, Antonella Zambon

Design and feasibility of a multicentric study assessing blood microbiota and diet in relation to adenoma and colorectal cancer risk **256**
Marcello Cintolo, Roberto Penagini, Pietro Carnevali, Aldo Airoldi, Pietro Soru, Federica Mascaretti, Clorinda Ciafardini, Rossella Bonzi, Giovanni Ferrari, Carmelo Magistro, Barbara Oreggia, Marcello Vangeli, Maurizio Vecchi, Carlo La Vecchia, Simone Guglielmetti, Massimiliano Mutignani, Marta Rossi

The impact of job change on the cardiovascular health: evidence from the belgian idewe dataset **261**
Maniscalco L., Godderis L., Boon J., Schouteden M., Matranga D.

Poster

Detecting differences in "cytotec" search queries using the functional analysis of variance approach applied to google trends **265**
Porreca Annamaria, Scozzari Francesca, Di Nicola Marta

The effect of different weighting score adjustment in logistic regression models for binary correlated responses **271**
Marchioni Michele, Schips Luigi, Di Nicola Marta

Elevated levels of eef1a2 protein expression in triple negative breast cancer relate with poor prognosis **276**
Giudici Fabiola, Petracci Elisabetta, Nanni Oriana, Bottin Cristina, Pinamonti Maurizio, Barbati Giulia, Zanconati Fabrizio, Scaggiante Bruna

Stima dei trend temporali utilizzando database amministrativi: considerazioni metodologiche ed un'applicazione ai dati Inail relativi allo screening per uso di sostanze stupefacenti in provincia di Como e Varese **280**
Raimondi Jessica, Cavicchiolo Marco, Dashi Oriana, Calderini Duccio, Aiani Maria Rita, Ottaviani Carlo, Tosetto Claudio, Olandese Raffaela, Borchini Rossana, Ferrario Marco Mario, Veronesi Giovanni

Italy and Sweden: a description on cognitive trend among older adults **283**
Ferraro Ottavia Eleonora, Rizzuto Debora, Grande Giulia, Guaita Antonio, Villani Simona

A bayesian sample size estimation procedure based on a b-splines semiparametric elicitation method **286**
Danila Azzolina, Paola Berchialla, Dario Gregori, Ileana Bald

I modelli di diffusione nella previsione e controllo di patologie infettive **288**
Di Traglia M., Vestri A.R.

Accuratezza del test di cameriere per la stima della maggiore età: una revisione sistematica **292**
Dajko Marianxhela, De Micco Francesco, Rosaria Gesuita, Andrea Bucci, Andrea Faragalli, Edlira Skrami, Flavia Carle, Cameriere Roberto, Luigi Ferrante

Prediction of treatment outcome in clinical trials: a machine learning approach **295**
Berchialla Paola, Lanera Corrado, Gregori Dario, Baldi Ileana

Sindrome coronarica acuta: dati dal mondo reale in Lombardia **300**
Ronco Raffaella, Rea Federico, Merlino Luca, Corrao Giovanni

Analisi degli errori di I e II tipo in alcuni test di statistica circolare per il riconoscimento di trend ciclici **304**
Rocchi Ettore, Sisti Davide, Amatori Stefano, Rocchi Marco B.L.

Un approccio non parametrico all'analisi della varianza a due vie **307**
Borrelli Paola, Gigli Berzolari Francesca, Pistono Cristiana, Mallucci Giulia, Monti Maria Cristina

Daily patterns of carbohydrate intake in adults by non-parametric multi-level latent class analysis: an application to a representative sample of british adults from the Uk national diet and nutrition survey (2008/09 - 2015/16) **311**
Chaochen Wang, Suzana Almoosawi, Luigi Palla

Measuring multimorbidity inequality across Italy through the multisource comorbidity score: a nationwide study **315**
Rea Federico, Cantarutti Anna, Merlino Luca, Zanier Loris, Di Lenarda Andrea, De Palma Rossana, Belotti Laura, Francesconi Paolo, Profili Francesco, Forni Silvia, Carle Flavia, Bucci Andrea, Davoli Marina, Di Martino Mirko, Lallo Adele, Lepore Vito, Martini Nello, Dondi Letizia, Scondotto Salvatore, Pollina Addario Sebastiano, Fantaci Giovanna, Garau Donatella, Sechi Giuseppe, Lispi Lucia, Carbone Simona, Corrao Giovanni and on behalf of the working group 'Monitoring and assessing diagnostic therapeutic paths' of the Italian Health Ministry

Identification of gene expression in the brain causally linked to multiple sclerosis: a two-sample mendelian randomization approach **318**
Nova Andrea, Fazia Teresa, Gentilini Davide, Adaikalavan Ramasamy, Berzuini Carlo, Bernardinelli Luisa

Relationship between total protein intake and kidney function in adults from the cooperative health research in south tyrol (chris) study **322**
Vukovic Vladimir, Gögele Martin, Rainer Johannes, Garcia-Larsen Vanessa, Domingues Francisco, Pramstaller P Peter, Pattaro Cristian

Dietary non-enzymatic antioxidant capacity and the risk of cardiovascular diseases in a large cohort of Swedish women **324**
Essi Hantikainen, Marie Löf, Alessandra Grotta, Ylva Trolle Lagerros, Mauro Serafini, Rino Bellocco, Elisabete Weiderpass

Effect of depression and low-grade systemic inflammation on mortality and hospitalization risk: data from the Moli-sani study **326**
Gialluisi Alessandro, Costanzo Simona, Di Castelnuovo Augusto, Bonaccio Marialaura, Bracone Francesca, De Curtis Amalia, Cerletti Chiara, Donati Maria Benedetta, De Gaetano Giovanni, Iacoviello Licia, on behalf of the Moli-sani Study Investigators

Lo studio viadana III: inquinamento atmosferico e salute in età pediatrica **328**
Silvia Panunzi, Pierpaolo Marchetti, Chiara Badaloni, Kees de Hoogh, Linda Guarda, Francesca Locatelli, Vanda Pironi, Caterina Silocchi, Massimo Stafoggia, Paolo Ricci, Alessandro Marcon

Long term incidence of symptomatic cardiac events in 2,809 breast cancer patients treated with adjuvant trastuzumab in real-world practice **332**
Franchi Matteo, Di Cosimo Serena, Trama Annalisa, Merlo Ivan, Minicozzi Pamela, Tarantini Luigi, Merlino Luca, Apolone Giovanni, Corrao Giovanni

Development of a food knowledge questionnaire for tanzanian women of childbearing age **335**
Maria Vittoria Conti, Marco Gnesi, Maria Cristina Monti, Rachele De Giuseppe, Angela G. Mkindi, Naelijwa Mshanga, Hellas Cena

Use of health information flows to estimate the prevalence of Alzheimer's disease and other dementias in Sicily **338**
Maniscalco L., Fantaci G., Pollina Addario S., Scondotto S., Monastero R., Matranga D.

Total anatomical reconstruction during prostatectomy: risk of post-operative incontinence prediction **341**
Stura Ilaria, Manfredi Matteo, Checcucci Enrico, Garrou Diletta, Amparore Daniele, Fiori Cristian, Porpiglia Franco, AND Migliaretti Giuseppe

Consumo di statine e sopravvivenza in una coorte di pazienti con neoplasia della mammella **343**
Comoretto Rosanna Irene, Folino Franco, Zanier Loris, Gregori Dario, Canova Cristina

Vulnerability and social connectedness in young migrants: an analysis based on bayesian networks **346**
Charrier Lorena, Berchialla Paola, Borraccino Alberto, Lemma Patrizia, Sciannameo Veronica, Dalmasso Paola

Requiem per il "one sample one stage design (a'hern's design)" **348**
Zolin Anna, Tresoldi Claudio, Antonelli Paolo, Cesana Bruno Mario

Characterising ratings of 12 commonly consumed foods with respect to a set of desirability attributes in a survey on women's dietary habits in rural Uganda **355**
Luigi Palla, Andrea Spray, AnnaRita Vestri, Gwen Varley, Joweria Nambooze, Kate Wellard, Elaine Ferguson

Analisi bayesiana del rischio di ricovero per malformazioni congenite in puglia. anni 2001-2013 **358**
Bartolomeo Nicola, Mincuzzi Antonella, Leogrande Simona, Serio Gabriella, Minerba Sante, Trerotoli Paolo

Valutazione dell'inquinamento quale fattore di rischio nella cancerogenesi del sistema emolinfopoietico: esperienza in territori molisani e abruzzesi a rischio – risultati preliminari **363**
De Gregorio Daniela, Recchia Laura, Nofroni Italo, Alunni Fegatelli Danilo, de Michele Isabella, Vitullo Felice, Cantò Chiara, Storti Sergio, Vestri Anna Rita

Towards a hybrid human-machine emergency triage: predicting clinical evaluations from emergency department triage information through machine learning models in order to optimize access to care priority lists **365**
Milanese Alberto, Vestri Annarita

Studying the effect of a brief meditation training on psychological indicators of well-being **368**
Teresa Fazia, Francesco Bubbico, Giovanni Berzuini, Salvatore Bruno, Luisa Bernardinelli

Studio retrospettivo di una coorte di soggetti esposti all'amianto residenti in provincia di Pavia **370**
Silvia D. Visonà, Laura Gavarini, Simona Villani, Antonio Osculati

Exposure to cleaning agents at home or at work decreaseS levels of Fractional exhaled Nitric Oxide (FeNO). The GEIRD study **374**
Pierpaolo Marchetti, Lorena Torroni, Francesca Locatelli, Mario Olivieri, Giuseppe Verlato

Un elemento di verifica e di valutazione dell'apprendimento nei corsi di laurea magistrale in Medicina e Chirurgia: dal Progress Test al Training Test, elaborazione dati finali **376**
Laura Recchia, Bruno Moncharmont

Identificazione e intervento su bambini a rischio di disturbi del linguaggio e/o del neurosviluppo, figli di migranti: usabilità del PVB-BIL **380**
Simonelli Ilaria, Pasqualetti Patrizio, Rinaldi Pasquale, Cristiana Sanalitro, Maria Cristina Caselli, Lorena Remi, Paola Pettenati, Maria Elena La Marchina, Silvia Stefanini

Statistical, clinical and ethical issues for design and analysis of single-arm multi-stage phase II studies in anestesiology **382**
Cofini Vincenza, Fusco Pierfrancesco, Petruccci Emiliano, Di Carlo Stefano, Necozione Stefano

Caratterizzazione del profilo psico-sociologico degli studenti immatricolati ai corsi di laurea in Medicina e Chirurgia presso "Sapienza" Università di Roma. Risultati preliminari **386**
Nofroni Italo, Vestri Anna Rita, Ferrara Carla, Cattaruzza Mariasofia, Liuccio Michaela, Carpi Matteo

Treatment patterns and clinical outcomes in patients receiving palbociclib for hormone receptor–positive, human epidermal growth factor receptor 2–negative advanced or metastatic breast cancer: a real-world experience in a single italian institution **390**
Piezzo Michela, Milanese Alberto, Paccosi Sara, Simenon Vittorio, Tanturli Michele, Vestri Annarita, Riemma Maria, Cianniello Daniela, Caputo Roberta, Di Lauro Vincenzo, De Laurentiis Michele

Prevalenza di asma e bpco in due comuni della lomellina: uno studio trasversale **393**
Puci Mariangela Valentina, Gnesi Marco, Ferraro Ottavia Eleonora, Gigli Berzolari Francesca, Monti Maria Cristina, Borrelli Paola, Oddone Enrico, Zottarelli Fabio, Montomoli Cristina, Villani Simona



Atti | Epidemiologia

EVOLUZIONE DELLA MORTALITÀ PER TUTTE LE CAUSE NEI RESIDENTI DELLA PROVINCIA DI PAVIA DAL 1980 AL 2015

Borrelli Paola[1,2], Trivelli Leonardo[1], Gnesi Marco[1], Cadum Ennio[2], Villani Simona[1]

[1] *U.O. di Biostatistica e Epidemiologia Clinica, Dip.to Sanità Pubblica, Medicina Sperimentale e Forense, Università degli Studi di Pavia, Pavia*

[2] *U.O.C. Salute Ambiente e Progetti Innovativi, ATS Agenzia di Tutela della Salute di Pavia, Pavia*

Introduzione

La mortalità costituisce il più solido degli indicatori epidemiologici, poiché i dati sono disponibili a livello nazionale e internazionale in modo esaustivo, con dettaglio territoriale e su ampie serie storiche. La descrizione prima della mortalità generale e poi di quella cause specifiche, rappresenta pertanto un elemento rilevante per la conoscenza dello stato di salute di una popolazione e contribuisce a costruirne il "profilo di salute". In merito a questo indicatore, la Provincia Pavese presenta negli ultimi anni tassi di mortalità generali superiori (508,07 x 100.000) rispetto alla media regionale (450,00 x 100.000) e a quella nazionale (472,84 x 100.000), come pure un maggior ricorso alle strutture di ricovero ospedaliero rispetto al resto della Lombardia, e paragonabile alla media Italiana [1]. A supportare queste evidenze è anche la presenza nella Provincia di Pavia di un Sito di Interesse Nazionale (SIN) dovuto all'impianto produttivo Fibronit. Lo stretto legame tra esposizione all'amianto e carichi di malattia rende necessario procedere alla definizione di un profilo dello stato di salute della popolazione residente nell'area pavese per migliorare la conoscenza sulla mortalità.

Obiettivi

Stimare la distribuzione geografica della mortalità per tutte le cause in provincia di Pavia in tre periodi temporali (1980-1991,1992-2003, 2004-2015).

Metodi

È stato disegnato un studio ecologico in cui l'unità statistica è rappresentata da ciascuno dei 188 comuni della Provincia Pavese. L'outcome primario è rappresentato dalla mortalità per tutte le cause, escluse quelle traumatiche ed avvelenamenti (codici 800-999 della ICD-9 CM) [2]. I dati di morte per comune della provincia di Pavia sono stati forniti dalla banca dati ISTAT, così come quelli relativi alla popolazione residente a livello comunale nei periodi indagati. Per il calcolo della standardizzazione indiretta è stata utilizzata come popolazione di riferimento quella provinciale ISTAT al 1° Gennaio 2012.
La distribuzione geografica della mortalità, a livello comunale, è stata modellizzata con un approccio bayesiano di tipo gerarchico che permette di filtrare il 'rumore' di fondo, in modo che il rischio di mortalità nei comuni dove non si verificano eventi sia bilanciato da quello dei comuni in cui accadono. L'approccio bayesiano, utilizzando la distribuzione *a priori*, valorizza la tendenza dei tassi specifici per aree limitrofe di non differire troppo, ottenendo così uno *smussamento* delle stime del rischio di mortalità attraverso mappe in grado di fornire un'immagine più chiara della vera variazione geografica [3-7]. Il modello utilizzato è una regressione di Poisson che stima i rischi relativi (RR) di ogni unità areale tenendo conto delle informazioni dalle aree limitrofe (correlazione spaziale). La distribuzione finale degli RR combina l'informazione del modello *a priori* con quella contenuta nei dati e fornisce una distribuzione *a posteriori*, qui calcolata usando l'algoritmo computazionale *Integrated Nested Laplace Approximation* (INLA) [8]. Nello specifico, è stato determinato il valore di RR mediano globale della provincia che fornisce l'RR di riferimento e successivamente è stata calcolata la probabilità a posteriori (PP) per ogni i-esimo RR di essere superiore

al valore mediano (PP= p(RRi>M): le probabilità a posteriori (PP) possono essere viste come l'analogo bayesiano dei valori di p-value. Una PP per un comune superiore a 0,90 indica che l'RR di quel comune è significativamente superiore a quello mediano provinciale, mentre una PP minore di 0,10 sta a indicare che nell'area esaminata l'RR è significativamente inferiore rispetto a quello di riferimento.
I risultati ottenuti dalle analisi sono stati mappati utilizzando scale di colore crescenti per classi crescenti di RR e di PP. Le analisi sono state condotte con R, versione 3.4.2 (pacchetti INLA 3.5 e *Leaflet*).

Risultati

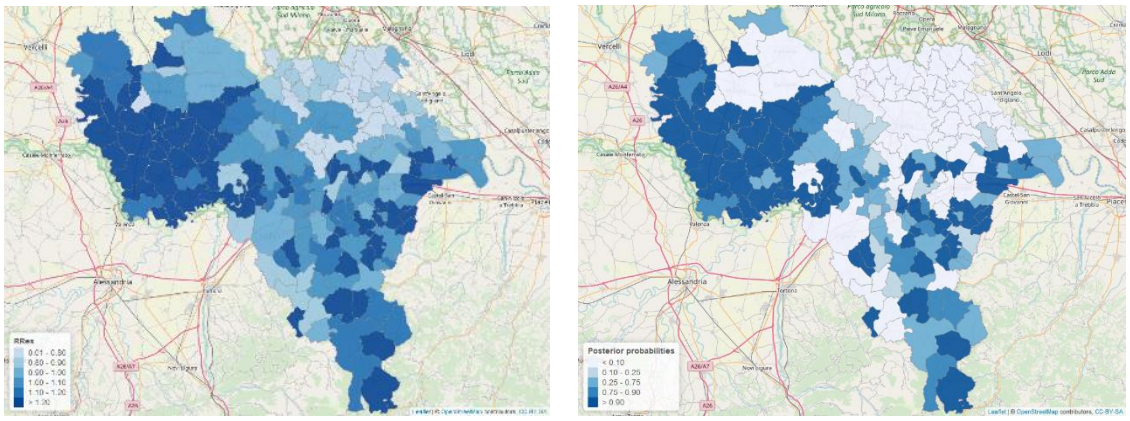
La distribuzione spaziale degli RR mostra un'evoluzione positiva della mortalità per tutte le cause nell'intera Provincia di Pavia (figura 1). Per il comune di Broni invece, Sito di Interesse Nazionale (SIN) per la presenza di un impianto di produzione di amianto come quello di Casale Monferrato, e per il suo hinterland si evince dalle mappe un peggioramento nel tempo della distribuzione spaziale degli RR. Precisamente, nel periodo 1980-1991 Broni presentava un RR di 1,00 (1711 casi) statisticamente inferiore rispetto l'RR mediano provinciale (1,08) mentre in quello successivo 1992-2003 era di 1,11 (1812 casi) statisticamente superiore a quello mediano del periodo pari a 1,05. Situazione simile anche per il periodo 2004-2015 (RR=1,08>1.01 [1702 casi]). Analogamente anche per il comune limitrofo di Campospinoso dove da valori di RR inferiori a quelli mediani provinciali nel tempo si è passati a valori significativamente superiori (1980-1991: RR=0,86 [174 casi]; 1992-2003: RR=1,12 [265 casi]; 2004-2015: RR=1,40 [367 casi]). Per gli altri comuni confinanti, come San Cipriano Po (RR=1,33 – 98 casi), Albaredo Arnaboldi (RR=1,42 – 65 casi), Pietra dè Giorgi (RR=1,22 – 221 casi) e Canneto Pavese (RR=1,24 – 323 casi) si evidenzia nel periodo 1980-1991 degli RR con valori significativamente (PP>0,90) maggiori rispetto a quello mediano provinciale, poi diminuiti negli altri due periodi temporali. I comuni di Barbianello e Redavalle rispecchiano lo stesso andamento. Il comune di Stradella e quello di Cigognola, invece mostrano per ogni periodo temporale degli RR statisticamente inferiori rispetto ai mediani provinciali (PP<0,10).

Conclusioni

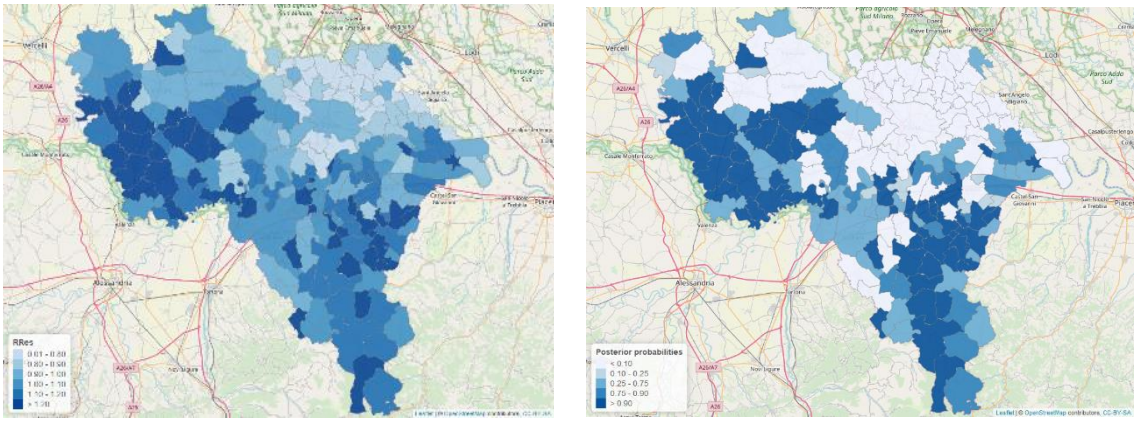
Il territorio pavese mostra, per quanto concerne la mortalità generale, l'esistenza di cluster di mortalità con RR significativamente maggiori di quelli mediani provinciali, così come inferiori. Più precisamente nel primo caso, il cluster di una più alta mortalità si trova nell'area della Lomellina, nel secondo caso in quella del Pavese (sotto rischio). Nell'Oltrepò si hanno indicazioni di cluster a bassa ed elevata mortalità.
I risultati suggeriscono di indagare la distribuzione spaziale di alcune patologie specifiche, verosimilmente correlabili alle fonti di esposizione presenti sul territorio. Nel caso specifico dell'Oltrepò Pavese si proseguirà con uno studio incentrato sulle patologie asbesto correlate (quali mesotelioma, tumore al polmone e alla pleura, ovaio, stomaco, laringe e faringe e colon) che consentirebbe di indagare sia l'aspetto epidemiologico di impatto delle fonti di rischio presenti nell'area, quali la Fibronit di Broni, che programmare successivi interventi di sorveglianza sanitaria sulla popolazione.

Figura 1. *Analisi della variazione geografica del rischio relativo (RR) bayesiano di mortalità per tutte le cause nei periodi in studio (1980-1991, 1992-2003, 2004-2015).*

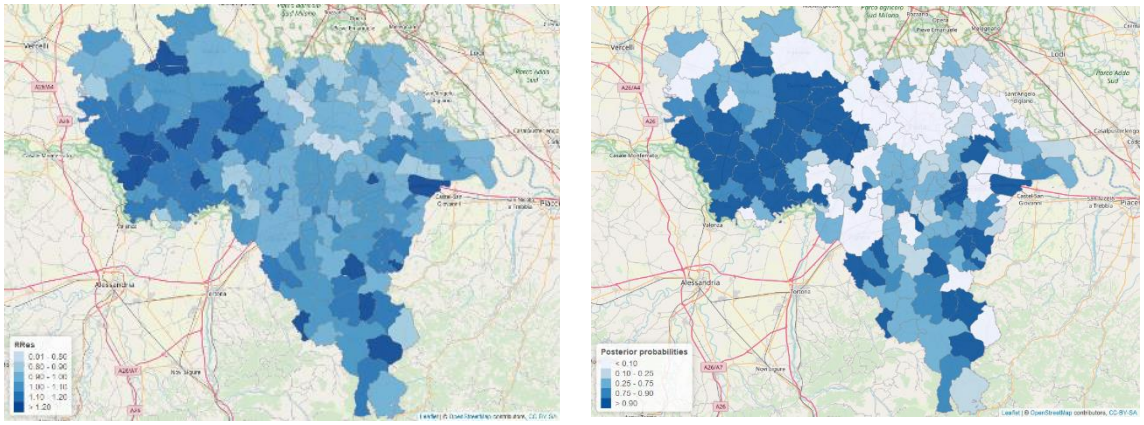
1980-1991



1992-2003



2004-2015



Bibliografia

[1] Piano Integrato Locale di Promozione della Salute - Anno 2019, ATS di Pavia, Link https://www.ats-pavia.it/sito.nsf/Organizzazione_Documenti.xsp/Documenti/-Piano-IntegratoLocale-di-Promozione-della-Salute-Anno-2019 (consultato il 24 giugno 2019).

[2] Classificazione Internazionale delle Malattie (ICD-9-CM).

[3] Claylon D, Kaldor J. Empirical Bayes estimates of age-standardizated relative risks for use in disease mapping. Biometrics 1987;, 43: 671-681.

[4] Besag J, York J, Mollie A. Bayesian image restoration with applications in spatial statistics (with discussion). Ann Inst Stat Math 1991;43: 1-59.

[5] Bernardinelli L, Montomoli C. Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. Stat Med 1992, Jun 15;11(8):983-1007.

[6] Clayton D, Bernardinelli L. Bayesian methods for mapping disease risk. In Cuzick J., Elliot P. and Stern R. (eds), Geographical and environmental epidemiology: Methods for small-area studies. Oxford: Oxford University Press, 1992: 205-220

[7] Clayton D, Hills M. Statistical Models in Epidemiology. Oxford University Press, Oxford 1993. No. of Pages: vii + 367. ISBN: 0-19-852221-5.

[8] Rue H, Martino S, Chopin N. (2009). Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with discussion), Journal of the Royal Statistical Society B, 71, 319-392.

GEOGRAPHIC HETEROGENEITY IN NON-COMMUNICABLE DISEASES IN RELATION TO QUALITY OF WORK IN EUROPE: EVIDENCE FROM THE SHARELIFE SURVEY

Domenica Matranga[1], Filippa Bono[2]

[1] *Dipartimento di Promozione della Salute, Materno-Infantile, Medicina Interna e Specialistica di eccellenza "G. D'Alessandro", Università degli Studi di Palermo*

[2] *Dipartimento di Scienze Economiche, Aziendali e Statistiche, Università degli Studi di Palermo*

Introduction

Adverse employment and working conditions are associated with worse health [1] and contribute to the social gradient in health [2]. There is a body of evidence that quality of work is associated with adverse health outcomes, including poor self-rated health, depression [3] and coronary heart disease [4]. Most of the studies concern small workers' populations [5] or unique general populations [6-7]. Little is known about the relation between health outcomes and quality of work, in case of individuals belonging to different populations.

Aim

The scope of this work is to investigate the relation between the number of non-communicable diseases and quality of work, using retrospective European data from the SHARELIFE Survey, 2017.

Methods

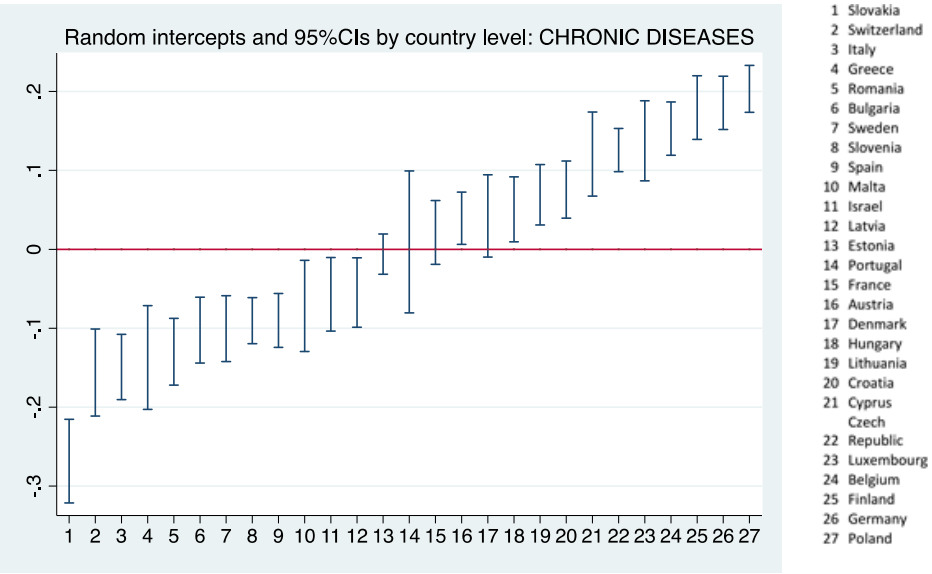
Data are drawn from the Survey of Health, Ageing and Retirement in Europe, which is a panel database of microdata on health, socioeconomic status and social and family networks of people aged 50 years and over, covering most of the European Union [8]. We combined data from the seventh wave (2017) of the Survey of Health, Ageing and Retirement in Europe (SHARE) with retrospective information on participants' life course from the third wave (2008-2009) (also called SHARELIFE), which includes details on previous employment histories and childhood conditions. We obtained information about twenty-nine European countries for a total of 75,701 people.
A multilevel negative binomial regression model was estimated to assess the number of chronic diseases related to the main job characteristics (sector of economic activity, job function) and quality of work (stressful work conditions, job satisfaction), adjusting for socio-demographic variables (gender, age, education, living in couple, current job status), behavioral risk factors for chronic diseases (BMI) and deprivation during early childhood (synthetic indicator of i) number of books at home at respondent's 10 years old, ii) housing crowding, iii) housing quality, iv) isco code of the breadwinner at respondent's 10 years old). Work quality was measured through two indicators: a binary indicator to assess satisfaction with the main job in the entire life and an 11-items indicator to measure the degree of adversity experienced in the main, according to the demand-control support model and the effort-reward imbalance model ([9]). To the aim of this analysis, we considered individuals nested within countries. The estimated model included the main effects and the first-order interaction between age and current job status. We chose 0.05 as cut-off level for statistical significance.

Results

The number of chronic diseases increased in the rate of 2% for one unit increase of age (IRR=1.02; 95%CI=[1.02-1.02]) and of 3% for one unit increase of bmi (IRR=1.03; 95%CI=[1.03-1.03]). It was higher for

women (IRR=1.09; 95%CI=[1.08-1.11]) and lower for high-educated people (IRR=0.95; 95%CI=[0.93-0.97] for upper secondary school and IRR=0.94; 95%CI=[0.91-0.97]). There was a significant interaction between age and current job status: compared to retired people, the number of chronic diseases was significantly lower for permanently sick people (IRR=0.99; 95%CI=[0.99-1.00]) and significantly lower for homemakers (IRR=1.01; 95%CI=[1.00; 1.01]).
Compared to people employed in Agriculture and fishery, there was an increasing risk to be affected by NCDs for those working in Manufacturing (IRR=1.05; 95%CI=[1.02-1.08]), Electricity, gas and water supply (IRR=1.08; 95%CI=[1.02-1.14]), Construction (IRR=1.06; 95%CI=[1.02-1.10]), Hotels and restaurants (IRR=1.12; 95%CI=[1.07-1.18]), Health and social work (IRR=1.05; 95%CI=[1.01-1.09]) and Other Community (IRR=1.04; 95%CI=[1.01-1.08]). Compared to employee (private sector), people working as civil servant (public sector) suffered more NCDs (IRR=1.05; 95%CI=[1.02-1.07]), while there was any statistically significant difference for self-employed.
Quality of work was a significant predictor of the number of chronic diseases only limited to the overall satisfaction with his/her own main job. Being unsatisfied was associated with a major risk of being affected by a number of NCDs (IRR=1.04; 95%CI=[1.02-1.05]), while the degree of adversity experienced in the main job was not a significant predictor of the number of chronic diseases (p=0.056).
Unobserved heterogeneity was statistically significant for the respondent's living country (variance component=0.02; 95%CI=[0.01; 0.03]). There were extremely statistically significant positive residuals for Belgium (u_j=0.15; 95%CI=[0.12; 0.19]), Finland (u_j=0.18; 95%CI=[0.14; 0.22]), Germany (u_j=0.19; 95%CI=[0.15; 0.22]) and Poland (u_j=0.20; 95%CI=[0.17; 0.23]) and extremely statistically significant negative residuals for Slovakia (u_j=-0.27; 95%CI=[-0.32; -0.22]), Switzerland (u_j=-0.16; 95%CI=[-0.21; -0.10]) and Italy (u_j=-0.15; 95%CI=[-0.19; -.11]) (see Figure)

Figure. *Variance components negative binomial regression model of the number of chronic diseases related to characteristics and quality of the main job, adjusted for socio-demographic variables, behavioral risk factors and deprivation during early childhood: country-level random residuals*



Conclusions

Job characteristics and quality of work were found as significant determinants of non-communicable diseases, with people enrolled in the sectors of Hotels and restaurants and Electricity, gas and water supply at major risk of being diseased. There was unexplained heterogeneity among countries, whose causes can be searched for at a regulatory, geographical and economic level.

References

[1] Marmot M, Bell R. 2010. Challenging health inequalities—implications for the workplace. Occupational Medicine 2010; 60:162–166.

[2] Borg V1, Kristensen TS. 2000. Social class and self-rated health: can the gradient be explained by differences in life style or work environment? Soc Sci Med. Oct;51(7):1019-30.

[3] Rugulies R1, Aust B, Madsen IE. 2017. Effort-reward imbalance at work and risk of depressive disorders. A systematic review and meta-analysis of prospective cohort studies. Scand J Work Environ Health. 2017 Jul 1;43(4):294-306.

[4] Dragano N1, Siegrist J, Nyberg ST, et al., 2017. Effort-Reward Imbalance at Work and Incident Coronary Heart Disease: A Multicohort Study of 90,164 Individuals. Epidemiology. 2017 Jul;28(4):619-626.

[5] Hodge B1, Wright B1, Bennett P1. Balancing Effort and Rewards at University: Implications for Physical Health, Mental Health, and Academic Outcomes. Psychol Rep. 2019 May 6:33294119841845.

[6] Åhlin JK1, LaMontagne AD2, Magnusson Hanson LL1. 2019. Are there bidirectional relationships between psychosocial work characteristics and depressive symptoms? A fixed effects analysis of Swedish national panel survey data. Occup Environ Med. 2019 May 4. pii: oemed-2018-105450.

[7] Myllyntausta S1,2, Salo P3,4, Kronholm E et al., 2019. Does Removal of Work Stress Explain Improved Sleep Following Retirement? The Finnish Retirement and Aging (FIREA) study. Sleep. May 7. pii: zsz109.

[8] Börsch-Supan A, Brandt M, Hunkler C, Kneip T, Korbmacher J, Malter F, Schaan B, Stuck S, Zuber S (2013) SHARE central coordination team. data resource profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). Int J Epidemiol 42:992–1001.

[9] Wahrendorf M., Siegrist J. 2014. Proximal and distal determinants of stressful work: framework and analysis of retrospective European data. BMC Public Health, 14:849

VARIATION OF MULTIPLE SCLEROSIS RISK IN THE PROVINCE OF PAVIA: INVESTIGATING PM2.5 EXPOSURE

Monti Maria Cristina[1], Trivelli Leonardo[1], Gerosa Leonardo[2], Pisoni Enrico[3], Borrelli Paola[1], Mallucci Giulia[2], Bergamaschi Roberto[2], Montomoli Cristina[1]

[1] *Dipartimento di Sanità Pubblica, Medicina Sperimentale e Forense - Unità di Biostatistica – Università degli Studi di Pavia,*

[2] *Multiple Sclerosis Centre, IRCCS Mondino Foundation, Pavia, Italy,*

[3] *European Commission, Joint Research Centre (JRC), Directorate for Energy, Transport and Climate, Air and Climate Unit, Via E. Fermi 2749, I-21027, Ispra, VA, Italy.*

Backgroud

Several studies identified Italy as a multiple sclerosis (MS) high risk zone, with a prevalence ranging from 110 to 330 per 100,000 inhabitants in the last decade [1]. Relatively few studies have explored the microscopic geographic variation in the distribution of MS, trying to unravel the multifactorial nature of MS risk and to solve out some of the statistical issues related to the fact that the disease is uncommon and distributed in small areas with relatively scarce population. Bayesian methodologies [2-4] had been used by our research group to provide more reliable estimates of the geographical distribution of MS. In the province of Pavia, Northern Italy, in the year 2000, we estimated an overall prevalence of 94 per 100,000 inhabitants and found evidence that the prevalence was greater in some areas than the median provincial rate [5]. Moreover, the same methodology was applied in south-west Sardinia [6] leading to the identification of a high-prevalence MS cluster. Clustering was more evident in the upper part of the region, where lead mines and steel industries are present. Although its etiology remains unknown, MS is believed to be a consequence of the interplay between environmental exposure and individual genetic susceptibility [7]. Hence, identification of spatial clusters of MS may suggest novel patogenetic hypotheses for the disease. In recent years, environmental air pollution has been associated with MS risk [8]. Moreover seasonal or temporal variation between MS relapses, possibly due to fluctuations in levels of inhalable particles, especially particulate matter (PM), has been observed [9-10]. Inhalable pollutants are known to cause inflammation at the level of the respiratory tract epithelium, possibly triggering the same immune responses involved in MS pathophysiology.

Aim

The aim of the study was to assess the spatial distribution of MS risk across the administrative areas of the Province of Pavia (northern Italy) taking into account air pollution distribution (considering PM2.5 concentrations) and adjusting for age and urbanization level, applying a Bayesian focused cluster-detection procedure.

Methods

The province of Pavia is one of the 11 provinces in the region of Lombardy, subdivided into 188 municipalities and has 547,251 inhabitants according to the 2017 inter-census Istat survey (http://demo.istat.it/pop2017/index4.html).
Cases were ascertained using the data-base of the MS Center of Pavia [11] as the primary source and were also drawn from the Data Warehouse of the Pavia Health Protection Agency (HPA), specifically searching for dispensation records of MS patients not included in the Pavia MS Center (specific code: 46.340). The

mean delay of MS case reporting was taken into account by setting the prevalence day on 31 December, 2016 [12].
Spatial data regarding PM2.5 concentrations (that is to say, concentrations of particulate matter with a diameter of 2.5 micrometers or less) were gathered from the "Open Source EMEP/MSC-W" model simulation results [13], as part of the database and tools developed in the frame of the UNECE (United Nation Economic Commission for Europe) "Long-range Transboundary Air Pollution (CLRTAP)" convention. The model simulation results (available at https://emep.int/mscw/mscw_moddata.html) are constituted by gridded concentrations of several air pollutants in 0.1° x 0.1° (long-lat) resolution from year 2000 onwards. According to the purpose of the study, gridded data from the Pavia province were extracted for the timeseries from 2010 to 2017. The grid was further subdivided into the 188 municipalities of the Pavia province; for each municipality, average seasonal (winter) concentrations of PM2.5 were made available and subdivided into tertiles (first tertile:6.74-25, second tertile:25.1-31.9, third tertile: 32-46.6 μg/m^3).
In order to stratify territories according to their respective level of urbanisation, the European degree of urbanisation (DEGURBA) (https://circabc.europa.eu/d/d/workspace/SpacesStore/b32fcc21-c566-45ea-8795-2a9ffe2ac58c/DGUR_ 2011_Tables_Maps_Countries.zip) classification was adopted. The classification of Local Administrative Units (LAUs) is based on three categories: 1-Cities (densely populated areas), 2-Towns and suburbs (intermediate density areas), 3-Rural areas (thinly populated areas). Due to the scarcity of territories classifiable as Cities (the Pavia and Vigevano areas only), we re-classified LAUs into Urban (category 1+2) *vs* Rural.
Crude age-specific prevalence rates were calculated by counting the number of MS patients that were alive the prevalence day and using, as denominator, the 2017 population of the province of Pavia.
Cases were assigned the ecological characteristics of the geographic unit (municipality) to which they were resident; the PM2.5 concentrations and urbanization matrix were extracted at municipality level and associated to the MS cases.
Using the age-specific prevalences of the whole province of Pavia at December 31, 2016, the expected counts (internal standardization) were calculated for each municipality. To provide a more precise estimate of the relative risk (RR) for each statistical area, we fit Bayesian log-linear Poisson regression models with Besag, York and Mollié random effects (BYM) to capture the over-dispersion [4], using informative prior beliefs based on the previous risk estimation conducted in the year 2000 [5].
The model may be written as follow:

$$log\theta_i = \mu + X'_i\beta + \eta_i + \phi_i$$

where θ_i is defined by the relative risk (RR) in area *i* and μ denotes the overall mean RR across the entire study region. We fit a model without any predictor (BYM) as well as a model (BYM_$PM_{2.5}$) where X_i are the categorical covariates PM2.5 and urbanization level with corresponding regression coefficient given by β and credible intervals. Covariates reference categories were the PM2.5 first tertile and the rural areas, respectively.
We defined η_i as the specified spatial random effects, and Φ_i denoted the unspecified features of area *i*, which do not incorporate spatial structure.
To assess posterior distributions of parameters and RR for each area, an Integrated Nested Laplace Approximation (INLA) approach was used [14]. To evaluate the area-specific RR posterior distribution, Posterior Probabilities (PP) were calculated by extracting 100,000 samples from the RR marginal posterior distribution and calculating, then, the relative frequency of samples where RR>1 for each area.
The posterior probabilities can assess our certainty defining an area as an "Area at high risk", that is an area with a higher RR estimate in reference to all municipalities of the province of Pavia more than 90% of the times (PP>0.90); likewise, an "Area at low risk" will have a higher RR estimates less than 10% of the times (PP<0.10).

Calculations were carried out using the statistical software Stata 15 and R version 3.4.2 (packages INLA 3.5 and *Leaflet*).

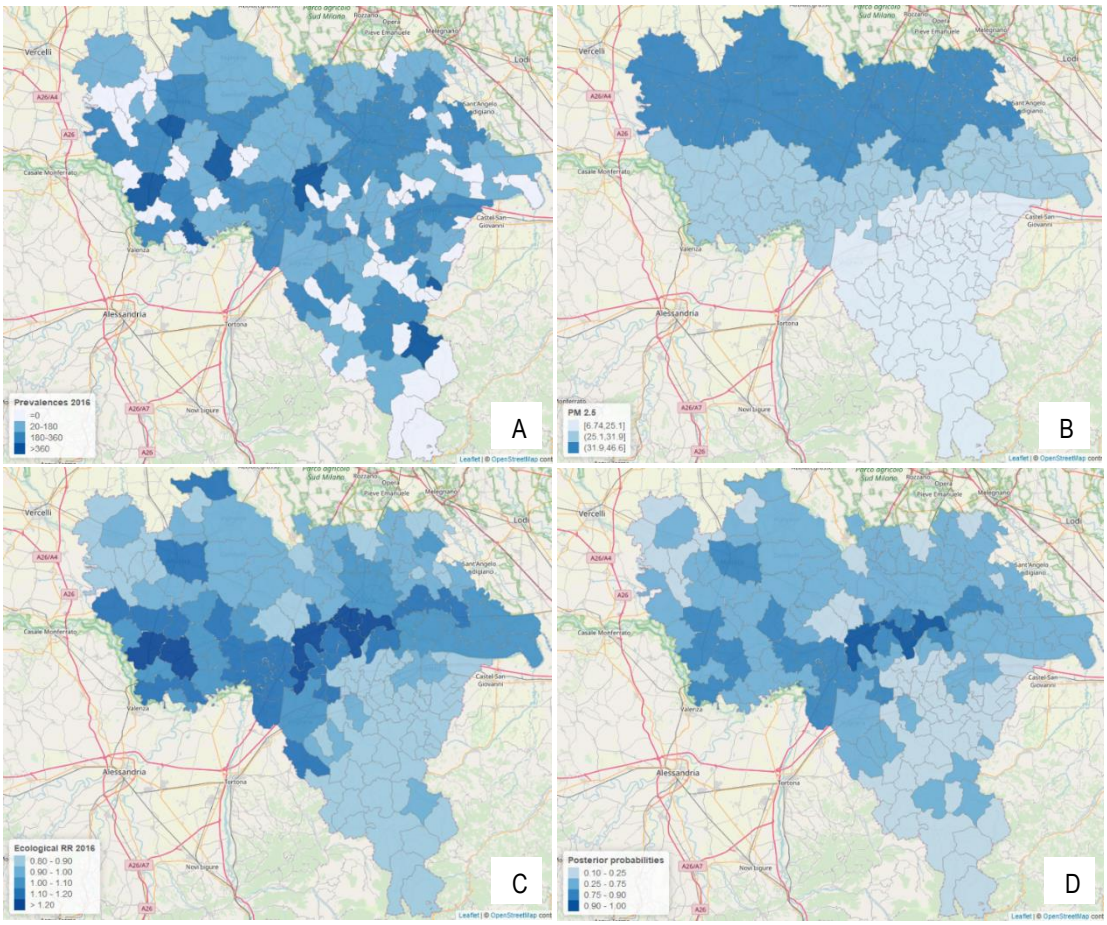
Results

We identified 927 MS-affected individuals who were alive and residents in the province of Pavia on December 31, 2016; the overall prevalence was 169.4 per 100,000 (95%CI:158.8-180.6).
Figure 1A shows the crude MS prevalence for each municipality. The map was heterogeneous, presenting a mixture of different values, probably due to random variability and differentially populated areas.
Figure 1B shows the map of PM2.5 concentration: the province was essentially subdivided into three stripes. Considering the effect of air pollution (model BYM_$PM_{2.5}$ described in the methods section), the MS risk was highest among those living in areas with PM2.5 concentration in the second tertile (RR=1.31, 95%CI:0.98-1.71, PP=0.97), adjusting for urbanization level. In areas with PM2.5 concentration in the third tertile (between 32 and 47 µg/m^3), the risk decreased to 1.10 (95%CI:0.82-1.46, PP=0.72). Urban areas appeared to have a higher MS risk than rural areas (RR=1.12, 95%CI:0.93-1.33, PP=0.88) controlling for PM2.5 concentration; credible intervals and PP were almost reaching significance level.
Figure 1C and 1D show the estimate of the ecological regression (BYM_$PM_{2.5}$) in terms of RR distribution and PP, respectively, for each municipality. Specifically, the PP map revealed consistent MS high risk clusters: not far from Pavia city, there were 3 adjacent municipalities where the RR was at least 20% higher and PP values >0.90; moreover, two adjacent municipalities facing Piedmont showed RR at least 20% bigger with PP values between 0.75 and 0.90. The selected areas were urban areas with an averaged PM2.5 concentration during winter season over 25 µg/m^3.

Conclusions

The study detected clustering of municipalities where an excess number of MS cases were identified taking into account PM2.5 and urbanization levels. Identification of such clusters may generate new hypotheses about environmental risk factors, which must be verified by appropriate study design.

Figure 1. *A: map of the crude MS prevalence per 100,000 inhabitants at December 31, 2016. B: PM2.5 concentration subdivided into tertiles. C: MS relative risks (RR). D: Posterior Probability (PP); a PP value higher than or equal to 0.90 indicated that the area-specific risk was consistently higher than the reference value (area at high risk), while a PP value lower than or equal to 0.10 indicated that the risk was consistently lower than the reference value (area at low risk).*



References

[1] Atlas of MS 2019, MSIF, 2019

[2] Clayton DG, Bernardinelli L, Montomoli C. Spatial correlation in ecological analysis. Int J Epidemiol. 1993 Dec;22(6):1193-202.

[3] Torabi M, Green C, Yu N, Marrie RA. Application of three focused cluster detection methods to study geographic variation in the incidence of multiple sclerosis in Manitoba, Canada. Neuroepidemiology.2014;43(1):38-48.

[4] A.B. Lawson. Hierarchical Modeling in Spatial Epidemiology. Bayesian Disease Mapping. Third Edition. Chapman and Hall/CRC Press, 2018.

[5] Bergamaschi R, Montomoli C, Candeloro E, PREMS (Pavia Register of multiple sclerosis) group et al (2006) Bayesian mapping of multiple sclerosis prevalence in the province of Pavia, northern Italy. J Neurol Sci 244:127–131.

[6] Cocco E, Sardu C, Massa R, Mamusa E, Musu L, Ferrigno P, Melis M, Montomoli C, Ferretti V, Coghe G, Fenu G, Frau J, Lorefice L, Carboni N, Contu P, Marrosu MG. Epidemiology of multiple sclerosis in south-western Sardinia. Mult Scler. 2011 Nov;17(11):1282-9.

[7] Marrie RA: Environmental risk factors in multiple sclerosis aetiology. Lancet Neurol 2004; 3: 709–718.

[8] Tateo F, Grassivaro F, Ermani M, Puthenparampil M, G. P. PM2.5 levels strongly associate with multiple sclerosis prevalence in the Province of Padua, Veneto Region, North-East Italy. Mult. Scler *J*. 1–9 (2018).

[9] Esmaeil Mousavi, S., Heydarpour, P., Reis, J., Amiri, M. & Sahraian, M. A. Multiple sclerosis and air pollution exposure: Mechanisms toward brain autoimmunity. Med. Hypotheses 100, 23–30 (2017).

[10] Bergamaschi, R. et al. Air pollution is associated to the multiple sclerosis inflammatory activity as measured by brain MRI. Mult. Scler. J. 24, 1578–1584 (2018).

[11] Trojano M, Bergamaschi R, Amato MP, Comi G, Ghezzi A, Lepore V, Marrosu MG, Mosconi P, Patti F, Ponzio M, Zaratin P, Battaglia MA; Italian Multiple Sclerosis Register Centers Group. The Italian multiple sclerosis register. Neurol Sci. 2019. Jan;40(1):155-165.

[12] Esbjerg S, Keiding N, Koch-Henriksen N. Reporting delay and corrected incidence of multiple sclerosis. Statistics in Medicine 1999; 18:1691-1706.

[13] Simpson D., et al. The EMEP MSC-W chemical transport model -- technical description. Atmospheric Chemistry and Physics, 2012, 12, 7825-7865.

[14] H. Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). Journal of the Royal Statistical Society, Series B, 71(2):319-392, 2009. R package version 3.5; 2018 Accessed: www.r-inla.org.

ADHERENCE TO THE MEDITERRANEAN DIET AND STOMACH CANCER RISK: RESULTS FROM THE STOMACH CANCER POOLING PROJECT

Bravi Francesca[1], Di Maso Matteo[1], Negri Eva[2], Palli Domenico[3], Lagiou Pagona[4,5], Lunet Nuno[6,7], Aragonés Nuria[8,9], Vioque Jesus[9,10], López-Carrillo Lizbeth[11], La Vecchia Carlo[1], Ferraroni Monica[1] on behalf of the StoP Project study group.

[1] *Dipartimento di Scienze Cliniche e di Comunità, Università degli Studi di Milano, Milano, Italia.*

[2] *Dipartimento di Scienze Biomediche e Cliniche, Università degli Studi di Milano, Milano, Italia.*

[3] *S.C. Epidemiologia dei Fattori di Rischio e degli Stili di Vita, Istituto per lo Studio, la Prevenzione e la rete Oncologica, Firenze, Italia.*

[4] *Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece.*

[5] *Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.*

[6] *Departamento de Ciências da Saúde Pública e Forenses e Educação Médica, Faculdade de Medicina da Universidade do Porto, Porto, Portugal.*

[7] *EPIUnit – Instituto de Saúde Pública, Universidade do Porto, Porto, Portugal.*

[8] *Department of Health of Madrid, Epidemiology Section, Public Health Division.*

[9] *CIBER Epidemiología y Salud Pública, Madrid, Spain.*

[10] *Department of Public Health, Miguel Hernandez University, FISABIO-ISABIAL, Alicante, Spain.*

[11] *Mexico National Institute of Public Health, Morelos, Mexico.*

Introduction

Stomach cancer is the fifth most frequent cancer and the third most common cancer-related cause of death worldwide [1]. Major recognized risk factors for this neoplasm include obesity, *Helicobacter pylori* infection, tobacco smoking, and alcohol drinking. In addition, some aspects of diet, including foods preserved by salting and meat consumption, have been suggested to be related to an increase in stomach cancer risk.
Mediterranean diet is a broad definition encompassing dietary habits typical of the populations from areas bordering the Mediterranean Sea, which are generally characterized by high intakes of fruit, vegetables, cereals, legumes, and fish, limited consumptions of meat, and dairy products, a moderate consumption of alcoholic drinks, and the use of olive oil. The Mediterranean diet has been investigated using various *a priori* scores and it has been inversely associated to different health outcomes, such as cardio and cerebrovascular diseases, and various types of cancers, including in particular those of the digestive tract [2, 3].

Aims

The aim of the present work is to investigate and quantify the strength of the association between the adherence to the typical Mediterranean diet and stomach cancer risk, through the analysis of pooled data from the stomach cancer pooling project (StoP) consortium, using different scores [4]. Moreover, we would

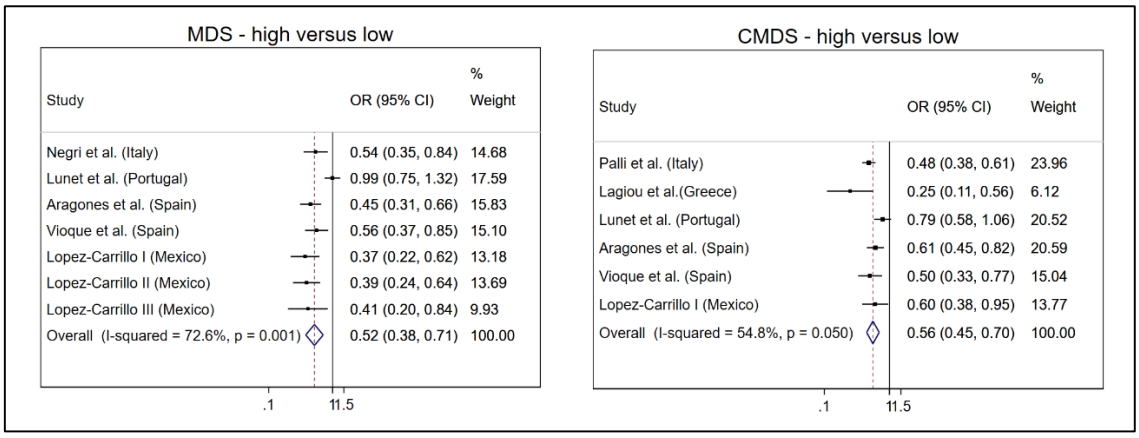
like to verify if the association is consistent across geographic area, comparing in particular countries within and outside the Mediterranean area.

Methods

The present work is based on the pooled data of the StoP consortium, including case-control and nested case-control studies, from several areas of the world. Participating studies provided complete original data, that were collected and harmonized by the coordinating center according to a standardized format. To date the StoP dataset release 3 includes 33 studies with a total of 12,753 stomach cancer cases and 30,682 controls [5]. For the present analyses we selected the studies with all the dietary information necessary to calculate a Mediterranean diet score, i.e. a total of 9 studies, 2 from Italy, 1 from Greece, 1 from Portugal, 2 from Spain, and 3 from Mexico, including overall 3322 cases and 8377 controls.
We selected 2 *a priori* scores to assess the adherence to the Mediterranean diet, with different characteristics in terms of dietary components included, cut-offs and scoring system to define the adherence, i.e. the Mediterranen Diet Score (MDS) proposed by Trichopoulou et al. [6], and the Cardioprotective Mediterranean Diet Score (CMDS) proposed by Martinez-Gonzalez et al. [7]. The MDS is the sum of 9 binary components, based on study-specific cut-offs: 1 point is assigned for a consumption above the study-specific median within sex group for components typical of the Mediterranean Diet (i.e. cereals, fruit and nuts, vegetables, legumes, fish, monounsaturated/saturated fatty acids ratio), and 0 otherwise; the reverse applies to non-Mediterranean components (i.e. meat, and dairy products), with 1 point assigned for a consumption below the study-specific median within sex group, and 0 otherwise; for alcohol, *a priori* defined cut-offs were used to credit moderate consumption (between 5 and 25 g of ethanol per day for women, and between 10 and 50 g of ethanol per day for men). Thus, the MDS ranges from 0 to 9, with higher scores indicating higher adherence to the Mediterranean diet.
The CMDS is the sum of 9 binary components, based on pre-specified cut-offs: 1 point is assigned to olive oil consumption greater or equal to 1 spoon per day, and 0 otherwise; 1 point to fruit consumption greater or equal to 1 serving per day, and 0 otherwise; 1 point to vegetable or salad greater or equal to 1 serving per day, and 0 otherwise; 1 point to fruit greater or equal to 1 serving per day and vegetable greater or equal to 1 serving per day, and 0 otherwise; 1 point to legumes greater or equal to 2 servings per week, and 0 otherwise; 1 point to fish greater or equal to 3 servings per week, and 0 otherwise; 1 point to wine greater or equal to 1 glass per day, and 0 otherwise; 1 point to meat lower than 1 serving per day, and 0 otherwise; 1 point to white bread lower than 1 serving per day and rice lower than 1 serving per week, or whole grain bread greater than 5 servings per week, and 0 otherwise. Thus, the CMDS ranges from 0 to 9, with higher scores indicating higher adherence to the Mediterranean diet.
For each study we harmonized data in order to create the study-specific components of the scores and overall MDS and CMDS. The overall scores were then categorized into 3 groups representing low, moderate and high adherence to the Mediterranean diet, based on tertiles in the study-specific distribution.
We used a two-stage approach to estimate the association between adherence to the Mediterranean diet and stomach cancer risk. In the first stage, we estimated study-specific odds ratios (ORs) and corresponding confidence intervals (CIs) of stomach cancer, running, separately for each study, multiple logistic regression models adjusted for age (5-year groups), sex, social class (low, intermediate, high, as defined in each original study based on education, income, or occupation), family history of stomach cancer (yes, no), tobacco smoking (never smoker, former smoker, light current smoker, moderate current smoker, heavy current smoker), and energy intake (quintile categories). In the second stage, we computed pooled estimates running a random effect model [8] on study-specific estimates. Furthermore, we performed a one-stage analysis including in a random effect model all the aforementioned covariates as well as the study id, and compared the results of this analysis with those of the two-stage analysis.

Results

About 82% of cases (2718 subjects) and 82% of controls (6883 subjects) included in the present analysis were from countries of the Mediterranean area, while the remaining subjects (604 cases and 1494 controls) were from Central America. Cases were more likely than controls to be of low socioeconomic status, smokers, heavy alcohol drinkers, and to have a family history of stomach cancer.
Original data available in the studies allowed to calculate the MDS in 7 studies (with the exception of 1 study from Italy, and 1 from Greece) and the CMDS in 6 studies (with the exception of 1 study from Italy, and 2 from Mexico).
Most studies showed significant inverse association between adherence to the Mediterranean diet and stomach cancer risk, considering both the MDS and the CMDS, with the only exception of 1 study from Portugal showing no significant association. In particular, the estimates for high versus low MDS were 0.53 (95% CI: 0.35-0.84) in the Italian study of Negri et al., 0.99 (95% CI: 0.75-1.32) in the Portuguese study of Lunet et al., 0.45 (95% CI: 0.31-0.66) in the Spanish study of Aragones et al., 0.56 (95% CI: 0.37-0.85) in the Spanish study of Vioque et al., 0.37 (95% CI: 0.22-0.62) in the Mexican study I of Lopez-Carrillo et al., 0.39 (95% CI: 0.24-0.64) in the Mexican study II of Lopez-Carrillo et al., and 0.41 (95% CI: 0.20-0.84) in the Mexican study III of Lopez-Carrillo et al.
The study-specific estimates for high versus low CMDS were 0.48 (95% CI: 0.38-0.61) in the Italian study of Palli et al., 0.25 (95% CI: 0.11-0.56) in the Greek study of Lagiou et al., 0.79 (95% CI: 0.58-1.06) in the Portuguese study of Lunet et al., 0.61 (95% CI: 0.45-0.82) in the Spanish study of Aragones et al., 0.50 (95% CI: 0.33-0.77) in the Spanish study of Vioque et al., 0.60 (95% CI: 0.38-0.95) in the Mexican study I of Lopez-Carrillo et al.
The Figure shows study-specific and pooled estimates based on two-stage random-effect models of stomach cancer risk for high versus low scores of adherence to the Mediterranean diet. The pooled estimate for high versus low MDS was 0.52 (95% CI: 0.38-0.71), in the presence significant heterogeneity (p=0.001). When we stratified the estimate by geographic area, the OR was 0.61 (95% CI: 0.42-0.91) for the Mediterranean area, and 0.39 (95% CI: 0.28-0.53) for the non-Mediterranean area.
We obtained similar results for high versus low CMDS (OR=0.56, 95% CI: 0.45-0.70), in the absence of significant heterogeneity (p=0.05). Corresponding estimates were 0.56 (95% CI: 0.45-0.70) for the Mediterranean area, and 0.60 (95% CI: 0.38-0.95) for the non-Mediterranean area.



Conclusion

In the present study, based on a uniquely large dataset, we found an inverse association between adherence to the Mediterranean diet and stomach cancer risk. The results were similar when we used the MDS – i.e. a score defining adherence on the basis of study-specific cutoffs, which largely varied depending on the study – and when we used the CMDS – which is based on predefined categories of adherence.
Furthermore, the inverse association between adherence to the Mediterranean diet and stomach cancer, was also confirmed in studies from outside the Mediterranean area, were this dietary pattern is supposed to be less common than around the Mediterranean basin, thus indicating that the favourable is likely real.

References

[1] Bray, F., J. Ferlay, I. Soerjomataram, et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin, 2018. 68(6): p. 394-424.

[2] Barak, Y. and D. Fridman, Impact of Mediterranean Diet on Cancer: Focused Literature Review. Cancer Genomics Proteomics, 2017. 14(6): p. 403-408.

[3] Sofi, F., C. Macchi, R. Abbate, et al., Mediterranean diet and health status: an updated meta-analysis and a proposal for a literature-based adherence score. Public Health Nutr, 2014. 17(12): p. 2769-82.

[4] Pelucchi, C., N. Lunet, S. Boccia, et al., The stomach cancer pooling (StoP) project: study design and presentation. Eur J Cancer Prev, 2015. 24(1): p. 16-23.

[5] Website, S.P. 31/05/2019]; Available from: http://stop-project.org/.

[6] Trichopoulou, A., T. Costacou, C. Bamia, et al., Adherence to a Mediterranean diet and survival in a Greek population. N Engl J Med, 2003. 348(26): p. 2599-608.

[7] Martinez-Gonzalez, M.A., E. Fernandez-Jarne, M. Serrano-Martinez, et al., Development of a short dietary intake questionnaire for the quantitative estimation of adherence to a cardioprotective Mediterranean diet. Eur J Clin Nutr, 2004. 58(11): p. 1550-2.

[8] DerSimonian, R. and N. Laird, Meta-analysis in clinical trials. Control Clin Trials, 1986. 7(3): p. 177-88.

IL RUOLO DEI DETERMINANTI EARLY-LIFE E MID-LIFE SULLO STATO COGNITIVO

Simeon Vittorio[1], Dolce Pasquale[2], Gallo Valentina[3], Signoriello Simona[1], Signoriello Giuseppe[1], Amalia Mattiello[4], Panico Salvatore[4], Chiodini Paolo[1]

[1] *Unità di Statistica Medica, Università degli Studi di Napoli "Luigi Vanvitelli", Napoli*

[2] *Dipartimento di Sanità Pubblica, Università Federico II, Napoli*

[3] *Centre for Primary Care and Public Health, Blizard Institute, Queen Mary University of London, London, UK*

[4] *Dipartimento di Medicina Clinica e Sperimentale, Università Federico II, Napoli*

Introduzione

Ritardare il declino cognitivo e migliorare le condizioni di vita degli anziani è di fondamentale importanza per contrastare quello che sta diventando la condizione più frequente nei paesi economicamente sviluppati [1]. La popolazione anziana sta aumentando rapidamente e, in particolare, la popolazione femminile costituisce la maggioranza a causa della maggiore aspettativa di vita rispetto ai maschi [2]. Il paradosso demografico femminile consiste nel fatto che le donne vivono più a lungo, ma non necessariamente godono di una vita "più sana" nella loro vecchiaia [3]. Di conseguenza, la conoscenza dei determinanti clinici e dello stile di vita del declino cognitivo nelle donne è diventata una delle maggiori priorità nelle indagini epidemiologiche sulle popolazioni anziane [4]. In un nostro precedente lavoro [5], nell'ambito della collaborazione di EPIC Italia, abbiamo preso in analisi una serie di fattori clinici e di stile di vita evidenziando il ruolo del carico glicemico alimentare, e quindi dell'alimentazione ad alta intensità energetica, come determinante dello stato cognitivo. Altri studi osservazionali hanno contribuito in modo significativo [6] aumentando il numero di evidenze e migliorando le conoscenze sul potenziale ruolo dei fattori intellettuali nel corso della vita (ad es. educazione, complessità del lavoro, e attività mentale) nel migliorare la riserva cognitiva di queste persone e ritardare l'insorgenza di disturbi cognitivi (ad esempio, demenza). La riserva cognitiva esprime la capacità del cervello nel contrastare gli effetti dell'invecchiamento e della patologia cerebrale legata alla demenza in modo che sia necessario un carico maggiore di patologia prima che i segni di declino cognitivo o sintomi di demenza diventino evidenti. Come questa riserva cognitiva venga costruita e preservata nel corso della vita è di fondamentale importanza e spiegare l'associazione tra i vari determinanti dello stato cognitivo può portare a percorsi di prevenzione e strategie di riduzione del rischio della demenza.

Obiettivi

Valutare il ruolo dei determinanti dello stato cognitivo legati agli anni dell'infanzia (fattori early-life) e della mezza età (fattori mid-life) in una coorte di donne residenti nell'area metropolitana di Napoli. Valutare il potenziale effetto di mediazione dei fattori mid-life sull'associazione tra early e stato cognitivo, analizzando l'effetto diretto ed indiretto dei fattori early.

Metodi

Lo studio è stato condotto sulla coorte napoletana di EPIC (Progetto ATENA), che tra il 1993 e il 1997 ha arruolato 5.062 donne di età compresa tra i 30 e i 69 anni [7]. Tra il 2008 e il 2009, alle donne con un'età uguale o superiore ai 65 anni (1.609 donne), è stato somministrato un questionario telefonico per valutare lo stato cognitivo (TICS - Telephone Interview to evaluate Cognitive Status). Il questionario TICS comprende i domini orientamento, memoria a breve e lungo termine, attenzione/calcolo e comprensione e ripetizione. Il TICS è stato ampiamente utilizzato e validato [8] e ha un punteggio totale di 39, ottenuto sommando il punteggio relativo ad ogni singolo dominio.

In seguito a selezione per criteri di esclusione, tra cui il rifiuto a partecipare e la presenza di patologie gravi o depressione, l'analisi è stata effettuata su un totale di 1.514 partecipanti.
Sono state raccolte le caratteristiche sociodemografiche (età e istruzione) e le misure antropometriche (peso e altezza). Inoltre, sono state valutate una serie di informazioni cliniche delle partecipanti (pressione sistolica e diastolica, colesterolo totale siero, HDL e LDL). I soggetti che hanno riportato diagnosi e/o una storia di trattamento del diabete di tipo II sono stati considerati diabetici. Le informazioni su abitudini alimentari e stile di vita sono state raccolte tramite questionario validato nell'ambito dello studio EPIC italiano. Tutte le informazioni e le variabili legate agli anni dell'infanzia (fino ai 18 anni) sono state classificate come fattori early-life, mentre quelle relative al momento del reclutamento sono state classificare come fattori di mezza età o fattori mid-life.
È stato valutato l'effect-size [9] per ogni incrocio delle variabili in studio. Nello specifico sono stati calcolati i seguenti indici di effect-size: i) r per due variabili continue – effetto piccolo se $0.1 \leq r < 0.3$, effetto intermedio se $0.3 \leq r < 0.5$, effetto grande se $r \geq 0.5$; ii) d di Cohen d per variabile continua e dicotomica – effetto piccolo se $0.2 \leq d < 0.5$, effetto intermedio se $0.5 \leq d < 0.8$, effetto grande se $d \geq 0.8$; iii) η^2 per variabile continue e categorica – effetto piccolo se $0.01 \leq \eta^2 < 0.06$, effetto intermedio se $0.06 \leq \eta^2 < 0.14$, effetto grande se $\eta^2 \geq 0.14$; iv) v di Cramer per due variabili dicotomiche o categoriche - effetto piccolo se $0.1 \leq V < 0.3$ con 1 df o $0.07 \leq V < 0.21$ con 2 df o $0.06 \leq V < 0.17$ con 3 df, effetto intermedio se $0.3 \leq V < 0.5$ con 1 df o $0.21 \leq V < 0.35$ con 2 df o $0.17 \leq V < 0.29$ con 3 df, effetto grande se $V \geq 0.5$ con 1 df o $V \geq 0.35$ con 2 df o $V \geq 0.29$ con 3 df.
Un modello di regressione lineare è stato utilizzato per stimare l'associazione indipendente tra il punteggio TICS e ciascuna delle variabili in analisi, scelte a priori e sulla base dell'analisi dell'effect-size. Le variabili che mostravano un alto effect-size sono state prese in considerazione, preferendo l'inserimento nel modello lineare delle variabili early-life.
Sono stati considerati i seguenti modelli: (1) modello multivariato dei fattori early-life aggiustato per età; (2) modello multivariato dei fattori early-life aggiustato per età e fattori mid-life.
Nel caso di modelli nested, è stato calcolato il Likelihood-ratio test (LR test) per valutare il miglioramento e la significatività del modello. Nel caso di modelli non-nested, sono stati calcolati l'Akaike Information Criterion (AIC) e il Bayesian Information Criterion (BIC) preferendo i modelli con valori di AIC e BIC inferiori.
È stato inoltre calcolato, per il modello finale, l'incremento di R^2 relativo dovuto all'aggiunta di ogni singola variabile o dell'intero blocco di variabili early-life e successivamente mid-life. Queste analisi statistiche sono state eseguite utilizzando il software statistico Stata versione 14 (Stata Corp., College Station, TX, USA).
È stato eseguito il metodo Partial Least Square Path Modeling (PLS-PM) [10], utilizzando il pacchetto seminr nel software R 3.6.0, al fine di valutare l'effetto dei fattori early sul TICS mediato dai fattori mid e calcolare l'effetto diretto ed indiretto. I fattori sono stati identificati sulla base dei risultati delle analisi precedenti. Il PLS-PM viene utilizzato per analizzare le relazioni lineari tra variabili latenti (VL) di input, intermedie e di output. Poiché le VL sono definite attraverso la combinazione lineare delle variabili osservate, il metodo PLS-PM è riconosciuto come un modello ad equazioni strutturali basato sui compositi, ed è più indicato nel nostro studio dato che le VL sono realmente dei compositi piuttosto che costrutti definiti come fattori comuni. Le relazioni tra le VL, che identificano il modello strutturale, sono quantificate attraverso dei path coefficients (β), che rappresentano l'influenza delle VL esplicative sulle VL dipendenti e possono essere diretti o indiretti a seconda del numero di percorsi che li collegano. Le connessioni tra VL e le corrispondenti variabili osservate formano i modelli di misura e sono quantificati attraverso i pesi utilizzati per costruire i compositi. La valutazione del modello strutturale si basa sul segno, sull'ampiezza e la significatività dei path coefficients. Inoltre, si valuta la l'effect size calcolando l'f^2 di Cohen, che esprime la variazione dell' R^2 quando un il corrispondente path viene omesso dal modello. Effect-size f^2 di 0.02, 0.15 e 0.35 rappresentano rispettivamente piccoli, medi e grandi effetti [9, 11].
È stato utilizzato il metodo bootstrapping, con un ricampionamento di 5000 sottocampioni, per determinare la significatività dei path coefficients nel modello strutturale.

L'adattamento del modello globale è stato valutato mediante lo standardized root mean square residual (SRMR: un valore di 0 per SRMR indica un adattamento perfetto). Studi di simulazione mostrano che un valore di cut-off di 0.08, come proposto da Hu e Bentler [12], sembra essere adeguato.

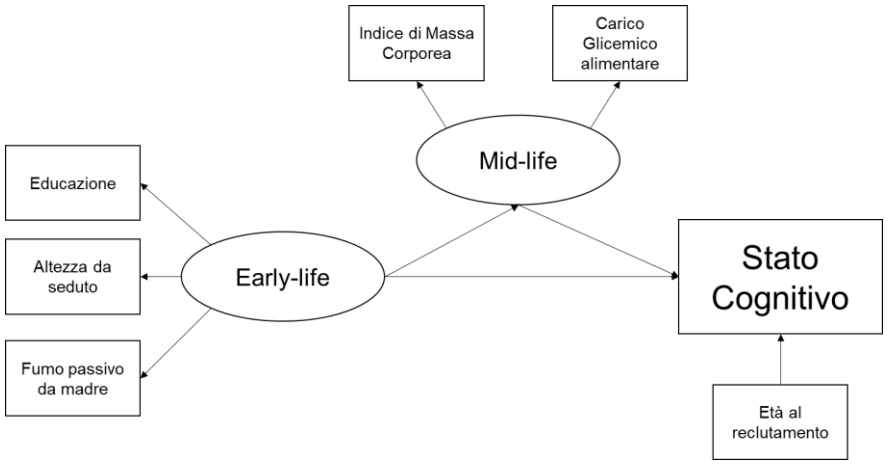
Risultati

Le variabili in analisi e le caratteristiche dei 1514 partecipanti sono state descritte in tutto il campione e in base ai terzili del punteggio TICS (punteggio basso 1-15, punteggio medio 16-21, punteggio elevato 22-39). Nel modello multivariato finale (modello 4) è stata osservata un'associazione significativa, e con andamenti coerenti, per le seguenti variabili: età al reclutamento (β -0.3; IC95% -0.34, -0.2); anni di scolarizzazione [(0-8 anni; ref.)(9-13 anni; β 3.9; IC95% 3.3, 4.6)(>13 anni; β 6.2; IC95% 5.2, 7.1)]; altezza da seduto (in centimetri) (β 0.2; IC95% 0.1, 0.3); fumo passivo della madre durante l'infanzia (β 0.96; IC95% 0.05, 1.86); indice di massa corporea (in kg/m^2) (β -0.12; IC95% -0.2, -0.05); carico glicemico alimentare (in g/d) (β -0.02; IC95% -0.03, -0.003).
Nel modello PLS-PM sono stati considerati come fattori che influenzano il punteggio TICS una VL di input (early-life), misurata da tre variabili osservate (altezza da seduto, livello di educazione e fumo passivo della madre) e una VL intermedia (mid-life), misurata da due variabili osservate (indice di massa corporea e carico glicemico alimentare), che media la relazione tra early-Life e il punteggio TICS. Gli effetti sono stati aggiustati per età al reclutamento (Fig 1). I pesi di tutte le variabili osservate sono positivi, sia per le VL early-life che per quella mid-life, il che significa che i punteggi delle VL aumentano all'aumentare delle variabili corrispondenti. Per la VL early-life, il livello di istruzione ha un peso maggiore (w = 0.82) nella costruzione del composito, mentre l'indice di massa corporea rappresenta la variabile osservata con peso maggiore (w = 0.81) nella definizione della VL mid-life.
La VL early-life influenza negativamente la VL mid-life (β = -0.24, p <0.001), con un effect size piccolo/medio (f^2 = 0.063) mentre ha un effetto diretto positivo sul punteggio TICS (β = 0.40, p <0.001), con un effect size medio/grande (f^2 = 0.198). Il mediatore mid-life ha un effetto negativo significativo sul punteggio TICS (β = -0.10, p <0.001), ma l'effect size è molto piccolo (f^2 = 0.012). L'effetto indiretto di early-life sul punteggio TICS mediato da mid-life è positivo e significativo (β = 0.02, p <0.001), ma con unffect size molto piccolo (f^2 = 0.014). La bontà di adattamento globale è soddisfacente (SRMR = 0.07).

Conclusioni

Lo studio evidenzia un'associazione tra educazione, altezza da seduto, fumo passivo dovuto alla madre durante l'infanzia, indice di massa corporea, carico glicemico alimentare e lo score dello stato cognitivo. Tra questi fattori, l'educazione e l'altezza da seduto sono noti proxy dello stato socioeconomico dell'infanzia. Tuttavia, anche il fumo passivo dovuto alla madre, rispetto al periodo storico considerato, potrebbe essere ritenuto un proxy dello stato socioeconomico dell'infanzia. Il modello PLS-PM ha mostrato come l'associazione dei fattori early-life con lo score TICS possa essere sia di tipo diretto che di tipo indiretto (mediato dai fattori mid-life). L'utilizzo di un approccio PLS-PM ha permesso di identificare, nell'analisi della variabilità dello stato cognitivo, il contributo dei diversi fattori identificando possibili pathway causali.
Considerando l'importante obiettivo di migliore le performance cognitive negli anziani e ridurre il rischio di declino cognitivo, è importante valutare i meccanismi con cui i fattori legati agli anni dell'infanzia e della mezza età siano in grado di influenzare lo stato cognitivo. Il miglioramento delle condizioni socioeconomiche e dei fattori di rischio modificabili legati allo stile di vita, potrebbe preservare la riserva cognitiva, stabilizzando e modificando le traiettorie di declino della funzione cognitiva.

Fig.1 *– Modello PLS-PM*



Bibliografia

[1] Bornebroek M., Breteler M.M. Epidemiology of non-AD dementias. Clinical Neuroscience Research., 2004, 3:349-361

[2] WHO | Women, ageing and health: A framework for action. Focus on gender. https://www.who.int/gender-equity-rights/knowledge/9789241563529/en/

[3] Van Oyen H., Nusselder W., Jagger C. et al., Gender differences in healthy life years within the EU: an exploration of the "health-survival" paradox. Int J Public Health. 2013; 58(1):143-55

[4] Ahmed T., Haboubi N. Assessment and management of nutrition in older people and its importance to health. Clin Interv Aging. 2010; 5:207-16

[5] Simeon V., Chiodini P., Mattiello A. et al., Dietary glycemic load and risk of cognitive impairment in women: findings from the EPIC-Naples cohort. Eur J Epidemiol. 2015; 30(5):425-33

[6] Fratiglioni L., Qiu C. Prevention of cognitive decline in ageing: dementia as the target, delayed onset as the goal. Lancet Neurol. 2011; 10(9):778-9

[7] Panico S., Dello Iacovo R., Celentano E. et al., Progetto ATENA, a study on the etiology of major chronic diseases in women: design, rationale and objectives. Eur J Epidemiol. 1992; 8(4):601-8

[8] Dal Forno G., Chiovenda P., Bressi F. et al., Use of an Italian version of the telephone interview for cognitive status in Alzheimer's disease. Int J Geriatr Psychiatry. 2006; 21(2):126-33

[9] Cohen J., Statistical Power Analysis for the Behavioral Sciences. Routledge. 1988. ISBN 978-1-134-74270-7.

[10] Tenenhaus M., Vinzi V.E., Chatelin YV. Et al., PLS path modelling. Computational Statistics & Data Analysis, 2005; 48(1):159-205

[11] Hair J., Hult T.M., Ringle M. et al., A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), Sage, 2014; 177-178

[12] Hu, L.T. and Bentler P.M., Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives, Structural Equation Modeling, 1999; 6(1): 1-55

COHORT ANALYSIS OF EPITHELIAL CANCER MORTALITY MALE TO FEMALE SEX RATIOS IN THE EUROPEAN UNION

Malvezzi Matteo[1], Bertuccio Paola[2], Carioli Greta[1], La Vecchia Carlo[1], Negri Eva[2]

[1] *Università degli Studi di Milano, Dipartimento di Scienze Cliniche e di Comunità, Milano, Italia*

[2] *Università degli Studi di Milano, Dipartimento di Scienze Biomediche e Cliniche L. Sacco, Milano, Italia*

Introduction

Male disadvantage in cancer incidence and mortality has long been documented [1, 2]. The main risk factors, aside from well established lifestyle exposures strongly associated with sex (e.g. smoking, alcohol consumption and numerous occupational exposures) to which these differences have been attributed, are related to carcinogenic susceptibility (including genetics), hormones, anthropometric measures, susceptibility and exposure to infections, and use of health care service and their accessibility [3-5].
Recently, sex ratios and sex differentials have been studied underlining both how different sites and types of cancer can be classified according to their sex differential etiologies, and the interest to study these sex differentials from a cohort perspective [3].

Aim

To update information on sex ratios in epithelial cancer mortality, with a focus on cohort patterns using both descriptive and graphical techniques.

Methods

We obtained official matrices of certified deaths and resident population from the World Health Organisation mortality database over the 1970-2014 period, for 28 countries of the European Union (EU as defined in July 2013) excluding Cyprus due to lack of data. When pooling data to construct the EU as a whole, if data for a cause or country were missing for a period, we used the nearest available (temporal) data [6].
We encoded causes of death using three successive Revisions of the International Classification of Diseases (ICD) over the studied period (ICD8, ICD9 and ICD10). All causes of cancer death were recoded according to the 10th Revision of the ICD. We considered 12 cancer sites: oral cavity and pharynx (C00-C14), oesophagus (C15), stomach (C16), colorectum (C17-C21, C26), liver (specified as primary) (C22.0-C22.7), pancreas (C25), larynx (C32), lung (C33-C34), skin (including melanoma) (C43-C44), bladder (C67), kidney and other urinary sites (C64-C66, C68) and thyroid cancer (C73).
From the matrices of certified deaths and resident populations, we computed age-specific rates and numbers of deaths for each cause of cancer death, sex, 5-year age group (from 0–4 to 85+ years) and quinquennium. We computed age standardized rates per 100,000 men and women, at all ages, using the direct method, on the basis of the world standard population. We calculated male to female Rate Ratios (RR) for both all ages standardised rates and raw age specific rates between the 30-34 and 85+ age groups.
To disentangle the effects of age, period of death and cohort of birth, we applied an Age-Period-Cohort (APC) model to quinquennial age-standardized rates and RR from 1970-74 to 2010-14 quinquenniums. Age groups were limited to those between 30 (30-34 years) and 84 (80-84 years) years, to avoid the effects of excessive variation in recent, young cohorts. Cohorts were defined according to their central year of birth, hence the earliest possible cohort is that of 1890 corresponding to 80-84 year olds who died in the 1970-74 quinquennium. We carried out the analysis using the APC R package [7]. Briefly, the age, period and cohort effects were estimated using a log-linear Poisson model with a likelihood penalizing function to solve the identifiability problem. To model RRs, instead of using log linear Poisson model, a normal model was applied

to log transformed RRs [8]. Confidence intervals (CIs) 95%, were calculated with a parametric simulation. The age effects are interpretable as mean age specific death rates per 100,000 person-years over the analysed period, in the case of RRs as the mean age specific male to female RR, while the period and cohort effects are expressed as multiplicative effects in relative terms against their weighted average set to unity.

Results

Male to female age standardised RRs were greater than unity for all the cancer sites considered with the exception of thyroid cancer in the EU over the whole considered calendar period. RRs were either descending or stable over the last decade, with the exceptions of skin, kidney and thyroid cancers. RR patterns for oral, oesophageal, laryngeal, lung, skin and bladder cancers followed the patterns in male all ages standardised mortality rates closely.
Figure 1 give the RR heatmap tables for the studied cancers and from 1970-74 to 2010-14 quinquenniums, in the EU. Calendar quinquenniums are given along the horizontal and age quinquenniums along the vertical, all ages age standardised RRs for the corresponding quinquennium along the top row. To highlight changes in values, each period and age cell is colour coded along a gradient colour scale from the lowest (green) to the highest (red) ones with white as the median, the scales are table specific.
The highest all ages age standardised RR (17.3) was in laryngeal cancer in the 1980-84, with the highest age specific RR (22.8) in the same quinquennium at age 45-49. In contrast, the lowest age standardised RR (0.48) was in thyroid cancer in 1970-74, with the lowest age specific RR (0.5) for the same cancer in 1975-79 at 80-84 years. Oral cavity, oesophagus, larynx and bladder cancers showed peaks in all ages standardised RRs before the 1990s and then presented descending values. Inspecting the age specific RRs for these cancers, we observed high values for young and old ages with peaks at middle age, similarly RR values were lower at the beginning of the studied period, relatively throughout the age groups, rose to reach a peak in the 1980s to then descend. Visually this corresponds to stronger red area centred on the age period peak (middle aged, 1980s) that weakened homogeneously along both axes turning to white and green. Stomach cancer age standardised RRs rose and stabilised around 2.2, age specific RRs were higher among the 50 to 75 year olds in 1970s and 1980s, with the age group moving up over the years to become higher among the age 55 to 80 in the 2010s. Cancers of the colorectum, liver and skin showed the highest RRs in the upper right quadrant of the tables, corresponding to the more recent years and older age groups, consequently showing rising age standardised RRs reaching 1.7 for colorectum and skin, and 2.9 for liver in 2010-15.
Results from the APC analysis showed that age effects prevalently had a similar pattern with a strong rise, a peak and then a fall for older ages in age specific male to female RRs. In oral, oesophageal, laryngeal, lung and thyroid cancers, this pattern showed an inverted “u” or “v” shape, with RRs in the young and old age groups with similarly low values. RR values for cancers of the stomach, colorectum, liver, skin, bladder, and kidney showed a peak at an older age to then slightly decrease. Age effects in male to female RRs for pancreatic cancer showed a much earlier peak and values fell much lower than where they started. Cohort effects showed a number of different patterns across cancer sites. Mostly descending patterns through the cohorts emerged for oral, pancreatic, laryngeal and lung cancers. Bladder cancer also showed smaller effects in recent cohorts, but falls only started after an initial plateau (i.e. those born after the mid 1940s). The opposite rising pattern in cohort effects can be seen in liver, kidney and thyroid cancers. Colorectal cancer also showed a rising trend in cohort effects, however this reached a plateau for more recent cohorts, after the 1940s. Skin cancer multiplicative cohort effects rose up the 1940s, reached a peak and then fell. A similar trend in cohort effects was observed in stomach cancer. Multiplicative cohort effects in male to female cancer mortality RRs for oesophageal cancer had no distinguishable pattern or were indeterminate due to excessive variation. Multiplicative period effects were of very modest entity or excessively variable in most

cases, however oral and lung cancers showed period effects rising for the earliest quinquenniums with a favourable downturn for the most recent ones.

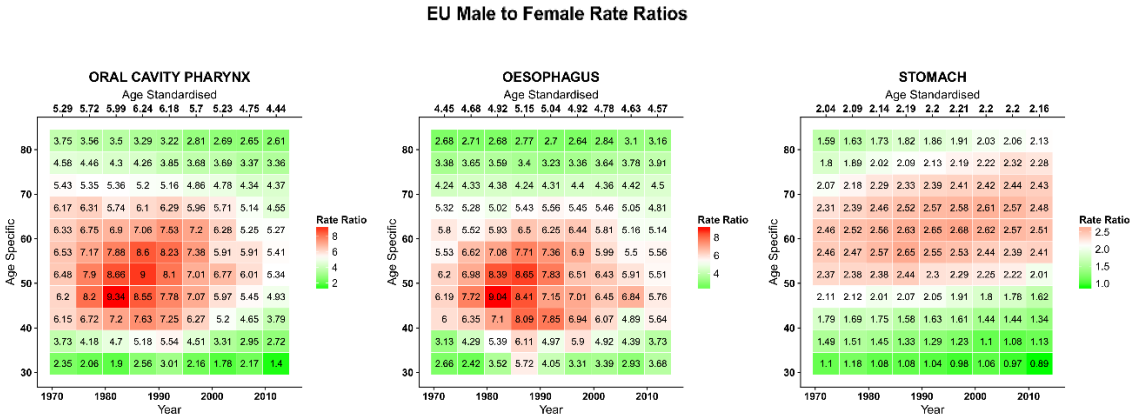
Conclusions

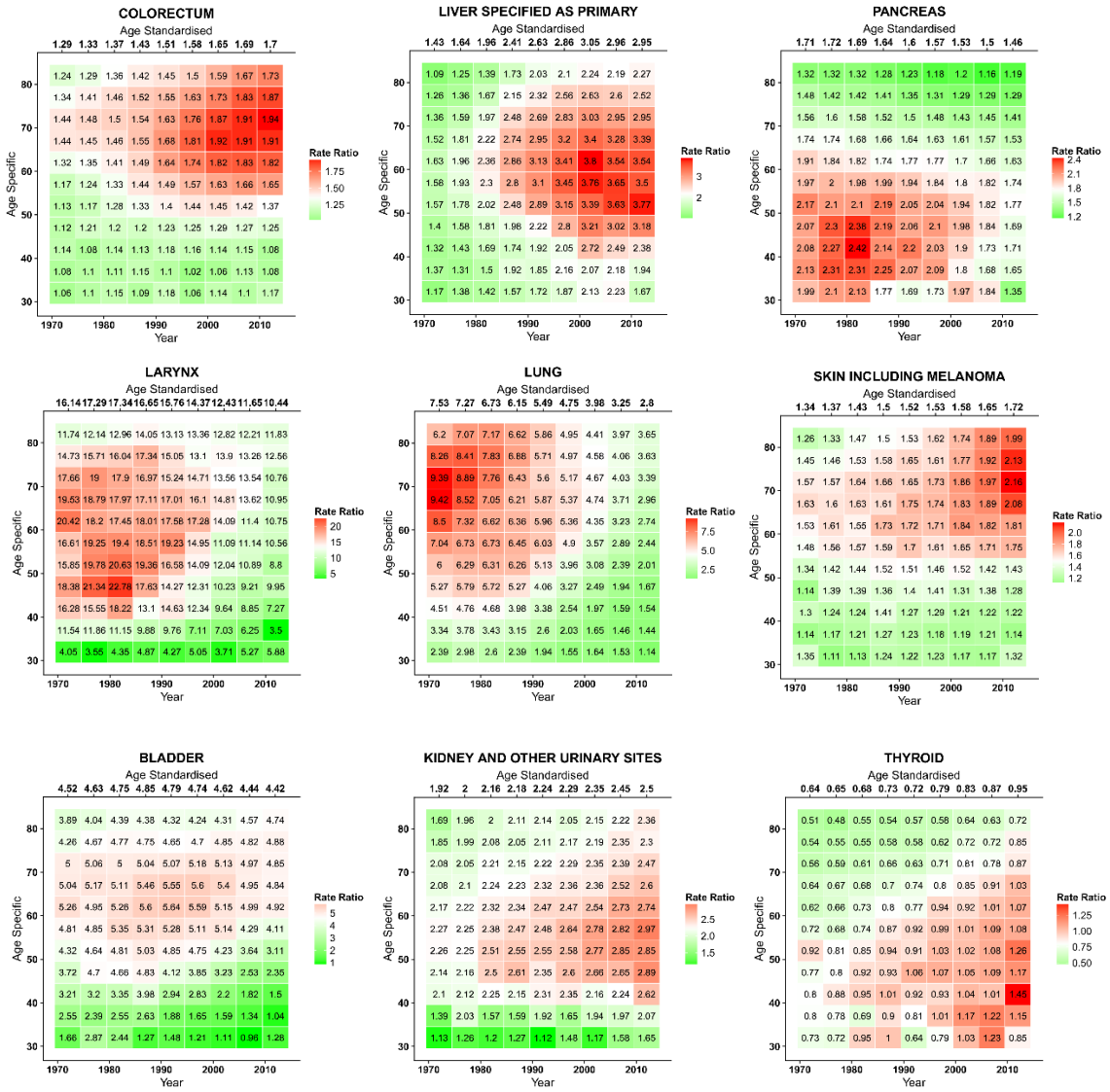
Cancer incidence and mortality have long known to be worse in men compared to women [1, 2]. However, for most cancer sites this difference has not been entirely or satisfactorily explained, in spite of a renewed interest on the issue [3, 5], in which researchers underline the lack of explicit focus on heterogeneity in cancer studies in favour of aetiology, and in particular differences between the sexes. Furthermore cohort effects in cancer sex ratios have never been comprehensively studied in a modelling context. Cohort patterns have been discussed speculatively by inspection of raw age specific rates, and by plotting the age curves at different points in time [2, 9].
To investigate cohort patterns cancer mortality male to female sex ratios, we used both visual tools on raw age specific RRs in the form of Lexis diagrams enhanced with colour gradients to highlight patterns in male to female RRs, and analytic modelling techniques in the form of a modified version of a penalised likelihood APC model [7]. We adapted a standard penalised likelihood model by selecting an appropriate distribution for the male to female RRs [8] and to compute the simulated 95% CIs for the age, period and cohort effect estimates (ratio from 2 binomial distributions determined form the parameters of the male and female age specific rates). This APC method inherits the limitations of its original implementation, i.e. that it favours cohort effects due to the higher number of factor levels and that younger cohorts tend to be unstable due to the smaller number at risk man years and hence numbers of events. In addition, the nature of male to female RRs also poses limitations, in the case of very low rates the effects this model estimates become unstable. Mortality sex ratios are less subject to the influence of changes in diagnosis, coding and classification of disease as well as preventive strategies [10]. However, since RRs are dimensionless quantities, their interpretation can be problematic, to this end information on the mortality rates for both men and women are provided for context and comparison. Other limitations are due to instability of RRs when mortality rates are low and issues of excessive random variation arise. To minimize this issue, we limited our study to a geographic area with a very large population (over 500 million) and high quality mortality data available from the WHO database. However, in interpreting the results from this analysis the differences within the EU populations must be considered, in particular between western and northern EU countries and eastern former non-market economy ones [11].
In this work, we analysed epithelial cell cancers because they are more susceptible to environmental factors [12] .Consequently, we expected to observe patterns in male to female RRs corresponding to known differences in exposures.
The more immediately visible patterns in male to female RRs are related to smoking and alcohol consumption. The smoking pattern is observed in lung cancer (85% attributable fracion) while alcohol in oesophageal cancer (50% attributable fraction in men, 30% in women) [13, 14].With these 2 patterns alone and their combination, the differences in mortality between the sexes can be largely explained according to their respective and combined attributable fractions in the following cancers: oral cavity and pharynx, oesophagus, larynx and lung cancer, and to a smaller degree colorectal, liver and pancreatic, cancers.
Another set of cancer risk factors that are susceptible to sex differentials are infectious diseases, which do not only have different prevalences but also prognostic differences. Stomach cancer for example, while associated with smoking, is strongly associated to *H. pylori*, diet and obesity. In the EU the pattern is more related to differences in *H. Pylori* infection - which is more common in men [15] and more probably lifestyle habits throughout successive cohorts [16, 17]. In general considering that gastric mortality is falling in both sexes, whether the sex RR rises or falls depends on the relative speeds at which male and female mortality rates fall. Liver cancer's main risk factors are hepatitis B (HBV) and hepatitis C (HCV), besides alcohol. The EU's heterogeneous population probably makes the patterns less defined, HCV is the main risk factor in

southern Europe, while central and northern Europe heavy drinking is an issue, as well as different timeframe of application of HBV primary prevention [18]. Pancreatic cancer also has its main known modifiable risk factor in smoking, however the observed age specific RR patterns do not appear to have much in common with those seen in lung cancer, but the cohort effects from the APC analysis are similar. This is possibly due to different attributable fractions and latency in cancer onset, other modifiable factors are obesity diabetes and heavy drinking [19]. The pattern for colorectal cancer male to female mortality RRs, as well as being related to risk factors such as alcohol, tobacco, obesity and sedentariness, and reduced risk from oral contraceptive hormonal replacement therapy use in women, could also be related to differences in adherence to screening programs since the ages and period affected coincide with their period and age of application [20]. Skin cancer and melanoma are related to UV exposure, however due to the excess mortality in men presenting itself in more recent years and only for older ages the more accredited explanations lie with greater at risk behaviours in males including both exposure (possibly occupational as well) and lack of protection (sun screen use), and lack of awareness, i.e. failure to recognise cancerous lesions and less frequent contact with medical personnel [21], however the lack of, or inverse pattern at ages under 40 years, point to other factors as well. Bladder cancer's strongest risk factor is smoking and this pattern is somewhat visible, however occupational exposures that were more prevalent in males also account for the differences [22]. The main risk factors for kidney cancer are obesity, smoking and hypertension, however the relation of these risk factors to the observed patterns in male to female RR are difficult to explain [23]. Thyroid cancer is the only cancer that presented a female disadvantage, however the fall in mortality due to this cancer in both sexes has brought the RR back to unity, the excess in incidence is mostly due to papillary thyroid cancer that has much higher survival rates, while follicular thyroid cancer has been mostly stable [24].

In conclusion the observed period and cohort patterns in male to female cancer mortality RRs are mostly explained by the differences in exposure to known avoidable risk factors. These are mostly tobacco, alcohol and obesity/overweight, and other lifestyle related factors, with primary prevention for HCV and HBV also playing role in liver cancer. These results underscore the importance of tobacco use and alcohol abuse prevention and the promotion of healthy lifestyles including diet and physical exercise in general.

Figure 1. *Heatmap tables of age specific and age standardised quinquennial male to female mortality rate ratios in epithelial cancers in the EU.*





References

[1] Ashley DJ. Sex differences in the incidence of tumours at various sites. Br J Cancer 1969; 23: 26-30.

[2] Levi F, La Vecchia C, Lucchini F et al. Trends in cancer mortality sex ratios in Europe, 1950-1989. World Health Stat Q 1992; 45: 117-164.

[3] Cook MB, Dawsey SM, Freedman ND et al. Sex disparities in cancer incidence by period and age. Cancer Epidemiol Biomarkers Prev 2009; 18: 1174-1182.

[4] Cook MB, McGlynn KA, Devesa SS et al. Sex disparities in cancer mortality and survival. Cancer Epidemiol Biomarkers Prev 2011; 20: 1629-1637.

[5] Edgren G, Liang L, Adami HO et al. Enigmatic sex disparities in cancer incidence. Eur J Epidemiol 2012; 27: 187-196.

[6] World Health Organization Statistical Information System. WHO mortality database Available at: http://www.who.int/healthinfo/statistics/mortality_rawdata/en/index.html (Last accessed December 2018).

[7] Decarli A, Vecchia CL, Malvezzi M et al. An R package for fitting age, period and cohort models. Epidemiology, Biostatistics and Public Health 2014; 11 doi 10.2427/9977.

[8] Koopman PAR. Confidence-Intervals for the Ratio of 2 Binomial Proportions. Biometrics 1984; 40: 513-517.

[9] Yang Y, Li T, Nielsen ME. Aging and cancer mortality: dynamics of change and sex differences. Exp Gerontol 2012; 47: 695-705.

[10] Nicholson WJ, Davis DL. Analyses of changes in the ratios of male-to-female cancer mortality. A hypothesis-generating exercise. Ann N Y Acad Sci 1990; 609: 290-297; discussion 297-299.

[11] Bertuccio P, Alicandro G, Malvezzi M et al. Cancer mortality in Europe in 2015, and an overview of trends since 1990. Ann Oncol 2019.

[12] Preston-Martin S, Pike MC, Ross RK et al. Increased cell division as a cause of human cancer. Cancer Res 1990; 50: 7415-7421.

[13] World Health Organization. WHO global report: mortality attributable to tobacco. In Organization WH (ed). Geneva: 2012.

[14] Praud D, Rota M, Rehm J et al. Cancer incidence and mortality attributable to alcohol consumption. Int J Cancer 2016; 138: 1380-1387.

[15] Ibrahim A, Morais S, Ferro A et al. Sex-differences in the prevalence of Helicobacter pylori infection in pediatric and adult populations: Systematic review and meta-analysis of 244 studies. Dig Liver Dis 2017; 49: 742-749.

[16] Malvezzi M, Bonifazi M, Bertuccio P et al. An age-period-cohort analysis of gastric cancer mortality from 1950 to 2007 in Europe. Ann Epidemiol 2010; 20: 898-905.

[17] Ferro A, Peleteiro B, Malvezzi M et al. Worldwide trends in gastric cancer mortality (1980-2011), with predictions to 2015, and incidence by subtype. Eur J Cancer 2014; 50: 1330-1344.

[18] Bertuccio P, Turati F, Carioli G et al. Global trends and predictions in hepatocellular carcinoma mortality. J Hepatol 2017; 67: 302-309.

[19] Lucas AL, Malvezzi M, Carioli G et al. Global Trends in Pancreatic Cancer Mortality From 1980 Through 2013 and Predictions for 2017. Clin Gastroenterol Hepatol 2016; 14: 1452-1462 e1454.

[20] Van der Vlugt M, Grobbee EJ, Bossuyt PM et al. Adherence to colorectal cancer screening: four rounds of faecal immunochemical test-based screening. Br J Cancer 2017; 116: 44-49.

[21] Swetter SM, Pollitt RA, Johnson TM et al. Behavioral determinants of successful early melanoma detection: role of self and physician skin examination. Cancer 2012; 118: 3725-3734.

[22] Fernandez MI, Brausi M, Clark PE et al. Epidemiology, prevention, screening, diagnosis, and evaluation: update of the ICUD-SIU joint consultation on bladder cancer. World J Urol 2019; 37: 3-13.

[23] Capitanio U, Bensalah K, Bex A et al. Epidemiology of Renal Cell Carcinoma. Eur Urol 2019; 75: 74-84.

[24] La Vecchia C, Malvezzi M, Bosetti C et al. Thyroid cancer mortality and incidence: a global overview. Int J Cancer 2015; 136: 2187-2195.

TRATTAMENTI FARMACOLOGICI POTENZIALMENTE INAPPROPRIATI E INTERAZIONI FARMACOLOGICHE A RISCHIO DI EVENTI AVVERSI

Lallo Adele[1], Alagna Michela[1], Onder Graziano[2], Ricci Roberto[3], Chiriatti Alberto[4], Di Martino Mirko[1], Davoli Marina[1]

per la Commissione Regionale del Farmaco della Regione Lazio (Co.Re.Fa.).

[1] *Dipartimento di Epidemiologia del S.S.R. – Regione Lazio – ASL Roma 1*

[2] *Fondazione Policlinico A. Gemelli IRCCS, Università Cattolica del Sacro Cuore, Roma*

[3] *Ospedale S. Spirito in Sassia, ASL Roma 1*

[4] *Medico di Medicina generale, ASL Roma 3*

Introduzione

Il trattamento farmacologico dell'adulto e dell'anziano è certamente un problema rilevante. I dati forniti dal rapporto 'L'uso dei farmaci in Italia' dell'Osservatorio Nazionale sull'Impiego dei Medicinali (OsMed) [1] mostrano che nella popolazione italiana, in particolare negli anziani, è molto comune il ricorso alla politerapia e, quindi, all'utilizzo concomitante di molteplici farmaci. Tuttavia, non sempre la terapia farmacologica porta con sé dei benefici clinici, e talvolta è causa di reazioni avverse, anche severe, che possono portare ad un aumentato rischio di ospedalizzazione e morte [2, 3]. La letteratura internazionale segnala che l'utilizzo di farmaci inappropriati, definiti come quei farmaci i cui rischi superano i possibili benefici clinici, e le interazioni tra farmaci possono essere comunemente causa di reazioni avverse che determinano un peggioramento delle condizioni cliniche dei pazienti [4]. Lo sviluppo di indicatori atti ad identificare questi due fenomeni nella popolazione italiana potrebbe rappresentare un passo importante per misurare e monitorare la qualità della prescrizione farmacologica a livello nazionale.

Obiettivi

- Identificare e quantificare le prescrizioni farmacologiche potenzialmente inappropriate e le associazioni terapeutiche ad alto rischio di interazione ed eventi avversi nella popolazione assistita e residente nella regione Lazio.
- Analizzare la variabilità nella prevalenza di prescrizioni inappropriate e nella prevalenza di associazioni terapeutiche a rischio di interazione tra le ASL della regione e tra i medici di Medicina generale (MMG).

Metodi

E' stato condotto uno studio di farmaco-utilizzazione utilizzando una "banca dati di popolazione" ottenuta integrando diversi Sistemi Informativi della Regione Lazio: Anagrafe degli Assistiti, Anagrafe Tributaria, Registro di Mortalità, Sistema Informativo Ospedaliero, Sistema Informativo della Farmaceutica Territoriale, Flusso della Farmaceutica ad Erogazione Diretta, Sistema Informativo delle Prestazioni Specialistiche Ambulatoriali, Archivio delle Esenzioni Ticket, Anagrafe dei medici di Medicina generale.
La popolazione in studio è costituita da tutti i soggetti assistiti e residenti nella regione Lazio al 31/12/2016 (data indice). Le prescrizioni farmacologiche sono state analizzate in un anno di *follow-up*, a partire dalla data indice. Sono stati considerati solo farmaci di fascia A.

Sono stati esclusi dalle analisi:

- i soggetti che non risultavano sempre assistiti dallo stesso medico di Medicina generale e i soggetti assistiti da MMG che hanno cambiato distretto o interrotto il servizio durante il periodo di *follow-up*.

Sono stati sviluppati una serie di indicatori volti a quantificare:

- i pazienti esposti a prescrizioni farmacologiche potenzialmente inappropriate;
- i pazienti esposti ad associazioni farmacologiche ad alto rischio di interazione ed eventi avversi.

L'esposizione 1) a farmaci potenzialmente inappropriati o 2) ad associazioni farmacologiche a rischio di eventi avversi è stata espressa in termini di 'prevalenza di prescrizione', misurata come segue:

1) numero di soggetti con almeno una prescrizione potenzialmente inappropriata nell'anno di *follow-up*, per 100 individui;
2) numero di soggetti con utilizzo concomitante di due o più farmaci ad alto rischio di interazione nell'anno di *follow-up*, per 100 individui.

Nel caso di frequenze estremamente basse, le prevalenze sono state espresse per 1000 individui.
L'utilizzo concomitante di 2 o più farmaci è stato definito come sovrapposizione delle coperture terapeutiche (*overlapping*) per almeno un giorno di trattamento. Per ciascuna prescrizione, la copertura è stata stimata utilizzando le dosi definite giornaliere [5].
Di seguito l'elenco degli indicatori proposti.

1. <u>Prescrizioni farmacologiche potenzialmente inappropriate</u>.

   1.1. *Proporzione di pazienti che utilizzano statine sul totale degli assistiti con età ≥ 80 anni in prevenzione primaria.*
   ATC: C10AA.
   1.2. *Proporzione di pazienti che utilizzano ticlopidina sul totale degli assistiti con età ≥ 65 anni.*
   ATC: B01AC05.
   1.3. *Proporzione di pazienti che utilizzano antidepressivi triciclici sul totale degli assistiti con età ≥ 65 anni*
   ATC: N06AA.
   1.4. *Proporzione di pazienti che utilizzano digitale ad un dosaggio superiore ai 0.125 mg/die sul totale degli assistiti con età ≥ 65 anni.*
   ATC: C01AA05.
   1.5. *Proporzione di pazienti che utilizzano citalopram ad un dosaggio superiore ai 20 mg/die o escitalopram ad un dosaggio superiore ai 10 mg/die sul totale degli assistiti con età ≥ 65 anni.*
   ATC: N06AB04 o N06AB10.
   1.6. *Proporzione di pazienti che utilizzano clorpropamide o glibenclamide sul totale degli assistiti con età ≥ 65 anni.*
   ATC: A10BB02 o A10BB01.

Nella costruzione dell'indicatore 1.1 vengono considerati diabetici:

- i soggetti dimessi, in regime ordinario o di *day-hospital*, con diagnosi di diabete in qualunque posizione, nel 2016 o nei 5 anni precedenti;
- i soggetti con almeno 2 prescrizioni di farmaci antidiabetici (ATC: A10) nel 2016 o in almeno uno dei 3 anni precedenti;
- i soggetti con codice esenzione per diabete (013) attivo alla data indice.

Vengono considerati, invece, pregressi eventi cardiovascolari e cerebrovascolari tutti i ricoveri per acuti, in regime ordinario, con i seguenti codici di diagnosi o procedura, in qualsiasi posizione, nei 10 anni che precedono la data indice:

- codici di diagnosi ICD-9-CM: 410-414, 433.x1, 434.x1, 435, 436, 438;
- codici di procedura ICD-9-CM: 00.66, 36.0, 36.1, 38.10, 38.11, 38.12, 00.61, 00.62, 00.63.

Per il calcolo dei mg/die relativi agli indicatori 1.4 e 1.5 è stata usata la formula:

$$\frac{N.di\ mg\ dipensati\ tra\ la\ prima\ e\ la\ penultima\ prescrizione\ (comprese)}{N.di\ giorni\ tra\ la\ prima\ e\ la\ penultima\ prescrizione}$$

2. Associazioni farmacologiche ad alto rischio di interazione ed eventi avversi.

2.1. *Proporzione di pazienti che utilizzano contemporaneamente 2 o più farmaci con effetto di allungamento del QT.*
I farmaci considerati sono:
- Macrolidi. ATC: J01FA.
- Chinolonici. ATC: J01MA.
- Antiaritmici ATC: C01B.
- Sotalolo. ATC: C07AA07.
- Citalopram o escitalopram. ATC: N06AB04 o N06AB10.
- Fluconazolo. ATC: J02AC01.
- Domperidone. ATC: A03FA03.
- Clorpromazina. ATC: N05AA.
- Aloperidolo. ATC: N05AD01

2.2. *Proporzione di pazienti che utilizzano contemporaneamente 2 o più farmaci che aumentano il rischio di sanguinamento.*
I farmaci considerati sono:
- FANS o COX2-i. ATC: M01A o M01B.
- Antiaggreganti. ATC: B01AC.
- Anticoagulanti: Rivaroxaban. ATC: B01AF01.
  Apixabam. ATC: B01AF02.
  Dabigratan. ATC: B01AE07.
  Warfarin. ATC: B01AA03.
  Acenocoumarol. ATC: B01AA07.

2.3. *Proporzione di pazienti che utilizzano contemporaneamente 2 o più farmaci che aumentano il rischio di insufficienza renale.*
I farmaci considerati sono:
- FANS o COX2-i. ATC: M01A o M01B.
- Spironolattone. ATC: C03DA01.
- ACE-i. ATC: C09A o C09B.
- Sartani. ATC: C09C o C09D.

2.4. *Proporzione di pazienti che utilizzano contemporaneamente cortisonici e chinolonici.*
- Cortisonici. ATC: H02.
- Chinolonici. ATC: J01M.

2.5. *Proporzione di pazienti che utilizzano contemporaneamente PPI e clopidogrel.*
- PPI. ATC: A02BC.
- Clopidogrel. ATC: B01AC04.

Per ciascun indicatore sono state analizzate le prevalenze di prescrizione per classe di età del paziente e per ASL di assistenza.

Infine, sono stati sviluppati tre 'indicatori compositi'.

3. Indicatori compositi:

   3.1. *Proporzione di pazienti con almeno una prescrizione inappropriata.*
   3.2. *Proporzione di pazienti con almeno un'associazione terapeutica a rischio di eventi avversi.*
   3.3. *Proporzione di pazienti con almeno una prescrizione inappropriata o un'associazione terapeutica a rischio di eventi avversi*.

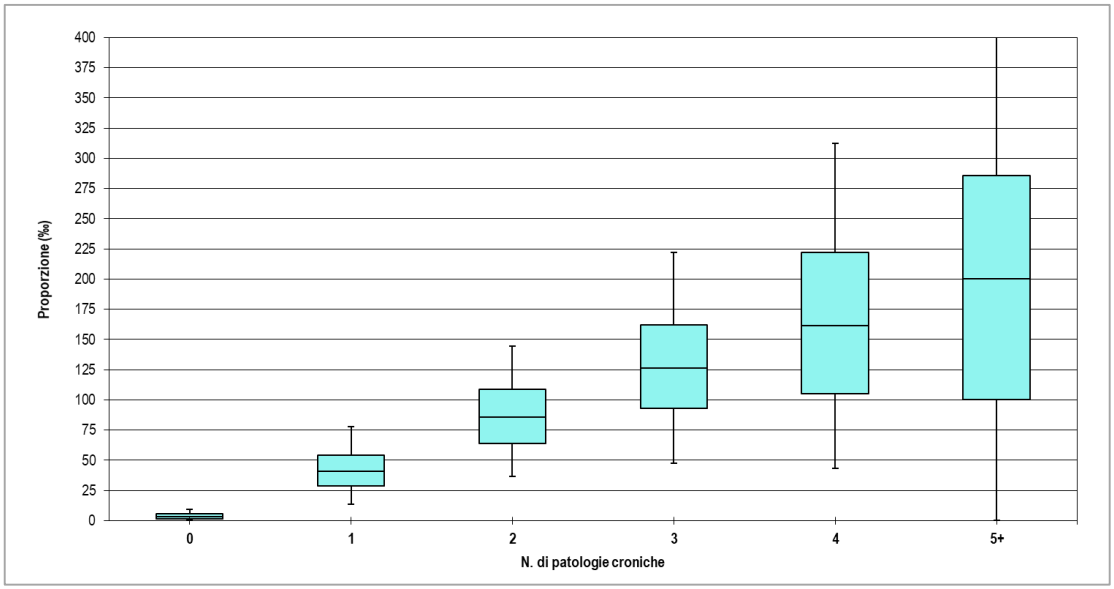
Per ciascun indicatore composito è stata misurata la variabilità tra ASL e per l'indicatore 3.3 viene riportata graficamente la variabilità tra MMG in funzione del numero di patologie croniche dei propri assistiti.
Le patologie croniche considerate sono: diabete, broncopneumopatia cronico-ostruttiva (BPCO), fibrosi polmonare idiopatica, sclerosi multipla, demenze, malattia di Parkinson, dializzati, ipertensione arteriosa, ipotiroidismo, tumore maligno, malattie cerebrovascolari, cardiopatia ischemica, malattie infiammatorie croniche dell'intestino (MICI) e artrite reumatoide.

Risultati

La popolazione in studio è costituita da 5 279 839 soggetti.
Tra gli indicatori volti a quantificare le prescrizioni di farmaci potenzialmente inappropriati, quello relativo alle statine presenta la prevalenza più alta in assoluto. La proporzione di ultra-ottantenni in prevenzione primaria, in trattamento con statine è pari al 24.4%. I grandi anziani, con 85 anni e più, sono maggiormente esposti, con una prevalenza del 28.2%. La percentuale di pazienti utilizzatori varia, rispetto alla ASL di assistenza, da un minimo del 21.4% ad un massimo del 26.8%.
Le prescrizioni inappropriate di ticlopidina riguardano circa 30 pazienti su 1000. La prevalenza di soggetti esposti aumenta all'aumentare dell'età ed è pari al 51.1‰ nei soggetti ultra-ottantacinquenni. Tra le ASL della regione la prevalenza varia da un minimo del 20.6‰ ad un massimo del 43.6‰.
Le prescrizioni inappropriate degli altri farmaci considerati hanno prevalenze decisamente più basse.
L'indicatore composito 3.1 è stato costruito considerando tutti i soggetti esposti ad almeno una prescrizione potenzialmente inappropriata. Tuttavia, essendo l'ordine medio di grandezza dell'indicatore 1.1 molto più elevato rispetto agli altri, le relative prescrizioni sono state escluse dal calcolo dell'indicatore composito, al fine di evitare che la sua variabilità "oscurasse" quella degli altri indicatori.
Complessivamente, la proporzione di soggetti esposti a prescrizioni inappropriate, sul totale della popolazione, è pari al 10.7‰. Considerando la ASL di assistenza, la proporzione minima è pari all'8.0‰, quella massima al 16.1‰.
Per quanto riguarda le associazioni terapeutiche a rischio di eventi avversi, l'utilizzo congiunto di farmaci che aumentano il rischio di insufficienza renale è quello più diffuso nella popolazione, con una prevalenza del 10.8‰. A meno di una riduzione osservata nei grandi anziani, la prevalenza aumenta all'aumentare dell'età, fino ad un massimo del 40.2‰ nella classe 75-84 anni. La proporzione di soggetti a rischio di eventi avversi varia, tra le ASL, da un minimo del 9.6‰ ad un massimo del 12.2‰.
La proporzione di pazienti che assumono contemporaneamente 2 o più farmaci che aumentano il rischio di sanguinamento e la proporzione di pazienti in trattamento con PPI e clopidogrel sono pari, rispettivamente, al 4.8‰ ed al 4.3‰.
Le altre associazioni terapeutiche a rischio di interazione hanno prevalenze inferiori all'1.0‰

Complessivamente, 18 pazienti su 1000 nella popolazione totale sono esposti ad almeno un'associazione ad alto rischio di eventi avversi (indicatore 3.2). La prevalenza varia da un minimo del 16.2‰ ad un massimo del 20.7‰ tra le ASL della regione.
Dal calcolo dell'indicatore composito (3.3) è emerso che la proporzione di soggetti esposti ad almeno una forma di inappropriatezza prescrittiva è pari al 27.6‰. La variabilità tra le ASL della regione è notevole: la prevalenza varia da un minimo del 24.9‰ ad un massimo del 35.5‰.
La figura 1 mostra la variabilità tra i medici di Medicina generale nella proporzione di pazienti esposti ad almeno una prescrizione inappropriata o un'associazione farmacologica a rischio di eventi avversi, in funzione del numero di patologie croniche del paziente.
Il 68.8% della popolazione in esame non ha alcuna patologia cronica tra quelle considerate, il 14.6% ne ha una, il 9.4% è affetto da 2 patologie, il 4.8% da 3, l'1.84 presenta 4 condizioni patologiche, il restante 0.64% ne presenta 5 o più.

Figura 1. *Variabilità nella proporzione di pazienti esposti ad almeno una prescrizione potenzialmente inappropriata o un'associazione farmacologica a rischio di eventi avversi. Box-plot per MMG (Lazio, 2017)*



Come è possibile vedere nella figura 1, all'aumentare del numero di patologie croniche cresce sia l'ordine medio di grandezza della prevalenza di pazienti esposti ad almeno una forma d'inappropriatezza prescrittiva, che la variabilità tra medici di Medicina generale: è proprio sul paziente "complesso", verosimilmente esposto ad una molteplicità di trattamenti farmacologici, che emergono le differenze tra i medici prescrittori.

Conclusioni

Lo studio identifica alcune aree d'azione per migliorare l'appropriatezza prescrittiva. Si propongono una serie di indicatori che valutano la qualità della prescrizione farmacologica. Questi indicatori, sviluppati in base alle evidenze provenienti dalla letteratura nazionale ed internazionale, sono stati applicati a livello regionale in una popolazione di oltre 5 milioni di persone, mettendo in evidenza come comune sia l'uso inappropriato di alcuni farmaci ed il rischio di interazioni farmacologiche, particolarmente nella popolazione anziana. Un ulteriore elemento di interesse è costituito dalla notevole variabilità tra medici di Medicina

generale nella "probabilità" di prescrivere farmaci inappropriati o a rischio di interazione. È importante sottolineare che gli indicatori sviluppati non possono sostituire il giudizio del medico e che possano esistere delle situazioni per cui il ricorso a farmaci inappropriati o a combinazioni che causano interazioni farmacologiche siano necessari per particolari esigenze cliniche. Il set di indicatori sviluppati può rappresentare la base per guidare interventi di salute pubblica finalizzati a migliorare la qualità della prescrizione farmacologica e ridurre i costi farmaceutici.

Bibliografia

[1] Osservatorio Nazionale sull'Impiego dei Medicinali, L'uso dei farmaci in Italia. Rapporto Nazionale 2017. Roma: Agenzia Italiana del Farmaco, 2018

[2] Onder G., Bonassi S., Abbatecola A.M., et al., on behalf of the Geriatrics Working Group of the Italian Medicine Agency, High prevalence of poor quality drug prescribing in older adults: a nationwide report from the Italian Medicine Agency (AIFA). J Gerontol A Biol Sci Med Sci., 2014 April; 69(4):430-437

[3] Onder G., Vetrano D.L., Cherubini A., Prescription drug use among older adults in Italy: a country-wide perspective. J Am Med Dir Assoc., 2014 Jul; 15(7): 531.e11-531.e15

[4] By the 2019 American Geriatrics Society Beers Criteria® Update Expert Panel, American Geriatrics Society 2019 Updated AGS Beers Criteria® for Potentially Inappropriate Medication Use in Older Adults. J Am Geriatr Soc., 2019 Apr; 67(4):674-694

[5] World Health Organization (WHO). Introduction to drug utilization research, 2003

TOTAL SERUM BILE ACIDS CONCENTRATION FOR THE DIAGNOSIS OF INTRAHEPATIC CHOLESTASIS OF PREGNANCY: A META-ANALYSIS USING INDIVIDUAL PATIENT DATA.

Casazza Giovanni [1], Manzotti Cristina [2], Nikolova Dimitrinka [3], Colli Agostino [3]

[1] *Dipartimento di Scienze Biomediche e Cliniche "L. Sacco" Università degli Studi di Milano; Milan, Italy*

[2] *UO Epatologia, Ospedale San Giuseppe, Università degli Studi di Milano; Milan, Italy*

[3] *Cochrane Hepato-Biliary Group, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Department 7812, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark*

Introduction

Intrahepatic cholestasis of pregnancy (ICP) is a pregnancy-specific liver disorder, associated with an increased risk of severe foetal adverse events. Total serum bile acids (TSBA) serum concentration is the most used biomarker for this diagnosis; it is widely used in clinical practice and quoted by scientific literature. TSBA concentration has been introduced as a diagnostic test for ICP based on case series and case-control studies that showed that the TSBA mean serum concentration was higher in women with intrahepatic cholestasis of pregnancy, when compared to healthy pregnant women [1, 2]. The most used cut-off value of TSBA concentration for ICP diagnosis is around 10 µmol/L to 14 µmol/L. In context of this knowledge, we planned a Cochrane systematic review to assess the accuracy of TSBA concentration for the diagnosis of ICP [3]. As we managed to obtain individual patient data (IPD) from authors of many of the included studies, we have been able to perform meta-analyses using methodological approaches that are not usually feasible if only cumulative data were available.

Objectives

The aims of the present work are: 1) to perform a meta-analysis of the study data in order to obtain estimates of sensitivity and specificity of TSBA at several different cut-off values; 2) to perform a meta-analysis of single study ROC (receiver operating characteristic) curves; 3) to assess the consistency between the two methods of meta-analysis.

Methods

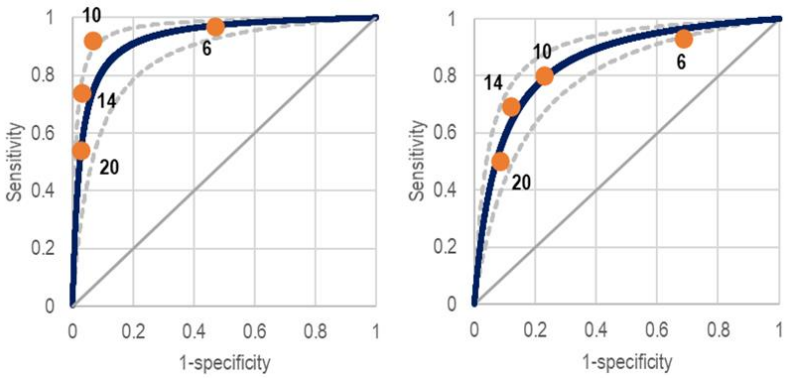
In our review, we included all the studies in pregnant women of any age, recruited in any clinical setting where the accuracy of TSBA for diagnosing ICP was assessed using clinical evaluation, including the period of follow-up after delivery, as the reference standard. We included thirteen studies of case-control and cross-sectional design, providing data for the 2x2 table. In case of missing information in the included studies, we contacted the authors. Thus, we obtained individual participants data from nine studies. In the present wok, we focus on the meta-analysis using only IPD provided by the authors of the included studies. We used a two-stage approach. In the first stage, we summarized each individual study by building 2x2 tables or ROC curves; in the second stage we pooled the individual study tables or curves across studies [4].
We identified four cut-off values for TSBA (6, 10, 14, and 20 µmol/L) as suggested by at least one of the included studies, and we performed separate meta-analyses for each single cut-off value. We calculated sensitivity and specificity for each study at each of the selected cut-off values, and then, we used the bivariate method for meta-analysis of sensitivity-specificity pairs [5]. Following this approach, a hierarchical model was fitted to jointly model the sensitivity-specificity pairs. At the first level, for sensitivity and specificity, we assumed a binomial distribution within studies. At the second level, we assumed that logit transformed sensitivity and specificity follow a bivariate normal distribution across studies, allowing for the possibility of

correlation between them [6]. Finally, pooled sensitivity and specificity with their 95% confidence intervals (CI) were obtained from the model-parameter estimates.
We also used all individual patient data to perform a meta-analysis of the ROC curves obtained from each individual study. Kester and Buntix proposed a straightforward method based on linearization of empirical ROC curves [7]. Briefly, a two-parameter model ($D=\alpha+\beta S$, where D is the sum and S is the difference of logit transformed sensitivity and logit transformed false positive ratio) was used to summarize the ROC curve of primary studies. α and β were estimated for each individual study, using a weighted linear regression, with weights proportional to the inverse of the variance of D. We used a bootstrap procedure to estimate the standard errors of α and β for individual studies. Then, we performed a meta-analysis, using the DerSimonian and Laird random-effects model, of the estimated values of α and β for all the nine studies, and obtained the pooled parameters (α_{pool} and β_{pool}). Finally, the summary ROC curve was obtained through back transformation of the pooled linear model in the ROC space.
We performed the two described meta-analyses by considering all the nine included studies, and we subsequently performed a sensitivity analysis by excluding three studies reporting an overestimation of diagnostic accuracy with both a sensitivity and a specificity of 100% for at least one cut-off (complete separation of data).
Furthermore, a visual comparison of the results obtained with the two methods was performed by plotting the results of both the analyses in the ROC space.

Results

Authors of nine of the included studies sent us individual data of TSBA and ICP status of 600 patients (ranging from 14 to 193 patients per study). Using the bivariate approach, we obtained sensitivity 0.917 (95% CI, 0.747 to 0.980) and specificity 0.935 (95% CI, 0.779 to 0.980) for the most commonly used cut-off of 10 µmol/L. Performing the analysis after the exclusion of the three studies with 100% sensitivity and 100% specificity, we obtained (analyses conducted on six studies, 422 patients) sensitivity 0.803 (95% CI, 0.626 to 0.909) and specificity 0.770 (95% CI, 0.705 to 0.824) for the most often used cut-off value of 10 µmol/L.
The obtained summary ROC curves, using the approach suggested by Kester and Buntix, are reported in Figure 1. The obtained pooled estimates with the bivariate approach are superimposed in order to allow a visual comparison of the results obtained with the two methods.

Figure 1. *Summary ROC curve with 95% confidence band and pooled estimates of sensitivity and specificity obtained with the bivariate method for the four prespecified cut-off values (orange dots). Numbers reported represent TSBA cut-off values (µmol/L). Left panel: meta-analysis of all the nine studies. Right panel: meta-analysis of six studies (sensitivity analysis).*



Conclusion

We found that the accuracy of TSBA for the diagnosis of ICP seems suboptimal: TSBA could provide some information, but TSBA should not be used as a unique marker for the diagnosis of intrahepatic cholestasis of pregnancy.
The two statistical approaches used in our analyses provided consistent results, especially after the exclusion of the three studies that had likely overestimated the diagnostic accuracy.
When individual patient data are available, a meta-analysis of individual ROC curves, in combination with meta-analysis of sensitivity and specificity at some prespecified cut-off points, should be considered as an additional, rather than an alternative, analysis. A strength of this approach is that it allows assessment of the overall accuracy of a test and, in addition, it can provide clinically useful information regarding the accuracy at specific operating points.

References

[1] Sjövall K, Sjövall J. Serum bile acid levels in pregnancy with pruritus (bile acids and steroids 158). Clin Chim Acta 1966;13:207-11

[2] Heikkinen J, Ylostalo P, Maentausta O et al. Bile acids in maternal serum, umbilical cord serum and amniotic fluid of healthy women, women with pruritis and patients with intrahepatic cholestasis of pregnancy. J Obstet Gynaecol 1983;4(1):17-20.

[3] Manzotti C, Casazza G, Stimac T et al. Total serum bile acids or serum bile acid profile, or both, for the diagnosis of intrahepatic cholestasis of pregnancy. Cochrane Database Syst Rev 2017, Issue 2. Art. No.: CD012546. DOI: 10.1002/14651858.CD012546.

[4] Debray TPA, Moons KGM, Abo-Zaid GMA, et al. Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 2013; 8(4): e60650. doi:10.1371/journal.pone.0060650

[5] Reitsma JB, Glas AS, Rutjes AW et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005; 58: 982-990.

[6] Harbord RM, Deeks JJ, Egger M et al. A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics 2007; 8: 239-251.

[7] Kester A, Buntix F. Meta-analysis of ROC Curves. Med Decis Making, 2000;20: 430-439

UTILIZZO APPROPRIATO DEL TRASTUZUMAB NEL LAZIO: SCENARI TERAPEUTICI E STIMA DEL RISPARMIO PER IL SSN

Alessandro Cesare Rosa[1], Felice Musicco[2], Paolo Marchetti[3], Nello Martini[4], Antonio Addis[1], Valeria Belleudi[1].

1 Dipartimento di Epidemiologia, Servizio Sanitario Regionale del Lazio. Roma, Italia;

2 IFO Istituti Fisioterapici Ospitalieri di Roma, Italia;

3 Università La Sapienza, Roma, Italia;

4 Fondazione Res, Ricerca e Salute Roma, Italia.

Introduzione

La disponibilità delle formulazioni sottocute (SC) ed endovena (EV) del trastuzumab e la recente immissione nel mercato della versione biosimilare EV (10/2018) aumentano l'offerta per il trattamento dei pazienti con carcinoma mammario o gastrico HER2 positivi. In questo contesto, risulta necessario disporre di dati sui quali costruire scenari terapeutici in grado di stimare il risparmio per il Servizio Sanitario Nazionale (SSN) tenendo conto dell'appropriatezza d'uso e delle preferenze del paziente. [1,2]

Obiettivo

Analizzare l'offerta trastuzumab per tipologia di somministrazione (EV/SC) nel Lazio nel 2018 individuando l'uso più appropriato; ipotizzare uno scenario economicamente vantaggioso per il SSN nel 2019.

Metodi

Attraverso il flusso della farmaceutica diretta, sono stati identificati i cicli di trastuzumab erogati nel 2018 separatamente per le formulazioni disponibili (EV/SC). Per ogni ciclo di terapia è stata ricercata, a partire dalla data di erogazione, l'eventuale presenza di un trattamento concomitante (± 2 giorni) tracciando il tipo e la modalità di somministrazione. Sono stati individuati i trattamenti concomitanti al trastuzumab per i quali si avevano modalità di somministrazioni discordanti. Inoltre è stata valutata l'offerta di SC per struttura erogatrice e, sulla base di tale informazione, è stato ipotizzato uno scenario di utilizzo che tenga conto dei costi e dei consumi plausibili.

Risultati

Nel Lazio nel 2018, si sono osservati 22.214 cicli di trattamento per il trastuzumab (33 milioni di euro) relativi a 2.407 pazienti, i nuovi utilizzatori erano il 52,2%. Il 46,8% dei cicli veniva somministrato per via SC; l'utilizzo del biosimilare è stato osservato a partire dal mese di Ottobre ed ha riguardato 143 cicli (0.6%). Nel 68,4% dei cicli il trastuzumab veniva erogato in monoterapia; tra le terapie associate al trastuzumab la più frequente risultava essere il pertuzumab (n=4258, %19,2 sul totale dei cicli), seguita dal paclitaxel (n=1364, 6,1% sul totale dei cicli). Tra i cicli di trastuzumab in terapia concomitante (N=7022), il 17,3% veniva somministrato per via SC nonostante la presenza di altro farmaco somministrato per via EV. Il pattern prescrittivo del trastuzumab risultava eterogeneo per struttura di erogazione. La somministrazione SC presentava una variabilità nell'offerta dal 26% al 70% (range interquartile) e non sembrava essere in relazione con la tipologia e volume di attività dell'ospedale.

Prevedendo per il 2019 un uso del biosimilare EV pari al 45% dei consumi e riducendo così l'uso dell'originator EV al 10% e dell'originator SC al 35%, si potrebbe ottenere un risparmio di oltre 7 milioni di euro.

Conclusioni

Attraverso i dati di utilizzo delle diverse formulazioni disponibili per il trastuzumab e tenendo conto dell'appropriatezza prescrittiva (e delle preferenze del paziente), è stato possibile identificare uno scenario per il SSN di convenienza economica grazie al maggior utilizzo del biosimilare.

Bibliografia

[1] Azienda Provinciale per i Servizi Sanitari. Provincia Autonoma di Trento relazione di carattere tecnico scientifico a supporto dell'utilizzo del Trastuzumab biosimilare nel tumore alla mammella. 14 febbraio 2019

[2] Cicchetti A. et al. Mini HTA di rituximab e trastuzumab in formulazione sottocutanea, Working Paper Altems n. 1, 2017. https://altems.unicatt.it/altems-HTA_Sottocute__Finale_intero_rev.pdf

VALUTAZIONE DEL PERCORSO DIAGNOSTICO TERAPEUTICO ASSISTENZIALE NEI PAZIENTI CON IL DEFICIT DELL'ORMONE DELLA CRESCITA IN ETÀ EVOLUTIVA NELLA REGIONE MARCHE. RISULTATI PRELIMINARI

Skrami Edlira[1], Dajko Marianxhela[1], Gesuita Rosaria[1], Bucci Andrea[1], Faragalli Andrea[1], Ferrante Luigi[1], Pricci Flavia[2], Rotondi Daniela[2], Agazio Elvira[2], Patregnani Luigi[3], Brugia Marina[4], Cherubini Valentino[5], Carle Flavia[1]

[1] *Centro di Epidemiologia, Biostatistica e Informatica medica, Università Politecnica delle Marche, Ancona*

[2] *Registro Nazionale Assuntori Ormone della Crescita, Istituto Superiore di Sanità*

[3] *Assistenza Farmaceutica, Agenzia Regionale Sanitaria, Regione Marche*

[4] *Laboratorio Analisi, Azienda Ospedaliero-Universitaria, Ospedali Riuniti di Ancona*

[5] *Divisione di Diabetologia Pediatrica, Dipartimento Materno Infantile, Azienda Ospedaliero-Universitaria, Ospedali Riuniti di Ancona, "G. Salesi"*

Introduzione

Il deficit dell'ormone della crescita è una condizione patologica caratterizzata da anomalie auxologiche, cliniche e metaboliche causate da una ridotta secrezione dell'ormone della crescita (GH) [1]. In età evolutiva, la sua incidenza si stima intorno a 1/4000 nati vivi [2] e può essere diagnosticata come deficit idiopatico o secondario a molteplici cause. La diagnosi, in particolare della forma idiopatica, rimane ancora oggi difficoltosa.
L'efficacia del trattamento sostitutivo con GH nel paziente pediatrico è condizionata alla correttezza della diagnosi e alla tempestività dell'inizio del trattamento [3]. L'Agenzia Italiana del Farmaco (AIFA) ha stabilito che per la diagnosi del deficit di GH e la conseguente concessione della terapia sostitutiva con GH, siano rispettati determinati parametri clinico-auxologici e che ci sia la risposta patologica a due test di stimolo per GH [risposta di GH < 8 ng/mL a due test farmacologici eseguiti in giorni differenti, o GH < 20 ng/mL nel caso in cui il test impiegato sia GHRH (Gonadotropin Releasing Hormone) e arginina] [4]. La prescrizione del GH è a carico del Sistema Sanitario Nazionale e viene concessa su diagnosi e piano terapeutico di centri specializzati. Nella Regione Marche, con determina n. 33/ARS del 5/07/2013, la prescrizione del GH è subordinata all'inserimento da parte dei centri specializzati prescrittori di GH, dei pazienti in trattamento nel Registro Nazionale Assuntori Ormone della Crescita. Il registro è gestito dall'Istituto Superiore di Sanità.
La valutazione e il monitoraggio nel tempo del Percorso Diagnostico Terapeutico Assistenziale (PDTA) del deficit di ormone della crescita in età evolutiva costituiscono elementi fondamentali per ottimizzare la diagnosi e l'accesso alla terapia sostituiva con GH.

Obiettivi

Nell'ambito della valutazione del PDTA del deficit di GH, in questo studio si intende valutare l'appropriatezza del percorso diagnostico e del conseguente trattamento sostitutivo con GH in soggetti in età evolutiva nella Regione Marche.

Metodi

Lo studio di natura osservazionale, tuttora in corso, include bambini e adolescenti di età compresa tra 2 e 18 anni, residenti nella Regione Marche e sottoposti al test di stimolo per GH nel periodo 01/01/2013 – 31/12/2017.

Per poter valutare l'appropriatezza del percorso diagnostico è stata costituita la coorte incidente dei soggetti sottoposti per la prima volta dal 1/11/2014 al 31/10/2017 al test di stimolo per il GH, utilizzando come fonte di dati il Registro degli esami di laboratorio Azienda Ospedaliera Universitaria Ospedali Riuniti di Ancona e l'Anagrafica dei residenti della Regione Marche. Il Registro Nazionale degli Assuntori dell'Ormone della Crescita e l'Archivio Regionale della Farmaceutica territoriale sono stati utilizzati come fonte di dati per ottenere le informazioni sul trattamento con GH, utilizzando un record -- linkage di tipo deterministico tra le suddette fonti e la coorte in studio. Per ogni soggetto della coorte è stata individuata la prima prescrizione di GH.
L'appropriatezza del percorso diagnostico è stata valutata attraverso l'aderenza alle linee guida proposte dall'AIFA riguardo ai risultati dei due test di stimolo per GH, necessari per accedere al trattamento. È stato definito come appropriato il percorso diagnostico di soggetti con: a) risposta patologica a due test di stimolo per GH; b) risposta non patologica ad un test di stimolo per GH e un tempo di follow up nella coorte superiore ad un anno dalla data del test (periodo considerato come tempo sufficiente di esposizione ad almeno un'ulteriore test di stimolo per GH); c) risposta non patologica al primo test di stimolo per GH, in assenza di un secondo test nei 4 mesi successivi (periodo considerato come tempo minimo per valutare la crescita prima di una ulteriore esposizione al test di stimolo in caso di continuo sospetto di deficit di GH da parte dello specialista endocrinologo).
L'appropriatezza del percorso è stata calcolata come proporzione di soggetti con percorso diagnostico appropriato rispetto al totale dei soggetti inclusi nella coorte. Quando l'informazione relativa al tipo di test per GH non era disponibile e il risultato del test non permetteva la classificazione della risposta come patologica/non patologica (8 ng/mL ≤ risultato test GH < 20 ng/mL) il percorso diagnostico non poteva essere definito a causa dell'incompletezza della fonte dei dati.
L'appropriatezza del trattamento è stata valutata come proporzione di soggetti con inizio del trattamento con GH nel periodo 01/01/2015 - 31/12/2017, tra coloro con risposta patologica ai test e follow up di almeno 1 anno dall'ultimo test di stimolo.

Risultati

Nella coorte in studio sono stati inclusi 316 soggetti residenti nelle Marche, 60.4% maschi, che avevano effettuato il primo test di stimolo con GH tra il 1/11/2014 e 31/10/2017 nel laboratorio dell'Azienda Ospedaliera Universitaria Ospedali Riuniti di Ancona. Nel 78% dei casi la richiesta di prescrizione del test di stimolo per GH è stata effettuata dai centri prescrittori dell'ormone della crescita della stessa struttura ospedaliera del laboratorio per il test di stimolo. La risposta al test di stimolo in seguito all'indagine diagnostica è risultata patologica nel 20% dei soggetti, non patologica nel 37%. Nel 43% dei soggetti non è stato possibile classificare la risposta del test farmacologico per mancata informazione riguardo al tipo di test di stimolo, per aver osservato un solo test patologico o due test con esiti discordanti.
L'appropriatezza del percorso diagnostico è risultata pari al 54% e la non appropriatezza del 13% con un numero totale di test di stimolo evitabili pari a 54 che nella Regione Marche corrisponde ad un costo di circa 14.000 euro (comprende il costo complessivo di ricovero ospedaliero in regime di day hospital richiesto per effettuare il test di stimolo con GH, il costo dell'ormone e i dosaggi seriati). Per il 4% dei soggetti il tempo di follow up non è stato sufficiente per valutare l'appropriatezza dell'indagine diagnostica così come non è stato possibile valutare l'appropriatezza per il 29% dei soggetti, a causa della scarsa completezza della fonte dei dati.
Quarantanove soggetti della coorte (16%) avevano iniziato il trattamento con GH nel periodo 2015-2017. L'appropriatezza del trattamento per coloro che sono stati sottoposti all'indagine diagnostica è risultata pari a 55%.

Conclusioni

I primi risultati di questo studio sottolineano un elevato livello di incompletezza della fonte dei dati utilizzata e un'appropriatezza del percorso diagnostico e del conseguente trattamento che rimane lontana da livelli ottimali. Ciò potrebbe avere un impatto considerevole sia sull'efficacia del trattamento con GH che sui costi dell'indagine diagnostica.
I risultati ottenuti evidenziano la necessità da una parte di ottimizzare la qualità delle fonti dei dati specifiche e dall'altra di indagare i determinanti della limitata appropriatezza osservata.

Bibliografia

[1] Alatzoglou KS., Webb EA., Le Tissier P. et al. Isolated Growth Hormone Deficiency (GHD) in Childhood and Adolescence: Recent Advances. Endocr Rev., 2014 Jun;35(3):376-432

[2] Murray PG., Dattani MT., Clayton PE. Controviersies in the diagnosis and management of growth hormone deficiency in childhood and adolescence. Arch Dis Child., 2016;101:96–100

[3] Maghnie M., Labarta JI., Koledova E. et al. Short Stature Diagnosis and Referral. Front Endocrinol (Lausanne)., 2018 Jan;11(8):374.

[4] Italia. Determinazione dell'Agenzia Italiana del Farmaco 19 giugno 2014. Modifica alla Nota AIFA 39. Gazzetta Ufficiale - Serie Generale n. 154 del 5 luglio 2014.

DAY-TIME PATTERNS OF ENERGY INTAKE DERIVED VIA PRINCIPAL COMPONENT ANALYSIS AND THEIR RELATIONSHIP WITH BODY MASS INDEX AND WAIST CIRCUMFERENCE IN ADOLESCENTS: AN ANALYSIS OF THE UK NATIONAL DIET AND NUTRITION SURVEY (2008-2012).

Luigi Palla [1] and Suzana Almoosawi [2]

[1] *Department of Medical Statistics, London School of Hygiene and Tropical Medicine, UK; Luigi.Palla@lshtm.ac.uk*

[2] *Brain, Performance and Nutrition Research Centre, Northumbria University,UK*

Introduction

Accumulating evidence suggests a relationship between timing of energy intake and obesity. Most of this evidence first emerged in chronobiology where variations in circadian rhythms of energy metabolism, appetite hormones and hunger perception have been reported [1] and it was corroborated by a number of epidemiological studies [2-4] which found associations between evening or night-time energy intake and obesity. Whilst evening and night-time energy intake may potentially act as important target points for dietary interventions, describing the association between one pre-defined eating occasion or meal slot (i.e. evening or night-time) and obesity falls short of providing an overall summary of dietary behaviour with respect to the diurnal temporal dimension. In fact, several studies have indicated the (negative) correlation of such eating behaviours, with greater evening or night-time energy intake associated with decreased energy intake at breakfast and a shift in energy intake towards later in the day [5-7]. Hence the need for adopting more comprehensive methods for analysing time-of-day data, removing the focus from single time slots towards defining whole diurnal eating patterns (DEP).

Objective

Recent epidemiological studies have described the use of latent class analysis [8], cluster analysis [9] to describe DEP in the general adult population in Australia and Germany respectively. However no studies previously addressed the UK population and how these may relate to obesity prevalence, particularly in adolescence. Understanding current DEP in this age group is critical considering that most eating habits are formed in adolescence and that obesity in adolescence increases the probability of adult obesity by over 50% [10]. The present study aims to describe DEP in a nationally representative sample of UK adolescents, and assess the relationship between DEP and body mass index (BMI) or waist circumference (WC) as markers of obesity.

Methods

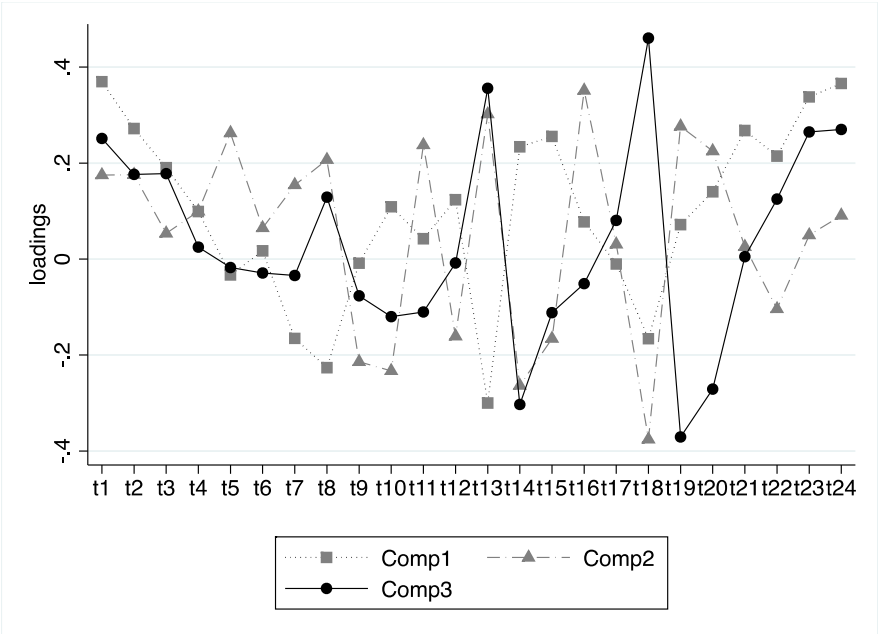
Study population. The study population consisted of a complete-case sample of 1438 adolescents aged 11-18yrs old who participated in the 2008-2012 National Diet and Nutrition Rolling Programme (NDNS RP) and who completed three or four days of dietary assessment. The NDNS RP is a cross-sectional rolling survey that collects yearly information on all food and drinks consumed from approximately a thousand randomly sampled individuals living in private households across the four regions in the UK [11].
Statistical analysis. Principal Component Analysis (PCA) on the correlation matrix was used to derive DEPs, where the 24 variables represented the average hourly energy intake across UK adolescents. Use of correlation matrix meant that that energy intake at each time slots was given an equal weight - regardless of the absolute amount of energy intake- relative to other eating occasions. Linear regression accounting for complex survey design was then used to model the association between the derived main diurnal pattern scores and BMI or waist circumference. To account for singleton units, analyses were conducted using the

centering method. Models were adjusted by sex, ethnic group, age, socioeconomic status, and total daily energy intake. Missing data for outcome and some covariates (assumed to be missing at random) were imputed using multiple imputations (50 imputed data sets were created and fitted by utilising the "ice" and "mim" packages in Stata (StataCorp LP, College Station, TX). Sensitivity analyses were conducted comparing the coefficients derived from imputed data and the complete case analyses.
To identify differences in DEP between boys and girls, PCA was conducted separately for boys and girls. Additional sensitivity analysis excluding time slots with energy intake less than 210kJ (50kcal) were undertaken in accordance with the definition of eating occasions published previously [12].

Results

Three orthogonal diurnal patterns (Figure 1) were interpretable as (i) a linear contrast (explaining 8% of total system variation) between breakfast & earlier lunch vs. later lunch, late dinner & evening/night snack which was named 'phase shift' DEP (DEP1); (ii) a linear contrast (6.0% of system variation) between mid-morning snacks, late lunch and early dinner vs. breakfast, early and late morning snacks, early lunch , mid-afternoon snacks and late dinner, renamed 'early eating and grazing' DEP (DEP2); (iii) a linear contrast (6.0% of system variation) between late main meals vs. early main meals and night snacks which was named 'anti-phase shift' DEP (DEP3). In the univariate analyses, only age and total daily energy intake were significantly associated (positively) with DEP1. For every one year increment in age, DEP1 score increased by 0.26units (P<0.001), indicating that a late/shifted eating pattern be more pronounced with age. Similarly, for every 100kcal increase in total daily energy intake, DEP1 score increase by 0.1 units (P<0.001). Daily energy intake was also associated positively with DEP2, such that every 100kcal increase in daily energy intake was associated with 0.3 unit increase in DEP2 (P=0.001), indicating that a shift towards eating late as well as an early and frequent eating were related to a higher energy intake. There were no associations between any of the other covariates and remaining dietary patterns.

Figure 1. *Loadings of diurnal eating patterns on the variables indicating energy intake across 24 hours of the day, based on performing Principal Components Analysis on the correlation matrix.*



After adjusting for sex, ethnicity, age, socioeconomic status and total mean energy intake over the day, every 1 unit increase in DEP3 was associated with an increase of BMI and waist circumference of 0.28kg/m2 (95%CI 0.05-0.51; P=0.016) and 1.16cm (95%CI 0.60-1.72; P<0.001), respectively. No association between DEP1 and DEP2 and BMI or waist circumference were observed in either the crude models or the adjusted models.
Overall, similar DEPs were derived compared with the initial analyses which did not exclude any eating occasions (below 210KJ or 50kcal). Results from regression analyses using these DEPs did not differ from original findings. Furthermore no interpretative differences in DEPs between boys and girls were observed.

Discussion

Our findings are consistent in part with the results of a recent analysis of the DEDIPAC-study, which demonstrated that young adults have DEPs characterised by less distinct peaks in the morning and midday and a shift towards meal times later in the day [13]. Leech and colleagues [8] recently identified through latent class analysis three patterns in Australian adults: conventional, late meal and grazing which were more recently also found in British adults by the same statistical technique [14]. In the current sample of UK adolescents and using a different, purely descriptive statistical technique (PCA), a grazing pattern was expressed by DEP2 while the late meal behavior is captured by both DEP1 and DEP3.
Mounting evidence suggests that each of breakfast skipping and increased energy intake in the evening is individually related to increased prevalence of overweight and obesity in epidemiological studies [15, 16]. However, findings from randomised controlled trials appear to be conflicting [17]. This discrepancy between epidemiological studies and clinical observations could be ascribed to the fact that few observational studies have investigated the potential intercorrelations between energy intake at various eating occasions or time slots and whether the observed eating behaviours form part of a broader DEP, such that skipping of breakfast may be associated with greater energy intake later in the evenings. In the present study, DEP1, which could be described as a phase delay DEP characterised by breakfast skipping and a shift in energy intake to later in the day, was not associated to BMI or waist circumference. Adolescence is generally known to be associated with a phase-delay with the prevalence of evening chronotype being higher in this population subgroup compared to younger children or adults [15]. Accordingly, it could be speculated that the DEP1 may be reflective of the phase delay generally observed in adolescents. The latter is consistent with the observations made by Rossbach and colleagues who found that in adolescents, chronotype is associated with timing of energy intake, such that a later chronotype is related to a shift in energy intake towards later in the day [5]. More specifically, Rossbach and colleagues found that every one hour delay in chronotype is associated with higher prevalence of breakfast skipping and energy intake in the evening [5]. On the balance of this evidence, it could be further hypothesised that, despite adolescents reporting later energy intake throughout the day, this DEP may not be detrimental to their health since such DEP aligns with the circadian phase shift generally observed in adolescents [5]. In contrast, DEP3 may reflect the pattern of social schedules constraints, thus exhibiting a more structured eating pattern characterised by an earlier lunch and dinner. The constraints to comply with socially defined meal times are then counterbalanced by late evening and night-time energy intake consistent with the circadian phase delay observed in adolescents. It could be speculated that in adolescents, the DEP3 may reflect a behavioural rhythm that is not synchronised with adolescents' internal circadian phase. Such misalignment between the biological clock and the behavioural phase may then explain the observed positive association between DEP3 and BMI or waist circumference [18].

Conclusions

The present study identified three main DEPs in a nationally representative sample of adolescents in UK. Adolescents who tended to eat early main meals combined with night-time energy intake rather than slightly later main meals without night-time energy intake had higher BMI and waist circumference. Findings from the present study provide evidence to inform public health strategies and dietary guidelines regarding potential eating occasions that may warrant further attention. Further substantive and methodological research is required to explore the determinants of DEP and explore the impact of the context of eating and other factors in the development of specific DEP and on disentangling the direction of causality between DEP and health outcomes.

References

[1] Garaulet M., P. Gomez-Abellan, Timing of food intake and obesity: a novel association. Physiol Behav, 2014. 134: p. 44-50.

[2] Bo, S.; Musso, G.; Beccuti, G et al., Consuming more of daily caloric intake at dinner predisposes to obesity. A 6-year population-based prospective cohort study. PLoS One, 2014. 9(9): p. e108467.

[3] Wang, J.B.; Patterson, R.E.; Ang, A.;et al., Timing of energy intake during the day is associated with the risk of obesity in adults. Journal of Human Nutrition and Dietetics, 2014. 27: p. 255-262.

[4] Almoosawi, S.; Prynne, C.J.; Hardy, R. et al., Time-of-day and nutrient composition of eating occasions: prospective association with the metabolic syndrome in the 1946 British birth cohort. Int J Obes (Lond), 2013. 37(5): p. 725-31.

[5] Rossbach, S.; Diederichs, T.; Nothlings, U et al., Relevance of chronotype for eating patterns in adolescents. Chronobiology International, 2018. 35(3): p. 336-347.

[6] Maukonen, M.; Kanerva, N.; Partonen, T.; et al., Chronotype differences in timing of energy and macronutrient intakes: A population-based study in adults. Obesity (Silver Spring), 2017. 25(3): p. 608-615.

[7] Lucassen, E.A.; Zhao, X.; Rother, K.I et al., Evening chronotype is associated with changes in eating behavior, more sleep apnea, and increased stress hormones in short sleeping obese individuals. PLoS One, 2013. 8(3): p. e56519.

[8] Leech, R.M.; Worsley, A.; Timperio, A. et al., Diurnal eating patterns: a latent class analysis approach. Int J Behav Nutr Phys Act, 2017. 14(1): p. 3.

[9] Eicher-Miller, H.A.; Khanna, N.; Boushey, C.J.et al., Diurnal Dietary Patterns Derived among the Adult Participants of the National Health and Nutrition Examination Survey 1999-2004 Are Associated with Diet Quality. Journal of the Academy of Nutrition and Dietetics, 2016. 116(2): p. 283-291.

[10] Whitaker, R.C.; Wright, J.A.; Pepe, M.S. et al., Predicting obesity in young adulthood from childhood and parental obesity. N Engl J Med, 1997. 337(13): p. 869-73.

[11] Lennox A. , Fitt E., Whitton C. Appendix A to National Diet and Nutrition Survey. Results from Years 1-4 (combined) of the Rolling Programme (2008/2009-2011/12). Dietary data collection and editing. 2014.

[12] St-Onge, M.P.; Ard, J.; Baskin, M.L.; et al., Meal Timing and Frequency: Implications for Cardiovascular Disease Prevention: A Scientific Statement From the American Heart Association. Circulation, 2017. 135(9): p. e96-e121.

[13] Wittig, F.; Hummel, E.; Wenzler, G. et al., Energy and macronutrient intake over the course of the day of German adults: A DEDIPAC-study. Appetite, 2017. 114: p. 125-136.

[14] Mansukhani R., Palla L. Investigating eating time patterns in UK adults from the 2008-2012 NDNS. Proc Nut Soc, 2018, 77 (OCE1)

[15]Almoosawi, S.; Vingeliene, S.; Karagounis, L et al., Chrono-nutrition: a review of current evidence from observational studies on global trends in time-of-day of energy intake and its association with obesity. Proc. Nutr. Soc. 2016. 75(4): p. 487-500.

[16] Szajewska, H.; Ruszczyński, M., and nutrition, Systematic review demonstrating that breakfast consumption influences body weight outcomes in children and adolescents in Europe. Crit. Rev. Food Sci. Nutr. 2010. 50(2): p. 113-119.

[17] Milanes, J.E.; Allison, D.B.; Brown, A.W et al., Effect of breakfast eating versus breakfast skipping on obesity related anthropometry: a systematic review. Energy Nutr. Metab. 2016. 30(1_supplement): p. lb394-lb394.

[18] Muñoz, J.; Cañavate. R.; Hernández. C.M.et al., The association among chronotype, timing of food intake and food preferences depends on body mass status. Eur. J. Clin. Nutr. 2017. 71(6): p. 736.

ANALISI E VALUTAZIONE DI APPROPRIATEZZA NEL PERCORSO DI CURA DEL CARCINOMA DELLA MAMMELLA IN REGIONE PIEMONTE. RISULTATI DELLO STUDIO E.PIC.A

Ricotti Andrea[1,2], Mistrangelo Marinella[3], Berchialla Paola[4], Dalmasso Paola[5], Pierini Andrea[6], Bono Danilo[7], Moirano Fulvio[8], Bertetto Oscar[3], Messori Ioli Giovanni[9]

[1] *Azienda Ospedaliera Ordine Mauriziano di Torino*

[2] *Scuola in Statistica Sanitaria e Biometria, Dipartimento di Scienze della Sanità Pubblica e Pediatriche, Università di Torino*

[3] *Dipartimento Funzionale Interaziendale Interregionale Rete Oncologica Del Piemonte e Della Valle D'Aosta*

[4] *Dipartimento di Scienze Cliniche e Biologiche, Università di Torino*

[5] *Dipartimento di Scienze della Sanità Pubblica e Pediatriche, Università di Torino*

[6] *Market Access, Roche S.p.A*

[7] *Direzione settore programmazione* sanitaria e socio-sanitaria dell'Assessorato alla Sanità, Regione Piemonte

[8] *Azienda per la Tutela della Salute Sardegna*

[9] *Azienda Sanitaria Locale di Asti*

Introduzione

La sostenibilità del sistema sanitario italiano è minata da diversi fattori: l'aumento della domanda di servizi e prestazioni da parte di cittadini e pazienti, il progressivo invecchiamento della popolazione, il costo crescente delle innovazioni, con particolare riferimento a quelle farmacologiche. Tutti questi elementi hanno sicuramente un elevato impatto socio-economico [1].
Per far fronte a queste problematiche il Decreto del Presidente del Consiglio dei Ministri (DPCM) del 12 gennaio 2017 pone le premesse per ridefinire i nuovi Livelli Essenziali di Assistenza (LEA), tenendo in considerazione anche le evidenze di efficacia ed il concetto di value, definito come il ritorno in termini di salute delle risorse investite in sanità (rapporto tra outcome e costi), pur senza esplicitare la metodologia da utilizzare per la loro definizione.
Governare l'estrema complessità delle organizzazioni sanitarie avendo come duplice obiettivo sia la promozione dell'efficacia/qualità assistenziale che dell'efficienza, in relazione alla sostenibilità economica, non è attuabile in modo scontato. Inoltre, si deve aggiungere il gap esistente tra professionisti (libertà professionale incondizionata) e manager (bilancio aziendale) [2].
È stato suggerito come alcuni miglioramenti si potrebbero registrare passando da una *Governance* basata sui budget a "silos" e quindi a singole variabili per lasciare strada a una gestione che tenga conto dell'intero percorso diagnostico terapeutico e assistenziale (PDTA) e quindi ad un approccio multidimensionale che permetterebbe di porre realmente il paziente al centro del suo percorso di malattia.
A tal fine è importante disporre di un sistema specifico di indicatori basati su dati di Real World Evidence per monitorare la qualità dell'assistenza sanitaria, in termini di sicurezza, efficacia, appropriatezza, partecipazione degli utenti, equità, efficienza [2]. Questo permetterebbe di attuare un piano di disinvestimento da sprechi e inefficienze, basato non solo su azioni puntuali di spending review, ma che preveda anche interventi strutturali e organizzativi di riallocazione di risorse [1].
A tal proposito la *Clinical Governance* fornisce, attraverso alcune delle sue aree (Evidence Based Medicine, Information & Data Management, Linee Guida e PDTA, Audit clinici), strumenti operativi utili per declinare azioni concrete mediante le quali le organizzazioni sanitarie si rendono responsabili del miglioramento

continuo della qualità dei servizi e del raggiungimento/mantenimento di elevati standard assistenziali, stimolando la creazione di un ambiente che favorisca l'eccellenza professionale [2].
Si è proceduto in tale direzione andando a esplorare un percorso di cura in ambito oncologico, mediante lo studio denominato E.Pic.A. (Appropriatezza Economica del Percorso integrato di cura) condotto con il supporto scientifico della Rete Oncologica Piemonte/Valle d'Aosta e basato sulle aree della *Clinical Governance* sopra evidenziate. L'approccio metodologico del progetto E.Pic.A si è basato sullo sviluppo di un modello di analisi, che a partire dal percorso post-diagnosi della paziente con tumore della mammella, è stato alimentato da dati di *Real World Evidence (RWE)*, estrapolati da database amministrativi e clinici delle aziende sanitarie locali.

Obiettivi

Il presente studio è stato condotto nell'ambito del progetto E.Pic.A con l'obiettivo di misurare un set di indicatori (Key Performance Indicators - KPI) proposti da Altini et al. [3] e selezionati da un panel di esperti sulla base di evidenze di letteratura per valutare le fasi (diagnostica, chirurgica e di chemio/ormo terapia) del percorso diagnostico terapeutico delle pazienti affette da tumore maligno della mammella.

Metodi

Lo studio è stato condotto retrospettivamente su un gruppo di pazienti affette da tumore maligno della mammella afferite per intervento chirurgico alle strutture piemontesi dal 1 gennaio al 31 dicembre 2015.
La coorte di casi incidenti è stata selezionata utilizzando il protocollo pubblicato da Altini M. et a. [3], basandosi sui dati presenti sulla scheda di dimissione ospedaliera (SDO).
La stadiazione del carcinoma mammario è stata definita mediante l'utilizzo della classificazione Tumor-Node-Metastasis (TNM) presente sulla banca dati della Scheda computerizzata per il controllo della Qualità del Trattamento del carcinoma Mammario (SQTM).
I dati per il calcolo dei KPI sono stati estratti andando a eseguire un cross-linkage tra i flussi SDO, prestazioni ambulatoriali, distribuzione per conto e farmaceutica convenzionata utilizzando la metodologia presentata da Altini M. et al. [3]. I risultati sono presentati con i relativi intervalli di confidenza al 95% calcolati con l'approssimazione di Wilson.

Risultati

È stata selezionata una coorte di 885 pazienti di sesso femminile di età superiore a 45 anni con la classe più rappresentativa compresa tra i 65 e 69 anni.
La coorte è stata ripartita in base allo stadiazione TNM: il 15% (129) in stadio 0, il 51% (453) in stadio I, il 28% (247) in stadio II, il 6% (55) in stadio III e una sola paziente in stadio IV.
Gli interventi di mastectomia alla data di immissione nella coorte sono stati il 16% (140) contro l'84% delle mammectomie (745). Il 29% (260) delle pazienti sono state sottoposte a chemioterapia, mentre il 65% (576) alla terapia ormonale. Il 74% (658) delle donne sono state sottoposte a radioterapia.
In tabella 1 sono riportati i KPI misurati sulla coorte di studio. Nei 60 giorni pre-intervento chirurgico e nei 60 giorni post-intervento chirurgico rispettivamente il 22% (154) e il 40% (279) delle pazienti in stadio I o II sono state sottoposte ad almeno uno tra i seguenti accertamenti: ECO epatica, TAC, RMN (eccetto torace o mammella), scintigrafia ossea o PET. Nei 365 giorni post-intervento chirurgico (escludendo i primi due mesi analizzati nel KPI 2) il 31% (218) delle pazienti in stadio I o II è stata sottoposta ad almeno una delle diagnostiche sopra citate.
Complessivamente, il 7% (10) delle pazienti sottoposte a mastectomia, come primo intervento, hanno eseguito uno svuotamento ascellare nei 90 giorni successivi l'intervento chirurgico; inoltre, nessuna delle 140 donne è stata sottoposta ad una ricostruzione della mammella nei 90 giorni successivi alla mastectomia.

Il 5% (35) delle pazienti sottoposte a mammectomia sono state sottoposte a re-intervento chirurgico.
Il 37% (96) delle pazienti sottoposte a chemioterapia adiuvante hanno iniziato il primo ciclo oltre 60 giorni dall'ultimo intervento chirurgico; mentre il 37% (211) delle pazienti sottoposte ad ormonoterapia hanno iniziato il trattamento oltre 60 giorni dall'ultimo intervento chirurgico.
Il 9% (43) delle pazienti non sottoposte a terapia adiuvante hanno eseguito la radioterapia oltre 180 giorni dall'ultimo intervento chirurgico; mentre il 18% (33) delle pazienti sottoposte a terapia adiuvante hanno eseguito la radioterapia oltre 180 giorni dall'ultimo intervento chirurgico.

***Tabella 1**. Descrizione dei KPI e risultati ottenuti dall'analisi con relative coorti e valori target.*

KPI	Descrizione del KPI	Popolazione	Risultati % (freq. assoluta)	Target	IC 95%
KPI 1	Percentuale di pazienti con tumore mammario di stadio I e II (stadiazione post-chirurgica secondo TNM) che pre-intervento (60 giorni) hanno eseguito uno tra i seguenti esami diagnostici: ECO epatica, TC, RMN, Scintigrafia ossea o PET .	Stadio I e II (n=700)	22% (154)	Non superare il 5%	19%-25%
KPI 2	Percentuale di pazienti con tumore mammario di stadio I e II (stadiazione post-chirurgica secondo TNM) che post-intervento (60 giorni) hanno eseguito uno tra i seguenti esami diagnostici: ECO epatica, TC, RMN, Scintigrafia ossea o PET .	Stadio I e II (n=700)	40% (279)	Non superare il 5%	36%-43%
KPI 3	Percentuale di pazienti con tumore mammario di stadio I e II (stadiazione post-chirurgica secondo TNM) che post-intervento (365 giorni, esclusi i primi 60 giorni) hanno eseguito uno tra i seguenti esami diagnostici: ECO epatica, TC, RMN, Scintigrafia ossea o PET .	Stadio I e II (n=700)	31% (218)	Non superare il 5%	28%-35%
KPI 4	Percentuale di pazienti che hanno ricevuto un intervento di svuotamento ascellare in intervento successivo alla mastectomia entro 3 mesi.	Interventi di mastectomia (n=140)	7% (10)	Tendere allo 0%	4%-13%
KPI 5	Percentuale di pazienti che hanno ricevuto un re-intervento entro 120 giorni da intervento chirurgico conservativo per tumore maligno alla	Interventi di mammectomia (n=754)	5% (35)	Tendere allo 0%	

	mammella (con identificazione del tempo mediano al re-intervento).				3%-6%
KPI 6A	Percentuale di pazienti che iniziano la chemioterapia adiuvante oltre 60 giorni dal primo intervento	Sottoposte a chemioterapia (n=260)	37% (96)	Tendere allo 0%	31%-43%
KPI 6B	Percentuale di pazienti libere da malattia che iniziano la terapia ormonale oltre 60 giorni dal primo intervento	Sottoposte a ormonoterapia (n=576)	37% (211)	Tendere allo 0%	33%-41%
KPI 7A	Percentuale di pazienti (sottoposti a terapia adiuvante) con resezione parziale che avviano il trattamento radioterapico oltre 180 giorni dalla data dell'ultimo intervento	Sottoposte a radioterapia senza previa terapia adiuvante (n=478)	9% (43)	Tendere allo 0%	7%-12%
KPI 7B	Percentuale di pazienti (sottoposti a terapia adiuvante) con resezione parziale che avviano il trattamento radioterapico oltre 270 giorni dalla data dell'ultimo intervento	Sottoposte a radioterapia previa terapia adiuvante (n=180)	18% (33)	Tendere allo 0%	13%-25%

Conclusioni

Sono stati utilizzati come valori di riferimento per la valutazione dei risultati ottenuti dei KPI le percentuali indicate nell'articolo di Altini M. et al. [3].
Durante la fase di diagnostica, è stato riscontrato un sovra-utilizzo delle metodiche sia nei due mesi che precedono l'intervento (KPI 1) in misura pari al 22% che per ciò che concerne la fase post-intervento (KPI 2) con una percentuale del 40% di prestazioni non appropriate eseguite (over-use). Anche per ciò che concerne la fase di follow-up è stato riscontrato un utilizzo del 31% ben al di sopra del 5% fissato dal panel di esperti di Altini M. et al. [3]. Tali risultati meritano un approfondimento idoneo attraverso *Audit* clinici.
L'analisi inerente all'attività chirurgica ha evidenziato con il KPI 4 che una quota di pazienti pari al 7% è stata sottoposta allo svuotamento dei linfonodi ascellari entro tre mesi dall'intervento di mastectomia. Tale fenomeno può essere riconducibile all'assenza in alcune realtà aziendali della diagnostica in sede operatoria finalizzata al riconoscimento della positività del linfonodo sentinella; altre strutture, invece, utilizzano una metodica di diagnostica intra-operatoria che nonostante il buon grado di sensibilità, può generare dei falsi negativi. Solo una minima parte delle strutture in Regione è dotata, infatti, di uno strumento ad alta sensibilità come la *One Step Nucleic acid Amplification (OSNA)* con elevata accuratezza diagnostica.

Sempre per ciò che riguarda la parte chirurgica, il KPI 5 ha mostrato il valore del 5% di re-interventi, in linea con i dati di lettura internazionale.
Andando ad analizzare i valori inerenti all'area della terapia adiuvante, il KPI 6A ha evidenziato che il 37% delle pazienti inizia la terapia oltre 60 giorni dalla data dell'ultimo intervento. Lo stesso vale per la terapia ormonale, infatti anche in questo caso il 37% delle donne inizia oltre i 60 giorni. Questo tipo di trattamento ha una forte correlazione con gli outcome e gli esiti di salute essendo l'efficacia della terapia adiuvante tempo dipendente.
Il percorso radioterapico (KPI 7) presenta sicuramente criticità inferiori anche se il 18% delle donne sottoposte terapia adiuvante inizia la radioterapia oltre i 270 giorni, mentre solo il 9% delle non trattate con la chemioterapia adiuvante sfora i 180 giorni.
In conclusione, lo studio, evidenziando nel complesso dei buoni livelli di qualità delle cure, ha messo in luce una serie di inappropriatezze nel percorso delle pazienti affette da carcinoma della mammella andando a sottolineare il concetto di come sia opportuno lavorare per rendere i clinici e i decisori più attenti ad efficacia ed efficienza.
Analizzando i risultati è emersa l'opportunità di pianificare delle discussioni strutturate con i clinici, al fine di giungere a indicazioni condivise in alcuni ambiti del percorso di cura là dove siano ancora poco definite le indicazioni di letteratura.
È inoltre emerso come "fare appropriatezza" non significhi unicamente risparmiare, in quanto tale situazione è abitualmente associata all'area dell'over-use; è infatti importante porre altrettanta attenzione anche all'area dell'under-use, concentrandosi quindi su pazienti che per il loro reale bisogno di salute dovrebbero ricevere un trattamento, ma che il sistema non eroga.

Bibliografia

[1] 2° Rapporto GIMBE sulla sostenibilità del SSN - presentato a Roma il 6 giugno 2017

[2] Il Governo Clinico nelle Aziende Sanitarie Versione 1.4 del 6 febbraio 2009 – GIMBE

[3] Altini M., Balzi W., Massa I. et al., Key performance indicators for monitoring the integrated care pathway in breast cancer: the E.Pic.A. project. AboutOpen – HTA & Market Access, 2019 Mag; 5(1): 31-38

DIFFUSE AND INTESTINAL HISTOTYPES PREVAIL RESPECTIVELY IN YOUNG AND OLD GASTRIC CANCER CASES

Verlato Giuseppe[1], Torroni Lorena[1], Bencivenga Maria[2], Giacopuzzi Simone[2], Antonucci Nadia[1], Vesentini Roberta[1], Negri Eva[3], La Vecchia Carlo[4], De Manzoni Giovanni[2]

[1] *Unit of Epidemiology & Medical Statistics, Dept. of Diagnostics & Public Health, University of Verona, Italy*

[2] *General and Upper GI Surgery Division, University of Verona, Verona, Italy*

[3] *Dept. of Biomedical & Clinical Sciences, University of Milan, Milan, Italy*

[4] *Dept. of Clinical Sciences & Community Health, University of Milan, Milan, Italy*

Background

Hereditary gastric cancer is rather rare, representing 1-3% of all cases [1]. It is usually diagnosed under 50 years of age, and it mostly has a diffuse histotype.

Aims

The present study aims at investigating clinico-pathologic characteristics of gastric cancer, arising at different ages.

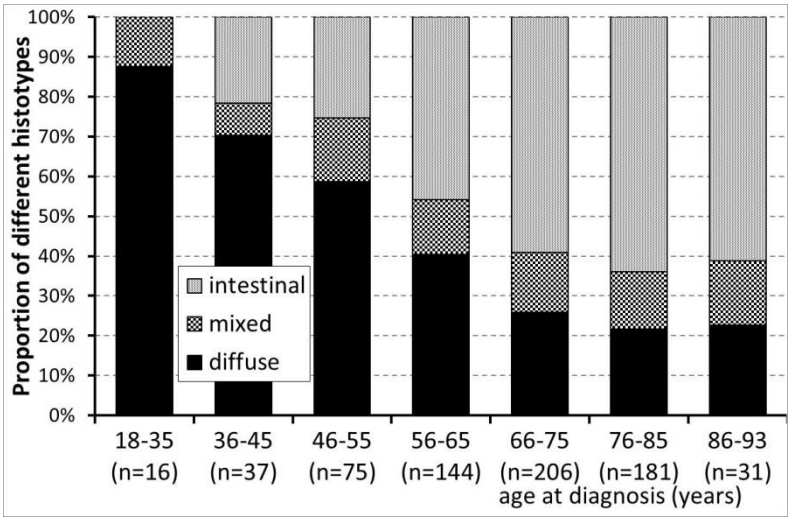
Methods

The present study was performed on 721 patients, undergoing gastrectomy for gastric cancer at the Division of Upper GastroIntestinal Surgery of Verona from January 1998 to May 2017. Of these 57 were aged 18-45 years at surgery, while 663 were aged 46-93 years.

Results

Patients aged 18-45 years were mostly females (61%) and mostly presented a diffuse histotype (75%), while older patients were mostly men (63.4%) and mostly had an intestinal histotype (53.7%) ($p<0.001$). As shown in Figure 1, the transition from diffuse to intestinal histotype progressively occurred with advancing age at diagnosis. The two age groups did not significantly differ as regards tumor site, or yTNM.
Patients aged 18-45 years mostly underwent extended (49%) or superextended (46%) lymphadenectomy, while a less aggressive surgical approach was used in older patients ($p=0.002$). Of these, a substantial proportion (23.2%) was treated with limited lymphadenectomy while only 28.3% with the superextended procedure.
Five-years overall survival was the highest in the 18-45 age group (55%, 95% CI 39-69%), intermediate in people aged 46-75 years (42%, 95% CI 37-47%) and the lowest in people aged over 75 years (29%, 95% CI 22-36%) (log-rank test: $p<0.001$).

Figure 1. *Proportion of different histotypes in gastric cancer cases, diagnosed at different ages.*



Conclusions

Gastric cancers, diagnosed at 18-45 years, has different features with respect to gastric cancer diagnosed at older age. The shift from diffuse to intestinal histotype, observed with increasing age at diagnosis, likely reflects a parallel shift from genetic to environmental determinants of cancer.

References

[1] Ekström A.M., Serafini M., Nyrén O et al., Dietary antioxidant intake and the risk of cardia cancer and noncardia cancer of the intestinal and diffuse types: a population-based case-control study in Sweden. Int J Cancer., 2000; 87(1):133-40

UN PROGETTO LAVORO DI INTERNAZIONALIZZAZIONE E ARMONIZZAZIONE DELLA RICERCA SU FARMACI INNOVATIVI, INNOVAZIONE DELLE RICERCHE EPIDEMIOLOGICHE, GENETICHE E DI SVILUPPO GLOBALE DELLA RICERCA IN EUROPA / ITALIA AI PAESI ORIENTALI / IRAN AZERBAIGIAN / KAZAKISTAN / CINA 2015 2025.

Franco Naccarella *, Domenico Criscuolo *, Giulio He **, Bin He **, Sophie Qi *, Lei Li $$ Janxiun Liu $$, Lei Sun *, Weynan Li ***, Youlin Qiao &&& Meray Abbasov &; Zeynal Akparov e Siamak Saber $, Pirooz Hebrahimi $. Reza Saedi, Antonella Tognin *, Mahmood Jamei Masjed *

*Unità di Farmacologia di Cardiologia Bologna, ***

*Gruppo TCM di Roma e Pechino, ****

BGI, & ANAS Risorse genetiche Baku, $

Theran Genetic Group, Mashhad Razavi Hospital £, CACMS Beijing $$,

China and Beijing CICAMS Union Medical College &&&

Obiettivi

Le nostre personali esperienze e ricerche internazionali hanno dimostrato che

1 differenze nelle mutazioni dei geni della stessa malattia, in Cardiologia (1,2,4), Oncologia, Neurologia (7,8), Ortopedia (6) Neonatologia (1-8) esistono in diverse popolazioni nell'East Side rispetto al West Side del mondo, come dimostrato per la Cina (Carcinoma Polmonare), l'Iran (Cerebellar Atrophia e Cardiovascular diseases) (1,3-6,7,8). **2** I targets genetici della terapia stanno diventando un approccio più razionale e più fattibile in molte di loro (1. 5), **3** Le differenze nella distribuzione e produzione di farmaci innovativi rappresentano un importante problema internazionale di ricerche e produzioni accettabili in tutto il mondo, unitamente al processo di armonizzazione per GMP, GLP e GCP. Così è stato avviato un Accademia del Training farmaceutico internazionale, in paesi selezionati per richiesta locale, (3.5)

Metodi

1 Un nuovo approccio alle ricerche epidemiologiche, confrontando i dati delle popolazioni non caucasiche a quelle caucasiche nelle ricerche epidemiologiche, è necessario, disposto a lavorare in progetti e laboratori di ricerca internazionali congiunti, saranno proseguiti **2** laboratori genetici aperti a screening precoci internazionali e BIG BANK DI DATI internazionale, **3** scuole di International Pharma Train con EMA, CFDA, IDO, AFDA, Azerimed e Avromed e Università internazionali sono state proposte

Risultati

1 Aspettando risultati preliminari per il **MAECI MOST (Ministero degli Affari Esteri Cooperazione Internazionale dell'Italia - e Ministro della Scienza e della Tecnologia di PrChina)**, Joint Sceientific and Economics Ventures e Laboratori di ricerca a conduzione bilaterale in Cina e Italia **2** Ancora attivo un dialogo continuo tra **Medicina tradizionale cinese (TCM) e medicina occidentale** su basi comuni di biologia molecolare e differenze genetiche di popolazioni e diversi profili farmacocinetici e farmacodinamici del metabolismo e distribuzione dei farmaci, in diverse popolazioni. **3** Lavorando per i dati preliminari del **MAECI ANAS (Ministero degli Affari Esteri Cooperazione Internazionale dell'Italia - Accademia Nazionale delle Scienze dell'Azerbaigian) MOU,** che istituisce un'Accademia Genetica per la Formazione

Internazionale e un Centro Diagnostico Genetico a Baku, per i Paesi circostanti del Medio Oriente e dell'Asia Centrale, compresi i genetisti iraniani e gli esperti cinesi di BGI (Beijing Genomic Institute)

Conclusioni

Nel mondo globalizzato di oggi, data la differenza tra paesi, in prevalenza, in incidenza e caratteristiche genetiche delle malattie

1 Le ricerche epidemiologiche comparate sono necessarie e obbligatorie per comprendere la realtà e la specificita' dei problemi

2 per approcciare correttamente i trattamenti diagnostici e specifici della stessa malattia, e

3 Nell'ambito di un approccio scientifico interculturale cooperativo, dovrebbe essere intrapresa l'armonizzazione nelle procedure, nella identificazione, fabbricazione e nella garanzia della qualità delle ricerche e la convalida e la produzione di farmaci. da attuare e continuare. (3)

References

[1] Jonhson Albert, Lewis Morgan, Raff Roberts. Molecular Biology of the Cell. Sixth Edition. Garland Science 2018

[2] Gussak Ihor, Antzelevch Charles, Editors, Electrical Diseases of the Heart. Vol 1 Basic Foundations and Primary Electrical Diseases of the Heart. Second Edition Springer 2013

[3] Rosier Jan A., Maartens Mark A., Thomas Josse R. Global New Drugs Development. Wiley Blackwell 2014

[4] International Pedriatric Cardiology Update. Carminati Mario, Frigiola Alessandro, Naccarella Franco, English Version Istituto Polilclinico San Donato 2019

[5] American Pharmacist Association Lexicomp. Drugs Information Handbook for Oncology. An Extensive Guide to Combination Chemotherapy Regimens. 15 th Edition Wolter and Kluwier 2018

[6] Lieberman Jay R., Friedlaender Gary E., Bone Regeneration and Repair. Biology and Clinical Applications Human Press. 2010

[7] AAVV. Cerebellar Atrophy, Visual Impairment, And Psychomotor Retardation; Phenotype-Gene Relationships

[8] AAVV Cerebellar degeneration | Genetic and Rare Diseases Information Center (GARD) – an NCATS Program

file:///E:/045%20MASJED%20MICHELUCCI/567%20Cerebellar%20degeneration%20_%20Genetic%20and%20Rare%20Diseases%20Information%20Center%20(GARD)%20–%20an%20NCATS%20Program.pdf



Atti | Metodi Biostatistici

POPULATION SIZE ESTIMATION WITH INTERVAL-CENSORED COUNTS AND EXTERNAL INFORMATION: PREVALENCE OF MULTIPLE SCLEROSIS IN ROME

Farcomeni Alessio[1], Danilo Alunni Fegatelli[1], Arianna Di Rocco[1], Vestri Anna Rita[1]

Department of Public Health and Infectious Diseases, Sapienza – University of Rome

Introduction & Aims

Population size estimation often proceeds by evaluating overlap of two or more lists of subjects. This usually involves specifying a log-linear model for the incomplete contingency table of cross-counts, where the cells corresponding to the situation of not being included in any list are missing. The parameter of interest is the population size *N*, which corresponds to the sum of observed and unobserved counts. In this work we focus on the case in which some observed counts are interval-censored, that is, only upper and lower bounds for the true underlying counts are available. This happens for instance when some individuals not belonging to the population of interest might have been wrongly recorded, and it is rather common but often overlooked. It is also related to the situation of open populations in which the population size varies during the observation period.
Our motivating example regards estimation of prevalence of multiple sclerosis (MS) in the metropolitan area of Rome. Six centers provided us with lists of their patients, and several patients appear in more than one list as they have been treated or visited in different centers after diagnosis. All centers also provided us with information regarding gender, age at prevalence date, date of last visit. One center (MS unit of Fondazione Santa Lucia) also carefully verified that patients were alive at the prevalence date, but the other ones did not. For patients not appearing in the Santa Lucia list we can guarantee alive status at the prevalence date only for subjects whose last visit occurred afterwards. We thus have five out of six lists with censored counts. Some patients in each list are nevertheless guaranteed to be alive at the prevalence date, therefore true counts not only have upper bounds (the total number of patients in the list) but also lower bounds (the total number of patients that could be confirmed to be alive at the prevalence date).

Methods

Suppose there are n cells in an incomplete contingency table, of which part are observed precisely, part are completely unobserved, and the remaining are interval-censored. For each censored cell we have observed upper and lower bounds. We specify a log-linear model, including possibly discrete covariates, for cell-specific means. We use either Poisson/Multinomial or Negative-Binomial assumptions (to account for over-dispersion). We fit the model in a Bayesian framework where the full conditional for the upper and lower bounds are clearly given by uniform distributions in the intervals below and above the currently sampled unobserved or censored counts. This allows us to adjust for censoring in a natural way. For model selection we use Structured Stochastic Search Variable Selection [1], and provide estimates through model averaging. In order to increase efficiency we also specify a different methodology in which we use external information on censoring. We argue censored counts to be independent of log-linear parameters conditionally on external information, so that their full conditional reduces to a binomial distribution with success probability given by the exogenously specified, cell-specific, survival rate.

Results

In this section we proceed with estimation of the number of patients with multiple sclerosis, and consequently prevalence estimation, for the metropolitan area of Rome. According to official (even if outdated) prevalence estimates the total number of patients with MS in the metropolitan area of Rome is slightly above 4,000. Results obtained ignoring the issue of censored counts, with log-linear models with unobserved heterogeneity in a frequentist framework, lead to *N=4,610 [2]*. This is reasonable but probably too large. The

resulting estimated prevalence would be 143 per 100,000; which is substantially larger than the largest prevalence ever reported for the area. As external information we use official age-gender-specific mortality rates for the general Italian population as published by ISTAT, adjusted using standardized mortality ratios specific to multiple sclerosis [3]. We begin our new analysis by using existing methods: Chao's estimators are either too large when ignoring censoring (*N=4,812*) or too small when using only patients guaranteed to be alive (*N=3,082*). Ignoring censoring leads to *N=4,229*, with a very large *95%* Credibility Interval (CI): *2,884-5,178*; removing censored cells is not feasible. Finally, existing approaches [4] based only on upper bounds lead to *N=184*, which is clearly wrong. Our approach with use of lower bounds leads to *N=4,333 (4,153-4,516)*, which is more credible than the previously published figure [2] and has a reasonably narrow 95% CI. Use of external information further improves the estimate to a final *N=4,423* (*4,300-4,545*). We note that previously published *N=4,610* lies outside the credibility intervals, hence this estimate can be deemed to have included dead patients in the prevalence estimate.

Conclusions

As confirmed also by an extensive simulation study in the accompanying paper, use of lower bounds and external information may greatly improve efficiency of the population size estimator in information scarce environments, where several lists provide censored cells. Surprisingly enough, even misspecified external information improves efficiency, up to a certain extent. Naively ignoring the censoring issue or removing censored cells clearly provides possibly severely biased estimates.

References

[1] Farcomeni A (2010) Bayesian Constrained Variable Selection. *Statistica Sinica*, **20**, 1043-1062.

[2] Farcomeni A, Cortese A, Sgarlata E, Alunni Fegatelli D, Marfia GA, Buttari F, et al. (2018) The prevalence of multiple sclerosis in the metropolitan area of Rome: a capture-recapture analysis. *Neuroepidemiology,* **50**, 105-110.

[3] Scalfari A, Knapperts V, Cutter G, Goodin DS, Ashton R, Ebers GC. (2013) Mortality in patients with multiple sclerosis. *Neurology*, **81**, 184-192.

[4] Overstall AM, King R, Bird SM, Huchinson SJ, Hay G. (2014) Incomplete contingency tables with censored cells with application to estimating the number of people who inject drugs in Scotland. *Statistics in Medicine*, **33**, 1564-1579.

EARLY PREDICTION OF EVIDENCE OF DISEASE ACTIVITY IN MULTIPLE SCLEROSIS PATIENTS

Trivelli L.[2], Mallucci G.[1], Colombo E.[1], Trojano M.[3], Amato MP.[4], Zaffaroni M.[5], Iaffaldano P.[3], Razzolini L.[4], Saraceno L.[5], Montomoli C.[2], Bergamaschi R.[1]

[1] *IRCCS Mondino, Pavia, Italy.*

[2] *Unit of Biostatistics and Clinical Epidemiology, Department of Public Health, University of Pavia, Italy.*

[3] *Department of Neurology, University of Bari, Bari, Italy.*

[4] *Department NEUROFARBA, University of Florence, Florence, Italy.*

[5] *ASST della Valle Olona, Hospital of Gallarate, Multiple Sclerosis Center, Italy*

Introduction

The common clinical practice needs a simple instrument to support therapeutic decisions for multiple sclerosis (MS) patients. Many studies have already analysed this problem with longitudinal methods; only few studies have been conducted with the aim to classify and predict the status of the patient in the early phase of the disease.

Objective

To create a simple and interactive prognostic tool easy to use and to predict the loosing of "no evidence of disease activity" (NEDA-3) status of patients with a clinically isolated syndrome (CIS) strongly suggestive for the onset of MS. This tool will be implemented to support clinical decisions and to inform patients about the risk of personal conditions.

Methods

We analysed demographic characteristics, clinical onset data, brain magnetic resonance imaging (MRI), evoked potentials (EPs), cerebrospinal fluid oligoclonal bands (OBs) of 279 CIS patients (183 female, 96 male; median age at disease onset 32.3 years range 12-56). The outcome of the proposed model is the NEDA-3 status loss in twelve months from the onset date (NEDA-3 to 'evidence of disease activity' -EDA). NEDA-3 is a combined measure defined as absence of either a clinical relapse, or disability worsening, or radiological activity [1]. To guarantee the best interpretability-accuracy trade-off, we used a mixed Frequentist/Bayesian approach. Firstly, in the feature selection stage, we used three stepwise logistic regression method (backward, forward and both, conditioning to the previous result) jointly based on Akaike information criterion. Secondly, in order to asses the marginal posterior distributions of the odds ratios, we have used a noninformative/weakly-informative Bayesian logistic regression model given the stepwise regression results. To simulate marginal posterior distributions, we used Hamiltonian Monte Carlo Markov Chains [2] sampling approach (HMCMC) based on 100000 for seven chains (after a warmup period of the 50% of the chain length) samples. Each chain reached the stationary status given a unitary $\hat{R}$ Convergence statistic [2] and "ad hoc" graphical trace plot analysis. Thirdly, to ensure marginal posterior insensitivity to the prior choice, we asses the same model with 100 different prior distributions (t-student prior distributions on parameters with degrees of freedom between 1 and 100). Also, we have evaluated posterior predictive checks, posterior probabilities of the marginal posterior distribution of parameters, marginal posterior predictive distribution of the probability of NEDA/EDA status of the patients. To evaluate the in-sample prediction accuracy of the model, we have used confusion matrix, receiver operating characteristic (ROC) and area under the receiver operating characteristic (AUC). To run the entire analysis we used R program [3], Stan program with packages rstan [4] and rstanarms [6] and R packages MASS [7] (to run the stepwise feature selection) and ggplot2 [8] (to create plots with functions).

Results

Selected variables by the frequentist step – AIC approach, that's are included in the Bayesian logistic regression model provide these marginal posterior distribution of odd ratios:

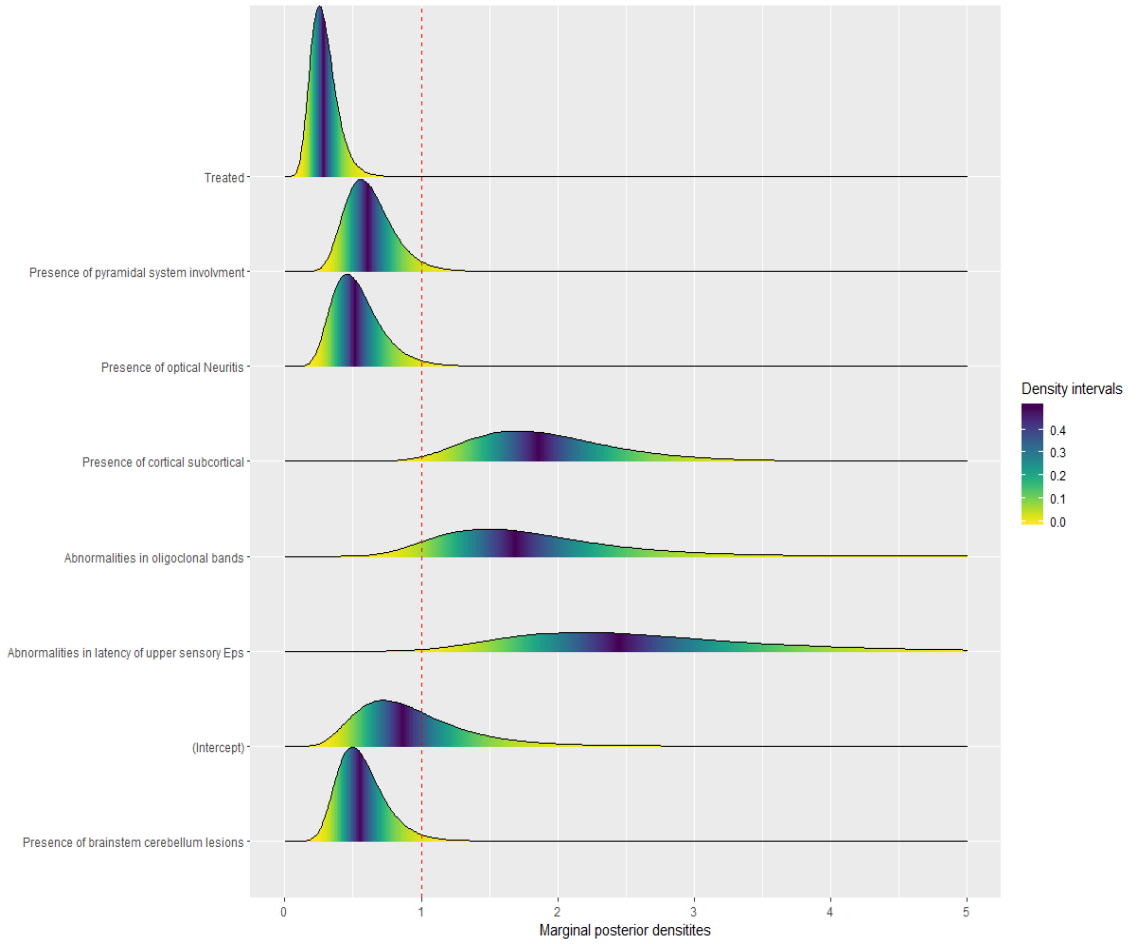
1. Dichotomous, absence or presence of therapy with disease modifying treatments at onset (mean = 0.288, sd = 0.094, percentile$_{0.025}$ = 0.143, percentile$_{0.25}$ = 0.221, median = 0.275, percentile$_{0.75}$ = 0.341, percentile$_{0.975}$ = 0.511, PP$_{>1.025}$ ~ 0 , PP$_{<0.975}$ ~ 1).
2. Dichotomous, noninvolvment or involment or of the pyramidal system at the onset (mean = 0.624 sd = 0.176, percentile$_{0.025}$ = 0.345, percentile$_{0.25}$ = 0.498, median = 0.601, percentile$_{0.75}$ = 0.725, percentile$_{0.975}$ = 1.033, PP$_{>1.025}$ = 0.027, PP$_{<0.975}$ = 0.960).
3. Dichotomous, absence or presence of optical neuritis at onset (mean = 0.535, sd = 0.184, percentile$_{0.025}$ = 0.260, percentile$_{0.25}$ = 0.403, median = 0.507, percentile$_{0.75}$ = 0.635, percentile$_{0.975}$ = 0.973, PP$_{>1.025}$ = 0.017, PP$_{<0.975}$ = 0.975).
4. Dichotomous, absence or presence, at MRI, of cortical and subcortical lesions at onset (mean = 1.930, sd = 0.552, percentile$_{0.025}$ = 1.080, percentile$_{0.25}$ = 1.537, median = 1.853, percentile$_{0.75}$ = 2.239, percentile$_{0.975}$ = 3.220, PP$_{>1.025}$ = 0.984, PP$_{<0.975}$ = 0.010).
5. Dichotomous, absence or presence of abnormal oligoclonal bands at onset (mean = 1.795, sd = 0.657, percentile$_{0.025}$ = 0.852, percentile$_{0.25}$ = 1.329, median = 1.683, percentile$_{0.75}$ = 2.134, percentile$_{0.975}$ = 3.384, PP$_{>1.025}$ = 0.923, PP$_{<0.975}$ = 0.058).
6. Centred continuous (to make intercept useful) age (given a sample mean of 32 years) at onset (mean = 0.969, sd = 0.014, percentile$_{0.025}$ = 0.941, percentile$_{0.25}$ = 0.960, median = 0.969, percentile$_{0.75}$ = 0.979, percentile$_{0.975}$ = 0.998, PP$_{>1.025}$ ~ 0 , PP$_{<0.975}$ = 0.637).
7. Dichotomous, absence or presence of increased latency of upper sensory Eps at onset (mean = 2.652, sd = 0.996, percentile$_{0.025}$ = 1.242, percentile$_{0.25}$ = 1.947, median = 2.476, percentile$_{0.75}$ = 3.156, percentile$_{0.975}$ = 5.078, PP$_{>1.025}$ = 0.994 , PP$_{<0.975}$ = 0.004)
8. Dichotomous, absence or presence of brainstem or cerebellum lesions at onset (mean = 0.566, sd = 0.177, percentile$_{0.025}$ = 0.294, percentile$_{0.25}$ = 0.440, median = 0.541, percentile$_{0.75}$ = 0.665, percentile$_{0.975}$ = 0.982, PP$_{>1.025}$ = 0.018, PP$_{<0.975}$ = 0.973).

Our model had an in-sample estimated accuracy of 67.38% (95% confidence intervals 61.54% - 72.85%) with a sensitivity of 61.36%, a specificity of 72.79% and an AUC of 0.71.

Conclusions

This prognostic model that was built through a Bayesian approach starting from defined clinical and instrumental variables at disease onset can predict the individual risk for CIS patients of loss NEDA-3 status after a short time. The limit of this study is that the small sample size does not allow to investigate multinomial outcomes and out of sample evaluation of predictivity. We are conducting other studies on larger CIS cohorts to investigate this kind of outcome with more complex models (with bigger sample size and higher dimensionality) such as "black box" classification models, to improve the accuracy of predictions and to test the model with an out – of sample prediction evaluation.

Figure 1. *Marginal posterior distribution of odd ratios. To run this analysis, we used t-Student priors with 7 degrees of freedom. The marginal posterior density of the regression parameter has been omitted from the plot because it is extremely narrow.*



Bibliography

[1] Giovannoni G, Turner B, Gnanapavan S, et al. Is it time to target no evident disease activity (NEDA) in multiple sclerosis? Mult Scler Relat Disord. 2015;4:329–33.

[2] Gelman A., Carlin J. B, Stern Hal S., et al., Bayesian Data Analysis. Third. London, 2013, Chapman &Hall/CRC Press.

[3] R Core Team, 2013, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

[4] Stan Development Team, 2018, RStan: the R interface to Stan. R package version 2.18.2. http://mc-stan.org/.

[5] Carpenter B., Gelman A., Hoffman M.D., et al., 2017, Stan: A probabilistic programming language. *Journal of Statistical Software* 76(1). DOI 10.18637/jss.v076.i01/.

[6] Goodrich B, Gabry J, Ali I & Brilleman S., 2018, rstanarm: Bayesian applied regression modelling via Stan. R package version 2.17.4. http://mc-stan.org/.

[7] Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

[8] H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

A BAYESIAN MENDELIAN RANDOMIZATION APPROACH TO IDENTIFY DISEASE CAUSING PROTEINS VIA PEDIGREE DATA, PARTIALLY OBSERVED EXPOSURES AND CORRELATED INSTRUMENTS.

Teresa Fazia[1], Leonardo Egidi[2], Burcu Ayoglu[3], Ashley Beecham[4], Jacob L. McCauley[4], Peter Nilsson[3], Carlo Berzuini[5], Luisa Bernardinelli[1]

[1] *Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy.*

[2] *Department of Economics, Business, Mathematics and Statistics "Bruno de Finetti" (DEAMS), University of Trieste, Trieste, Italy.*

[3] *SciLifeLab, Dept. of Protein Science, KTH Royal Institute of Technology, Stockholm, Sweden.*

[4] *John P. Hussmann Institute for Human Genomics and Dr John Macdonald Foundation, Department of Human Genetics, University of Miami, Miller School of Medicine, Miami, USA.*

[5] *Centre for Biostatistics, University of Manchester, Manchester, United Kingdom*

Introduction

Genome-wide association studies (GWASs) have found a wealth of genetic associations with diseases and traits. The functional meaning of these associations remains largely unknown. This may be due to: i) the associated variant being in Linkage Disequilibrium (LD) with, but distinct from, the causal one, ii) the partner causal gene of the variant not necessarily being the nearest to it and, iii) lack of longitudinal assessments of gene function [1].

A useful tool in this elucidation effort is the class of methods under the Mendelian Randomization (MR) heading [2-5]. The bare bones of the MR idea are that for an exposure (X) to be a causal influence on an outcome (Y), we, under certain assumptions, expect there to be a genetic variant (Z) that modulates X to likewise affect Y. Information about Z can then be used as an instrument to assess the causal effect of X on Y, despite confounding; in which case Z is called an instrumental variable (IV) for the effect of X on Y.

Objectives

Technical difficulties hamper applicability of standard Mendelian Randomization (MR) methods in situations where sample individuals are not independent. A notable example are samples of individuals organized in families with a known pedigree structure. Interest in extending MR to this kind of data is, in part, motivated by the possibility of making the analysis less vulnerable to biases due to unobserved population stratification, and to the confounding effect of unobserved environmental variables, especially if the studied families belong to an isolated population, as is the case in our illustrative study. Such benefits are extremely valuable in a MR context, since, in this context, they contribute to reducing unobserved exposure-outcome confounding, and, in this way, the source of possible violations of the untestable assumptions at the basis of the method.

Methods

We here show that an analysis of family data is possible by using a Bayesian approach to MR. This we do by specializing the method of Berzuini and colleagues [6] to deal with (i) family data, and (ii) weakly correlation (r^2<0.20) between finely mapped genotype instrumental variables (IVs), and (iii) presence of missing exposures. The proposed method allows pedigree-induced correlation between individuals to enter the specification of a multivariate outcome distribution. By incorporating in the model the family information we (to some degree) reduced the biasing effect of unobserved gene-related confounders.

The proposed method also allows to treat missing exposures as additional parameters to be estimated from the data (in such a way not to incur in a two-sample bias) and, finally it also acknowledges possible correlation between instruments.
In our sample as we jointly analyse two samples of individuals. One of them (Sample 1), contains individuals with a complete set of values for X, Y, and Z. The remaining individuals (Sample 2), have no measured value for X. From a Bayesian point of view, the distinction between samples is artificial, as we may just think in terms of a single sample with some missing exposures, to be treated as additional parameters to be estimated from the data, as is customary in Bayesian analysis.
A final consideration is that Berzuini's approach estimates the direct (pleiotropic) effects of the IVs on outcome (as you would wish to do in a mediation analysis), thereby identifying those instruments that are likely to exhibit pleiotropy, for a subsequent, more focused, clarification of their role.
We illustrate the method with the aid of a study of twenty multiplex Multiple Sclerosis (MS) Sardinian families for a total of N=936 individuals (98 cases and 838 unaffected relatives), whose aim is to identify plasma proteins that cause MS.
Four candidate proteins (EVI5, IL12A, IRF8 and STAT4) were selected a priori on the basis of Genome-Wide Significant (GWS) association between MS and genetic variants located within (e.g. exonic, intronic, in the UTR) or in the proximity of the protein-coding genes (e.g. downstream, intergenic) [7]. For each protein we collected the variants acting as IVs both *in cis* and *in trans* according as that variant is located in the encoding gene, in which case we say it is acting *in cis*, or in the distal region of the gene, in which case we say it is acting *in trans*.

Results

The results obtained in our study, highlighted four biologically characterized proteins. Of foremost interest among them is an immunomodulatory cytokine (i.e, IL12A) that can act as a growth factor for activated T and natural killer (NK) cells, enhance the lytic activity of NK/lymphokine-activated Killer cells, and stimulate the production of IFN-gamma by resting PBMC35 [8]. The highlighted protein is highly involved in autoimmune system activation and its involvement in MS pathogenesis and development is extremely plausible. Extremely interesting is the fact that the IVs which our method highlights as pleiotropic are precisely those which control the protein in trans.

Conclusions

Standard MR methods have been frequently applied to the study of high-level exposures, such as obesity. Our illustrative study turns instead the attention to disease-causing agents that operate at an earlier stage of the causal chain: the proteins. By focusing on proteins, we are professing interest in exposures that are (in some sense) closer to the genetic instruments, and that are more likely to directly point to fundamental mechanistic explanations of disease.
Our proposed method, allow to possibly study family-structured samples which are likely to facilitate detection of causality at a protein level. In fact, sample individuals belonging to the same family tend to display a lower degree of genetic heterogeneity, especially if they are drawn from a founder and isolated population as Sardinian, as in our study. One reason being that, in this case, the individuals also share a large portion of environmental variability. The low heterogeneity facilitates detection of causality at the level of proteins.

References

[1] Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predict complex trait gene targets. Nat Genet. 2016;48(5):481–487.

[2] Katan MB. Apoupoprotein {E} Isoforms, Serum Cholesterol, and Cancer. Lancet. 1986;327(8479):507–508.

[3] Davey Smith G, Ebrahim S. Mendelian Randomization: Can Genetic Epidemiology Contribute to Understanding Environmental Determinants of Disease? Int J Epidemiol. 2003;32:1–22.

[4] Lawlor DA, Harbord RM, Sterne JAC, et al. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133–1163.

[5] Didelez V, Sheehan NA. Mendelian Randomisation as an Instrumental Variable Approach to Causal Inference. Stat Methods Med Res. 2007;16:309–330.

[6] Berzuini C, Guo H, Burgess S, Bernardinelli L. A Bayesian approach to Mendelian randomization with multiple pleiotropic variants. Biostatistics. 2018 Aug 1. [Epub ahead of print]

[7] Consortium IMSG, Beecham AH, Patsopoulos NA, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013;45(11):1353–1360.

[8] Trinchieri G. Interleukin-12 and the regulation of innate resistance and adaptive immunity. Nat Rev Immunol. 2003;3(2):133–146.

PATTERN DIETETICI E MALATTIE RESPIRATORIE IN GEIRD

Chamitava Liliya[1], Zanolin Maria Elisabetta[1], Garcia-Larsen Vanessa[2], Ferrari Marcello[3], Cerveri Isa[4], Pirina Pietro[5], Verlato Giuseppe[1], Cazzoletti Lucia[1]

[1] *Sezione di Epidemiologia e Statistica medica, Dipartimento di Diagnostica e Sanità pubblica, Università di Verona, Verona, Italia*

[2] *Dipartimento di salute internazionale, Scuola di Sanità pubblica Johns Hopkins Bloomberg, Baltimora, Maryland*

[3] *Sezione di Malattie Respiratorie, Dipartimento di Medicina Interna, Università di Verona, Verona, Italia*

[4] *Divisione di Malattie Respiratorie, Istituto di Ricovero e Cura a Carattere Scientifico Fondazione Ospedale "San Matteo", Università di Pavia, Pavia, Italia*

[5] *Istituto di Malattie Respiratorie, Università di Sassari, Sassari, Italia*

Razionale

I recenti sviluppi dell'epidemiologia nutrizionale sono orientati allo studio di pattern dietetici piuttosto che di singoli cibi o nutrienti. L'analisi dei pattern dietetici considera gli alimenti in modo sinergico e mira a riconoscere potenziali interazioni tra sostanze nutritive e alimenti, e sembra in grado di affrontare meglio il problema del confondimento rispetto alle analisi dei singoli alimenti o nutrienti. Uno tra gli approcci più utilizzati nella derivazione di pattern dietetici è l'analisi delle componenti principali (PCA) (1, 2). I questionari sulla frequenza alimentare (FFQ) sono tra gli strumenti più comunemente usati per raccogliere informazioni dietetiche a livello individuale e, in particolare, nell'analisi dei pattern dietetici negli studi epidemiologici (3, 4).
Si suppone che abitudini alimentari caratterizzate da un elevato consumo di frutta/verdura potrebbero essere inversamente associate alla prevalenza di malattie respiratorie come l'asma e, d'altra parte, avere un'associazione diretta con una buona funzionalità respiratoria negli adulti. Al contrario, pattern dietetici con un elevato consumo di alimenti trasformati potrebbero essere direttamente associati all'insorgenza di malattie respiratorie e ad una funzionalità polmonare peggiore.

Obiettivo

Valutare le associazioni tra pattern dietetici e outcome respiratori (FVC, Capacità Vitale Forzata; FEV_1, Volume Espiratorio Forzato in 1 secondo) in adulti italiani partecipanti allo studio GEIRD (Gene Environment Interactions in Respiratory Diseases).

Metodi

Disegno dello studio. Il progetto GEIRD è uno studio multicentrico, di popolazione, che coinvolge sette centri italiani (Ancona, Palermo, Pavia, Sassari, Terni, Torino e Verona). Nella prima fase dello studio, è stato spedito un questionario riguardante sintomi respiratori a nuovi campioni casuali della popolazione generale o a preesistenti coorti di soggetti estratti casualmente dalla popolazione generale (20-84 anni di età, maschio/ femmina = 1/1). Nella seconda fase, tutti i soggetti che hanno riferito sintomi di bronchite cronica (CB), broncopneumopatia cronica ostruttiva (BPCO) o asma e campioni casuali di soggetti che hanno riportato sintomi di rinite, e di soggetti senza sintomi respiratori, sono stati invitati a sottoporsi ad alcuni esami clinici e di funzionalità polmonare (spirometria, test di broncoprovocazione alla metacolina/test di broncodilatazione, skin prick test).
Per questa analisi abbiamo preso in considerazione i soggetti reclutati nei centri di Pavia, Torino, Verona, Ancona e Sassari, escludendo gli altri centri a causa della mancanza o incompletezza delle informazioni sulla dieta.

Il numero di soggetti idonei per la fase di screening in questi centri era pari a 21558, di questi 12739 (59.1%) hanno compilato il questionario postale. Complessivamente, su 6586 soggetti invitati alle cliniche, 3046 (46.2%) hanno partecipato alla fase clinica. Dei 3046 soggetti che hanno partecipato alla fase clinica, 1480 (48.5%) hanno compilato il FFQ EPIC (European Prospective Investigation in to Cancer and Nutrition, Indagine prospettica europea su cancro e nutrizione), contenente 260 alimenti e strutturato per pasti (porzioni standard). Il FFQ utilizzato è la versione italiana del questionario validato EPIC (5).
Dei 1480 adulti di Pavia, Torino, Verona, Ancona e Sassari che hanno compilato il FFQ, 1375 soggetti (93%) hanno fornito informazioni complete relativamente a dati dietetici e clinici. Solo i soggetti che hanno risposto ad almeno l'80% delle domande del FFQ sono stati inclusi nel database (1334 osservazioni, 97%); inoltre, solo i soggetti compresi nell' intervallo 0.5 – 99.5% del rapporto intake calorico totale (Total Energy Intake, TEI) / metabolismo basale (Basal Metabolic Rate, BMR) sono stati considerati nell'analisi (1329 osservazioni, 99.6%), dove il BMR è stato stimato utilizzando equazioni specifiche per sesso per adulti (≤ 60 anni) e per soggetti anziani (> 60 anni) (6, 7).
1177 adulti di 3 centri (Pavia, Verona, Sassari) sono stati inclusi nel database finale. Gli altri 2 centri (Torino e Ancona) sono stati esclusi a causa della bassa numerosità campionaria. Nel presente studio, si sono derivati i pattern dietetici nel gruppo dei controlli (420 osservazioni). La relazione tra pattern dietetici (varabili indipendenti) e funzionalità respiratoria (variabili dipendenti), è stata indagata separatamente nei soggetti con asma (305 osservazioni), e nei controlli.
Si è utilizzato il software NAF (Analisi nutrizionale dei questionari di frequenza alimentare, National Cancer Institute, Milano, Italia) per ricavare l'assunzione giornaliera di cibo (grammi) e stimare il TEI (8).
Dai 233 cibi del FFQ sono stati eliminati 6 cibi che hanno presentato più del 20% di valori mancanti (227 cibi rimanenti, 97.4%); successivamente, sono stati eliminati 52 cibi (175 cibi rimanenti, 77.1%), non consumati da più dell'80% dei soggetti intervistati.

Analisi statistiche. Si sono utilizzati il test esatto di Fisher e il test di Kruskal-Wallis per analizzare le differenze di FVC, FEV1 tra centri, separatamente negli asmatici e nei controlli; con gli stessi test, si sono studiate anche le differenze tra centri dei potenziali fattori confondenti (età, peso, altezza o BMI (indice di massa corporea), genere, abitudine al fumo (mai, ex-, fumatore corrente), TEI, condizione lavorativa (occupato, disoccupato) e livello di educazione (età alla fine degli studi prima o dopo i 16 anni)), con livello di significatività statistica $p < 0.05$.

Pattern dietetici. È stato utilizzato il metodo di Rosenthal (9) per preparare i dati per la meta-analisi. Si è applicata la PCA, con rotazione varimax, alla matrice dei coefficienti di correlazione aggregata, per ottenere le componenti (pattern dietetici), dalle quali, a loro volta, si sono ricavati i punteggi dei pattern dietetici per ogni centro. I pattern dietetici sono stati costruiti con gli alimenti che presentavano un peso fattoriale > 0.3 in valore assoluto. Il numero dei componenti (due pattern dietetici) è stato deciso sulla base dello screeplot. In ogni centro, è stata fatta una regressione quantilica con bootstrapping e 500 replicazioni per valutare l'associazione tra ciascun pattern dietetico e funzionalità polmonare (FEV_1 e FVC, variabili dipendenti) in asmatici e controlli separatamente, aggiustando per l'altro pattern dietetico e i possibili confondenti. I risultati delle regressioni sono stati aggregati utilizzando una procedura di meta-analisi. L'eterogeneità tra i centri è stata valutata tramite il test statistico I-quadrato.

Risultati

Le caratteristiche del campione dei soggetti con asma e dei soggetti senz'asma (controlli) nei 3 centri sono presentati in tabella 1.
La PCA ha individuato due componenti con autovalore > 10: una componente "dieta prudente", associata al consumo di verdure, olio d'oliva e frutta, e l'altra componente "dieta occidentale", caratterizzata da un maggiore consumo di pasta, carne, formaggi, antipasti e dolci.

La meta-analisi delle associazioni tra pattern dietetici e funzionalità polmonare (FEV_1 e FVC) negli asmatici nei 3 centri italiani (Pavia, Verona e Sassari) ha mostrato un livello di FEV_1 (ml) significativamente ridotto nei soggetti che consumano una "dieta occidentale" β [95% CI] = -417[-799; -34] ml, dove β è il coefficiente di regressione aggiustato.
Non si sono rilevate altre evidenze di associazioni tra pattern dietetici e funzionalità polmonare né negli asmatici né nei controlli: per il gruppo asmatici / "dieta prudente": FEV_1 β[95% CI] = -193[-593; 207] ml; per il gruppo asmatici /"dieta prudente": FVC β[95% CI] = -37[-554; 479] ml; per il gruppo asmatici / "dieta occidentale": FVC β[95% CI] = - 98[-617; 420] ml; per il gruppo di controllo / "dieta prudente": FEV_1 β[95% CI] = 33[-130; 196] ml; per il gruppo di controllo / "dieta occidentale": FEV_1 β[95% CI] = 194[-75; 464] ml; per il gruppo di controllo / "dieta prudente": FVC β[95% CI] = 69[-148; 285] ml; per il gruppo di controllo/ "dieta occidentale": FVC β[95% CI] = 167[-181; 515] ml.

Conclusioni

In questo studio caso-controllo multicentrico di popolazione, abbiamo utilizzato un approccio meta-analitico alla PCA che permette di affrontare il problema della diversità della dieta nei centri coinvolti nello studio.
I risultati delle analisi condotte mostrano che negli asmatici adulti il valore di FEV_1 si riduce all'aumentare del consumo di cibi tipici della "dieta occidentale", quali pasta, carne, formaggi, antipasti e dolci.

Tabella 1. Caratteristiche dei campioni dei soggetti asmatici e dei controlli (senza asma)

SOGGETTI ASMATICI				
	Centri			
Variabili	*Pavia (n = 28)*	*Verona (n = 187)*	*Sassari (n = 28)*	*p*
	Mediana (IQR)			
FEV$_1$, *ml*	2880 (2030-3080)	3190 (2600-3770)	2280 (1550-2720)	<0.001
FVC, *ml*	3620 (2920-4290)	4005 (3470-5000)	3005 (2450-3460)	<0.001
Età, *anni*	56.5 (46.7-63.1)	46.2 (39.0-53.2)	70.5 (54.2-75.9)	<0.001
TEI, ***kcal***	1735.6 (1185.5-2204.4)	1960.3 (1599.2-2494.8)	1590.3 (1393.6-2010.7)	<0.01
BMI, *kg/m^2*	26.0 (22.5-29.1)	24.6 (22.1-27.6)	24.7 (23.2-27.6)	0.395
Altezza, *cm*	169.5 (165.0-173.5)	169.0 (163.0-176.0)	162.5 (160.0-172.0)	<0.05
Peso, *kg*	75.0 (66.5-84.0)	71.0 (60.0-81.0)	66.5 (58.0-75.0)	0.156
	N(%)			
Sesso, *maschi*	13 (46.4)	91 (48.7)	18 (64.3)	0.286
Livello educativo, *≤16 anni*	7 (25.0)	42 (22.5)	9 (32.1)	0.531
Occupazione, *si*	26 (92.9)	179 (96.2)	28 (100)	0.333
Fumo				0.517
mai	11 (40.7)	94 (50.3)	12 (42.9)	
passato	10 (37.1)	53 (28.3)	12 (42.9)	
corrente	6 (22.2)	40 (21.4)	4 (14.2)	
Attività fisica				0.071
leggera	20 (71.4)	108 (57.8)	22 (78.6)	
moderata	6 (21.4)	69 (36.9)	4 (14.3)	
vigorosa	2 (7.2)	10 (5.3)	2 (7.1)	
CONTROLLI				
	Centri			
Variabili	*Pavia (n=93)*	*Verona (n=288)*	*Sassari (n=39)*	*p*
	Mediana(IQR)			
FEV$_1$, *ml*	2880 (2480-3460)	3275 (2795-3995)	2720 (2520-3210)	<0.001
FVC, *ml*	3720 (3120-4290)	4000 (3325-4865)	3550 (2910-4330)	<0.01
Età, *anni*	59.0 (53.5-63.7)	48.7 (41.5-56.7)	71.3 (56.1-74.1)	<0.001
TEI, ***kcal***	1903.2 (1516.8-2450.5)	1888.9 (1484.8-2471.7)	1634.1 (1128.3-2095.7)	<0.05
BMI, *kg/m^2*	25.0 (22.3-27.2)	24.8 (22.3-27.6)	25.1 (23.0-27.9)	0.687
Altezza, *cm*	167.0 (162.0-172.0)	168.0 (162.0-177.0)	166.0 (162.0-172.0)	0.191
Peso, *kg*	70.0 (62.0-78.0)	70.5 (61.0-83.0)	71.0 (61.0-80.0)	0.883
	N(%)			
Sesso, *maschi*	38 (40.9)	135 (46.9)	27 (69.2)	<0.05
Livello educativo, *≤16 anni*	22 (23.7)	65 (22.6)	12 (30.8)	0.513
Occupazione, *si*	89 (95.7)	276 (95.8)	38 (97.4)	1.00
Fumo				0.170
mai	52 (55.9)	161 (56.1)	20 (51.3)	
passato	30 (32.3)	87 (30.3)	18 (46.1)	
corrente	11 (11.8)	39 (13.6)	1 (2.6)	
Attività fisica				<0.05
leggera	59 (63.5)	153 (53.1)	30 (76.9)	
moderata	31 (33.3)	115 (40.0)	7 (18.0)	
vigorosa	3 (3.2)	20 (6.9)	2 (5.1)	

Bibliografia

[1] Gorsuch RL. Factor analysis. 2015.

[2] Bakolis I, Burney P, Hooper R. Principal components analysis of diet and alternatives for identifying the combination of foods that are associated with the risk of disease: a simulation study. Br J Nutr. 2014;112(1):61-9.

[3] Shim JS, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health. 2014; 36: e2014009.

[4] Willett W. Foreword. The validity of dietary assessment methods for use in epidemiologic studies. Br J Nutr. 2009;102 Suppl 1: S1-2.

[5] Pisani P, Faggiano F, Krogh V, Palli D, Vineis P, Berrino F. Relative validity and reproducibility of a food frequency dietary questionnaire for use in the Italian EPIC centres. Int J Epidemiol. 1997;26 Suppl 1:S152-60.

[6] Harris JA, Benedict FG. A Biometric Study of Human Basal Metabolism. Proc Natl Acad Sci U S A. 1918;4(12):370-3.

[7] Fredrix EW, Soeters PB, Deerenberg IM, Kester AD, von Meyenfeldt MF, Saris WH. Resting and sleeping energy expenditure in the elderly. Eur J Clin Nutr. 1990;44(10):741-7.

[8] Pala V, Sieri S, Palli D, Salvini S, Berrino F, Bellegotti M, et al. Diet in the Italian EPIC cohorts: presentation of data and methodological issues. Tumori. 2003;89(6):594-607.

[9] Rosenthal R. Meta-analytic procedures for social research. Newbury Park: Sage Publications; 1991.

POTENTIAL IMPACT FRACTION FOR A CONTINUOUS RISK FACTOR: ASSESSING THE BURDEN OF ORAL AND PHARYNGEAL CANCER ACCORDING TO ADHERENCE TO THE HEALTHY EATING INDEX

Di Maso Matteo[1,2], Tomaino Laura[1], Garavello Werner[3], Crispo Anna[4], Serraino Diego[5], Ferraroni Monica[1], La Vecchia Carlo[1], Edefonti Valeria[1] and Bravi Francesca[1]

1 Department of Clinical Sciences and Community Health, Branch of Medical Statistics, Biometry and Epidemiology "G.A. Maccacaro", Università degli Studi di Milano, Milan, Italy

2 Department of Public Health and Pediatric Sciences, Università degli Studi di Torino, CTO Hospital, Turin. Italy

3 Department of Otorhinolaryngology, School of Medicine and Surgery, Università degli Studi di Milano-Bicocca, Monza, Italy

4 Unit of Epidemiology, Istituto Tumori Fondazione Pascale, IRCSS, Naples, Italy

5 Unit of Cancer Epidemiology, Centro di Riferimento Oncologico di Aviano (CRO) IRCCS, Aviano (PN), Italy

Introduction

The attributable fraction (AF) is the candidate tool to quantify individual shares of risk factors on the disease burden in a population. It is widely used in epidemiology to express the proportion of cases attributable to one or more risk factors and in public health to plan and prioritize prevention strategies [1].
Levin proposed the classical definition of AF for a single binary risk factor [2], as a measure that accounts for both the association between risk factor and disease occurrence and risk factor prevalence:

$$AF_{X_{binary}} = \frac{(RR-1)\cdot P(X)}{1+(RR-1)\cdot P(X)}$$

where $P(X)$ is the prevalence (or distribution) of the binary risk factor X in the population and RR is the relative risk. The AF can be also formulated in terms of probabilities according to the MacMahon and Pugh's formula [3]:

$$AF_{X_{binary}} = \frac{P(D)-P(D\mid \bar{X})}{P(D)}$$

where the dominator is the disease probability in the overall population, whereas the numerator is the difference between this probability and the one that would be observed if all people in the population were unexposed. Previous formulas were extended to take into account multiple risk factors and/or confounders (adjusted, sequential and average AFs), different study design (cohort with count and time-to-event outcomes, cross-sectional, and case-control), and types of risk factors (categorical, and continuous) [4].
Implicitly, the AF methodology assumes the *maximum prevalence reduction* scenario in which all subjects in the population were shifted to the unexposed (or lowest exposure) level. Obviously, if the factor considered has a protective effect the shifting will be toward the highest exposure level. In practice, the *maximum prevalence reduction* scenario is not easily feasible and alternative public health scenarios to reduce the disease burden may be considered. For instance, halve the risk factor prevalence for all subjects or changing the risk factor prevalence for specific subgroups at higher risk could be more reasonable alternatives.

A generalization of the AF concept is the potential impact fraction (PIF) defined as the proportion of disease burden that would occur if the distribution of the risk factor in population was modified to an alternative (or counterfactual) scenario [5, 6].
For a continuous risk factor X, the PIF is:

$$PIF_{X_{continuous}} = \frac{\int_{x=0}^{m} RR(x) \cdot P(x) dx - \int_{x=0}^{m} RR(x) \cdot P'(x) dx}{\int_{x=0}^{m} RR(X) \cdot P(X) dx}$$

where $RR(x)$ is the relative risk for the continuous risk factor X at exposure level x, $P(x)$ is the population distribution (or prevalence) of X at exposure level x, $P'(x)$ is the counterfactual distribution of X at level x and m is the maximum exposure level. The denominator represents the exposure-weighted risk of disease in the population under the observed risk factor prevalence and the numerator represents the difference between this risk of disease and one obtained under the counterfactual scenario. Clearly, the *maximum prevalence reduction* scenario underlying the AF methodology is one of the possible alternatives.

Aim

Using data from an Italian multicentre hospital-based case-control study, we will estimate the potential impact fraction (PIF) for oral and pharyngeal cancer attributable to healthy eating index (HEI) adherence according to 6 counterfactual HEI prevalence distributions. Furthermore, we will compare PIF estimates, will give their interpretations, and will discuss shortcomings.

Methods

Data collection. Data came from a multicentre hospital-based case-control study conducted between 1992 and 2009 in different Italian areas. Study areas included the greater Milan area and the provinces of Pordenone and Gorizia (Northern Italy), the province of Latina (Central Italy), and the urban area of Naples (Southern Italy). Cases were 946 patients (756 men and 190 women; median age 58 years, range 22–79 years) admitted to major hospitals in the study areas, with incident histologically confirmed cancer of the oral cavity and pharynx. The control group comprised 2492 cancer-free patients (1497 men and 995 women; median age 58 years, range 19–82 years) frequency-matched to cases by study area, study period, sex, and age. Controls were admitted to the same network of hospitals as the cases for a wide spectrum of acute, non-neoplastic conditions, unrelated to known or likely risk factors for oral and pharyngeal cancer (e.g. tobacco and alcohol consumption, human papilloma virus infection, etc.) or to a long-term diet modification. Overall, 19% of controls were admitted for traumas, 21% for non-traumatic orthopaedic disorders, 51% for acute surgical conditions, and 9% for other miscellaneous illnesses. All study participants signed an informed consent in accordance with the recommendations of the Board of Ethics of the study hospitals.
Exposure assessment. Trained interviewers administered a structured questionnaire to study participants during their hospital stay. The questionnaire included information on socio-demographic characteristics, anthropometric measures, lifestyles (e.g. lifetime smoking habit, physical activity, etc.), use of drugs, personal history of disease, family history of cancer, and dietary habits. A reproducible and validated food frequency questionnaire (FFQ) [7-10] was used to assess patients' usual diet with reference to the two years prior the interview. The FFQ included information on weekly intake of 78-food (or recipes) items, according to the following sections: (i) milk, hot beverages and sweeteners; (ii) bread, cereals and first courses

(including soups); (iii) second courses (e.g., meat, fish, and other main dishes); (iv) side dishes (i.e., vegetables and potatoes); (v) fruit; (vi) sweets, dessert and soft drinks. An additional section collected the lifetime consumption of alcoholic beverages. Serving size was defined either in "natural" units (e.g., 1 cup of milk, 1 coffee spoon of sugar, 1 egg, etc.) or as an Italian average serving (e.g., 80 g of pasta, 100 g of salad, 125 g of potatoes. 120 g of beef, etc.). Each serving size was converted into the corresponding approximately weight (grams) or capacity (millilitre). Seasonal variation in fruit and vegetable consumption was also considered to account for the fluctuations within the year. Occasional consumption (i.e., an intake lower than once a week, but higher than once per month) was coded as 0.5 serving per week. The FFQ-derived dietary data were used to calculate the HEI-2015 score for each study subject.

Health Eating Index. The HEI is an energy density-based measure for assessing diet quality, specifically the degree to which a set of foods is in compliance with the 2015-2020 Dietary Guidelines for Americans (DGA) [11, 12]. In particular, the HEI is composed of 9 adequacy components (i.e., total and whole fruits, total vegetables, greens and beans, whole grains, milk/dairy products, total protein foods, seafood and plants, proteins, and fatty acids) and 4 moderation components (i.e., refined grains, sodium, added sugars, and saturated fats). The components are generally scored between 0 (non-compliance) and 10 (highest compliance) except for those further classified in two subcategories (e.g., total and whole fruits) that are scored between 0 and 5. Each of the components is scored on standard densities (e.g. component amount per 1000 kcal/day) according to DGA cut-offs. For instance, if the total fruit consumption for an individual (expressed as total fruit amount per 1000 kcal/day) is exactly 2/3 of the maximum total fruit standard density, it will be scored as 3.33. Thus, the HEI is a sum of the 13 components resulting in a continuous measure defined in the interval [0,100] with higher HEI values indicating greater compliance to the DGA recommendations.

Risk estimation and counterfactual scenarios. The PIF method as described in the previous section can correctly be estimated in cross-sectional and prospective cohort studies. Instead, case-control studies select participants among diseased and not diseased subjects rather than exposed and unexposed ones. Since the ratio of controls to cases in the sample is fixed a priori, the risk factor distribution will be biased as will the PIF. Ferguson et al. [13] proposed an approach to account for the imbalance between case and controls by incorporating disease prevalence in the log-likelihood function of the regression model used to estimate disease risk. Briefly, the method assigns a weight of 1 to each case and a weight of $(1-p)/(v \cdot p)$ to each control, where p is the prevalence of disease in the population and v is the number of controls. We set a 81×10^{-6} prevalence of oral and pharyngeal cancer according to figures of the Italian Consortium of Cancer Registries [14].

We estimated odds ratio (OR) and 95% confidence interval (CI) for oral and pharyngeal cancer risk and HEI score adjusting for age, sex, study centre, year of study enrolment, years of education, body mass index (BMI), tobacco smoking, alcohol drinking, and total non-alcohol energy intake.

We designed the following 6 scenarios:

1. maximum prevalence reduction (i.e., assigning the highest HEI value for all study subjects);

2. mild global prevalence intervention (i.e., 10-point increment in the HEI for all study subjects);

3. strong prevalence intervention on low and lower-middle HEI subjects (i.e., 1/3 prevalence increment for subjects in the 1st and 2nd quartile of the HEI distribution);

4. *mild prevalence intervention on low and lower-middle HEI subjects* (i.e., 1/3 prevalence increment for subjects in the 1st quartile and 1/4 prevalence increment for subjects in the 2nd quartile of the HEI distribution);

5. *strong prevalence intervention on lower- and upper-middle HEI subjects* (i.e., 1/3 prevalence increment for subjects in the 2nd quartile and 1/4 increment for subjects in the 3rd quartile in the HEI distribution);

6. *mild prevalence intervention on lower- and upper-middle HEI subjects* (i.e.; 1/4 prevalence increment for subjects in the 2nd and 3rd quartile in the HEI distribution).

We conducted all analyses with the pifpaf [15] package in R (R Foundation for Statistical Computing, Vienna Austria) [16].

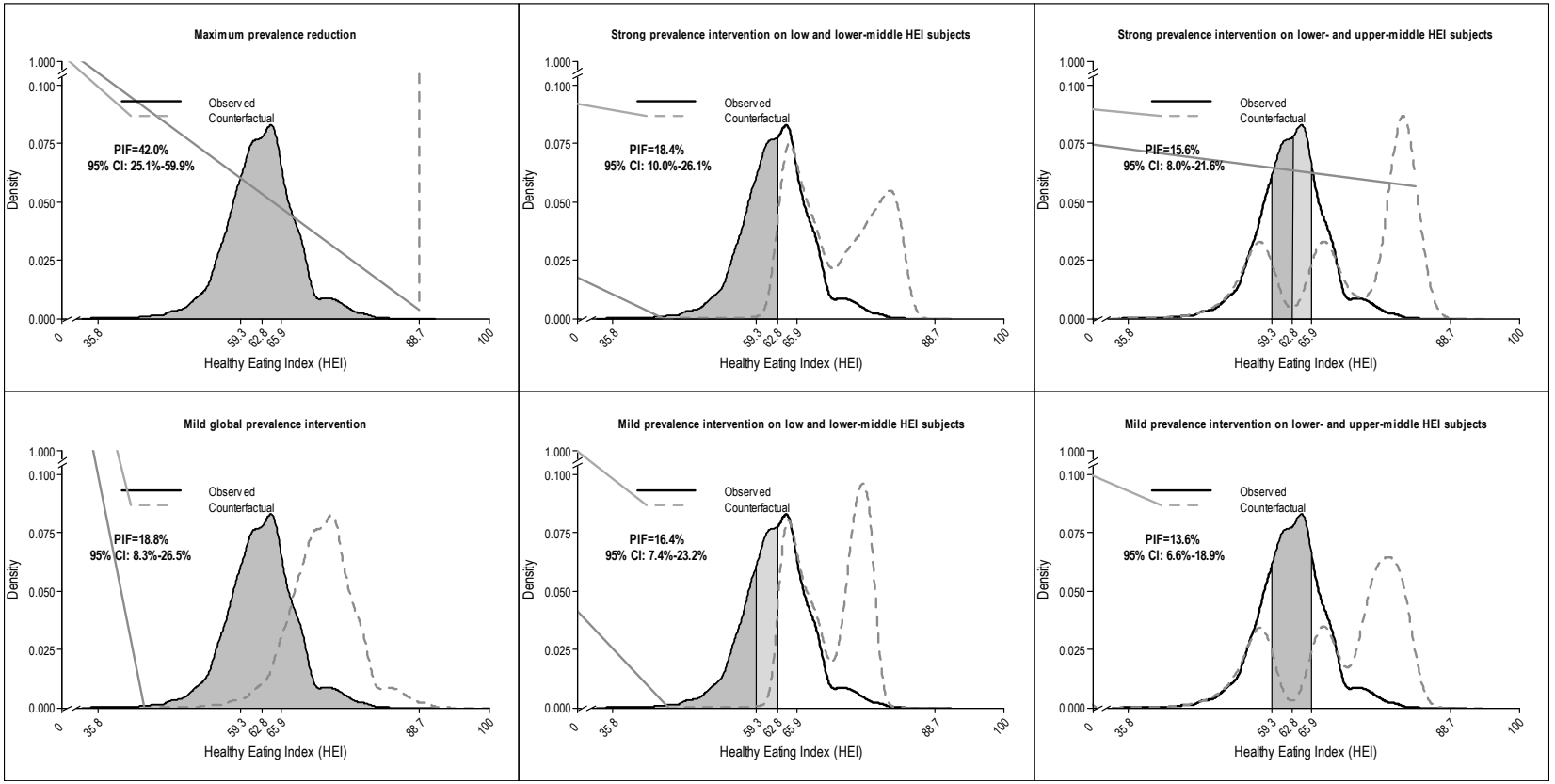
Results

Overall, the HEI ranged from 35.8 to 88.7 with a median of 62.8 (interquartile range: IQR=6.6; figure 1). The OR of oral and pharyngeal cancer for 1-point increment in the HEI was 0.98 (95% CI: 0.97-0.99). The fraction of oral and pharyngeal cancer cases attributable to HEI under the *maximum prevalence reduction* scenario was 42.0%, 95% CI: 25.1%-59.9% (figure 1 top left panel). This figure was of 18.8% (95% CI: 8.3%-26.5%) according to the *mild global prevalence intervention* scenario (figure 1 bottom right panel). Whereas, the shares of cases attributable to changing the prevalence only for subjects belonging to the 1st and 2nd quartile in the HEI distribution were 18.4% (95% CI: 10.0%-26.1%) and 16.4% (95% CI: 7.4%-23.2%) according to the *strong prevalence intervention on the low and lower-middle HEI subjects* scenario (figure 1 top central panel) and the *mild prevalence intervention on the low and lower-middle HEI subjects* scenario (figure 1 bottom central panel), respectively.
The fraction of attributable cases were 15.6% (95% CI: 8.0%-21.6%) under the *strong prevalence intervention on the lower- and upper-middle HEI subjects* scenario and 13.6% (6.6%-18.9%) under *mild prevalence intervention on the lower- and upper-middle HEI subjects* scenario (figure 1 top and bottom left panel).

Conclusions

A suitable epidemiological tool to address preventive strategies have to take into account the magnitude of risk factors and their prevalence in the population. The PIF fairly accomplish this task because it considers the risk of disease and the change in the risk factor prevalence according to a counterfactual scenario. The well-known AF measure is a special case of the PIF that expresses the number of cases when the risk factor (or risk factors) prevalence would be reduced to unexposed (or lowest exposure) level. Furthermore, the PIF offers the possibility of examining continuous risk factors. Indeed, risk factors exposure is often a continuous phenomenon, but in practice researchers and epidemiologists have mostly been using categories. However, when evaluating effects of a continuous risk factor and corresponding public health interventions, the interest is mostly far from categories.
In our example, we used an Italian multicentre case-control study to assess the proportion of oral cavity and pharynx cancer cases attributable to HEI adherence according to different counterfactual scenarios. Clearly, the theoretical maximum disease reduction was observed under the *maximum reduction prevalence* scenario (42%), but it is hardly achievable. The *mild global prevalence intervention* scenario could be a more realistic option with a fraction of preventable cases of approximately 19%. However, this scenario assumed a 10-point increment in the HEI for all study subjects, including also those subjects with a high HEI adherence. Instead, the remaining 4 scenarios involved only half of the study subjects (i.e., people with low and medium adherence), showing similar results (from 13.6% to 18.4%) compared to the *mild global prevalence intervention* scenario. Conversely, these scenarios required a more pronounced efforts by target subjects with an HEI increment up to 20.9 points.
In the original notation, the PIF was designed for prospective studies. Ferguson et al. [13] proposed an interesting approach to estimate the AF in case-control studies. This method consists in weighting the likelihood function of the model for the disease prevalence. Although this method can be easily applied, the assumption that there are no factor influencing the selection of controls other than the prevalence of disease used to weight the regression model is crucial [17].
In conclusion, the need of the epidemiologist and public health officers to translates risk factor prevalence and disease occurrence in useful numbers under more realistic and feasible counterfactual scenarios makes the PIF to be in a mounting epidemiological importance.

Figure 1. *Potential impact fraction (PIF) of oral cavity and pharyngeal cancer cases attributable to healthy eating index (HEI) adherence according to 6 scenarios.*



Shadow areas represent target subjects involved in the prevalence change intervention. Colour gradient represent the magnitude of the prevalence intervention.

References

[1] Bray F, Soerjomataram I. Population attributable fractions continue to unmask the power of prevention. Br J Cancer 2018; 118(8):1031-32.

[2] Levin ML. The occurence of lung cancer in man. Acta Unio Internationalis contra Cancrum 1953; 9(3):531-41.

[3] MacMahon B, Pugh TF. Epidemiology Principles and Methods. Boston, MA, Little, Brown and Company 1970.

[4] Di Maso M, Bravi F, Polesel J, et al. Attributable fraction for multiple risk factors: Methods, interpretations, and examples. Stat Methods Med Res 2019:962280219848471.

[5] Eide GE, Heuch I. Attributable fractions: fundamental concepts and their visualization. Stat Methods in Med Res 2001; 10(3):159-193.

[6] Ezzati M, Hoorn SV, Rodgers A, et al. Estimates of global and regional potential health gains from reducing multiple major risk factors. Lancet 2003; 362(9380):271-80.

[7] Decarli A, Franceschi S, Ferraroni M, et al. Validation of a food-frequency questionnaire to assess dietary intakes in cancer studies in Italy. Results for specific nutrients. Ann Epidemiol 1996; 6(2):110-8.

[8] Ferraroni M, Decarli A, Franceschi S, et al. Validity and reproducibility of alcohol consumption in Italy. Int J Epidemiol 1996; 25(4):775-82.

[9] Franceschi S, Negri E, Salvini S, et al. Reproducibility of an Italian food frequency questionnaire for cancer studies: results for specific food items. Eur J Cancer 1993; 29A(16):2298-305.

[10] Franceschi S, Barbone F, Negri E, et al. Reproducibility of an Italian food frequency questionnaire for cancer studies. Results for specific nutrients. Ann Epidemiol 1995; 5(1):69-75.

[11] Krebs-Smith SM, Pannucci TE, Subar AF, et al. Update of the Healthy Eating Index: HEI-2015. J Acad Nutr Diet 2018; 118(9):1591-1602.

[12] Reedy J, Lerman JL, Krebs-Smith SM, et al. Evaluation of the Healthy Eating Index-2015. J Acad Nutr Diet 2018;118(9):1622-1633.

[13] Ferguson J, Alvarez-Iglesias A, Newell J, et al. Estimating average attributable fractions with confidence intervals for cohort and case-control studies. Stat Methods Med Res 2016; 27(4):1141-52.

[14] Associazione Italiana di Oncologia Medica (AIOM), Associazione Italiana dei Registri Tumori (AIRTUM). I numeri del cancro in Italia 2014. Intermedia Editore 2014; AIOM/AIRTUM: Brescia.

[15] Zepeda-Tello R, Camacho-García-Formentí D. pifpaf: Potential Impact Fraction and Population Attributable Fraction for Cross-Sectional Data. R Foundation for Statistical Computing 2017, Vienna, Austria. https://CRAN.R-project.org/package=pifpaf. Accessed March 19, 2019.

[16] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing 2018, Vienna, Austria. URL https://www.R-project.org. Accessed March 19, 2019. 2018.

[17] Haaf KT, Steyerberg EW. Methods for individualized assessment of absolute risk in case-control studies should be weighted carefully. Eur J Epidemiol 2016; 31(11):1067-1068.

A MACHINE LEARNING APPROACH FOR INSPECTING RELATIONSHIPS BETWEEN PARKINSON DISEASE AND METAL EXPOSURE

Vezzoli Marika[1], Renzetti Stefano[1], Parrinello Giovanni[1], Calza Stefano[1]

1 Department of Translational Medicine, University of Brescia, Italy

Keywords: PD, Ensemble Methods, Random Forest, Variable Importance, Accuracy

Introduction

Parkinson's disease (PD) is a severe neurodegenerative disorder featured by motor dysfunctions such as bradykinesia, tremor, rigidity, as well as other symptoms including chronic fatigue, depression and sleep disturbances impacting on the quality of patients' life. Environmental influence, especially heavy metals, is a solid etiology of PD, widely investigated in many studies.
From a methodological point of view, many researches involving PD deal with fat matrixes (more covariates than subjects), containing variables of different nature (qualitative and quantitative), some of them with a high number of missing values or strongly correlated (multicollinearity).
For analysing this dataset and addressing with the abovementioned problems, we need robust methods able also to model non-linear relationships between the outcome (typically the diagnosis) and a high number of covariates.

Aim

The aim of this research is to model the relationships between the exposure to heavy metals and PD using robust methods belonging to machine learning approach. In detail, we apply the Random Forests [1] algorithm which reduce the instability of single trees [2], providing accurate results in terms of sensitivity and specificity and, more important, identifying, among a huge number of covariates, which are the predictors that have a major impact on the diagnosis.

Methods

Valcamonica is an Italian valley where ferroalloy industries have been active for a century and with an increased prevalence of parkinsonism. In this research, a case-control study containing 169 observations (93 cases, 55.03% and 76 controls, 44.97%) compares patients with PD and control subjects coming from Valcamonica and other areas around Brescia town (Italy). The protocol includes information on age, sex, occupational, residential history, life habits, neuro-psychological testing, and the assessment of genetic polymorphism. Subjects are screened also for serum Cu, Zn, Fe, Mn in blood (MnB) and urine (MnU), transferrin, peroxides, alanine (ALT) and aspartate (AST) transaminases and direct bilirubin. All the PD patients underwent Unified Parkinson's Disease Rating Scale (UDPRS) and metal deterioration battery, namely a neuropsychological assessment using for epidemiological and clinical purposes in Italy. Some genetic analysis was also performed.
Since the dataset is affected by the following problems, (i) high number of missing values, (ii) variables are strongly correlated, and (iii) n<<p (number of observations is less than number of variables), we apply the Random Forests (RF) algorithm for modelling the diagnosis: such an algorithm is indeed an ensemble method extremely flexible and well suited to handle these complications.
RF are an extension of Classification/Regression trees (Breiman et al., 1984) which combine the predictions obtained from many trees grown on perturbed versions of the dataset. RF give also a variable importance measure, the Mean Decrease in Accuracy (MDA), which identifies the covariates which exert the higher impact on the prediction of the outcome (diagnosis). Analytically, within each tree of the RF all the values of the r-th variable are randomly permuted and new predictions are obtained on this new data set (Y, Xr); next,

a loss function Lr is computed and compared with the original loss function of the ensemble L. This procedure is repeated m times on different bootstrap samples and the MDA measure for the r-th variable is given by the following average (av) on m:

$$MDA_r = av_m(L_r - L).$$

The measure obtained is relativized, namely each value is divided by the maximum importance obtained and then multiplied by 100. If the measure assumes a large value (≥ 60), variables have good predictive ability; on the contrary, low values (<60), zero or negative values identify non-informative variables.
By repeating this approach for each variable, we then obtain an importance measure for all potential predictors in the dataset.

Results

In preliminary results, we grow a forest of 10,000 classification trees where we model the diagnosis using heavy metals as covariates. From the forest obtained, we extract the variable importance measure identifying, as determinant of the PD, two heavy metals: copper (CU) and zinc (Zn). These results are in line with other studies that address the same topic ([3], [4]) but based on other methodologies. The main advantage of RF approach is the overperformance of the predictions obtained respect other parametric methods, producing a reliable model.

Conclusions

Compared to a classification tree or to logistic regression, RF increase the accuracy of the predictions by solving the problem of instability (small changes in the data induce big changes in the results) typical of classification trees.

Bibliography

[1] Breiman L., Random Forest. Mach. Learn., 2001; 45(1):5-32.

[2] Breiman L., Friedman J., Olshen R., Stone C., Classification and Regression Trees, Wadsworth Inc., California, 1984.

[3] Lucchini R., Albini E., Benedetti L., et al., Neurological and neuropsychological features in Parkinsonian patients exposed to neurotoxic metals. G Ital Med Lav Ergon, 2007, 29(3 Suppl):280-281.

[4] Squitti R., Gorgone G., Panetta V., et al., Implications of metal exposure and liver function in Parkinsonian patients resident in the vicinities of ferroalloy plants. J Neural Transm, 2009, 116(10):1281-1287.

SILVER RUSSELL SYNDROME: RBF-PSO MATHEMATICAL MODEL FOR THE GROWTH EVALUATION IN PATIENTS EXPOSED AND NOT EXPOSED TO GH TREATMENT

Stura Ilaria[1], Matarazzo Patrizia[2], Brondino Francesca Maria[2], De Sanctis Luisa[1,2] and Migliaretti Giuseppe[1]

[1] *Dept of Public Health and Pediatric Sciences, University of Turin, Turin, Italy.*

[2] *Pediatric Endocrinology, Regina Margherita Children Hospital, University of Turin, Turin, Italy.*

Introduction

The Silver-Russell syndrome (SRS) is a rare disorder characterized by heterogeneous clinical phenotypes and molecular mechanisms. The main clinical characteristics found in SRS patients are: a reduced prenatal growth, i.e. they are also Small for Gestational Age (SGA), a markedly decreased post-natal growth, a typical triangular face with prominent forehead and skeletal asymmetries. Thus, since molecular tests fail to identify the mutation in about 60% of the cases [1], the evidence of some recurrent phenotypic traits of the syndrome makes the diagnosis prevalently clinical [2]. Both SGA and SRS patients are treated with Growth Hormone (rhGH) treatment [3], even if little data exist so far on the response of the last category.

Aim

The study proposes a comparison of growth in SRS and SGA diagnosis with and without GH treatment, divided also by sex. A mathematical model is performed to predict the growth trend in the two diagnostic categories, also providing a predictive estimate of the growth over time.

Methods

13 SRS were paired with 13 SGA with similar characteristics. Patients were followed between August 2010 and December 2018 at the Regina Margherita Hospital (Turin, Italy). Neonatal, genetic and follow-up data have been reported.
In order to describe the auxological general trend of the patients with the two conditions, with and without GH treatment, the RBF-PSO mathematical model [4], [5], already validated in a study about the response to GH treatment in idiopathic and organic GH-deficient patients [6], was performed. The RBF-PSO is a mixed method that combines the Radial Basis Functions (RBFs) approximation technique with the Particle Swarm Optimization (PSO), a stochastic optimization tool. Briefly, it considers a theoretical curve for human growth (the gompertzian one) and estimates its parameters in order to describe the behaviour of the real data. In particular, two parameters are considered in growth description: H_∞, i.e. the carrying capacity, is maximum height that each person of his/her population could reach (a "growth hope"); r_i is the growth velocity rate, which is strictly related to the patient's characteristics.

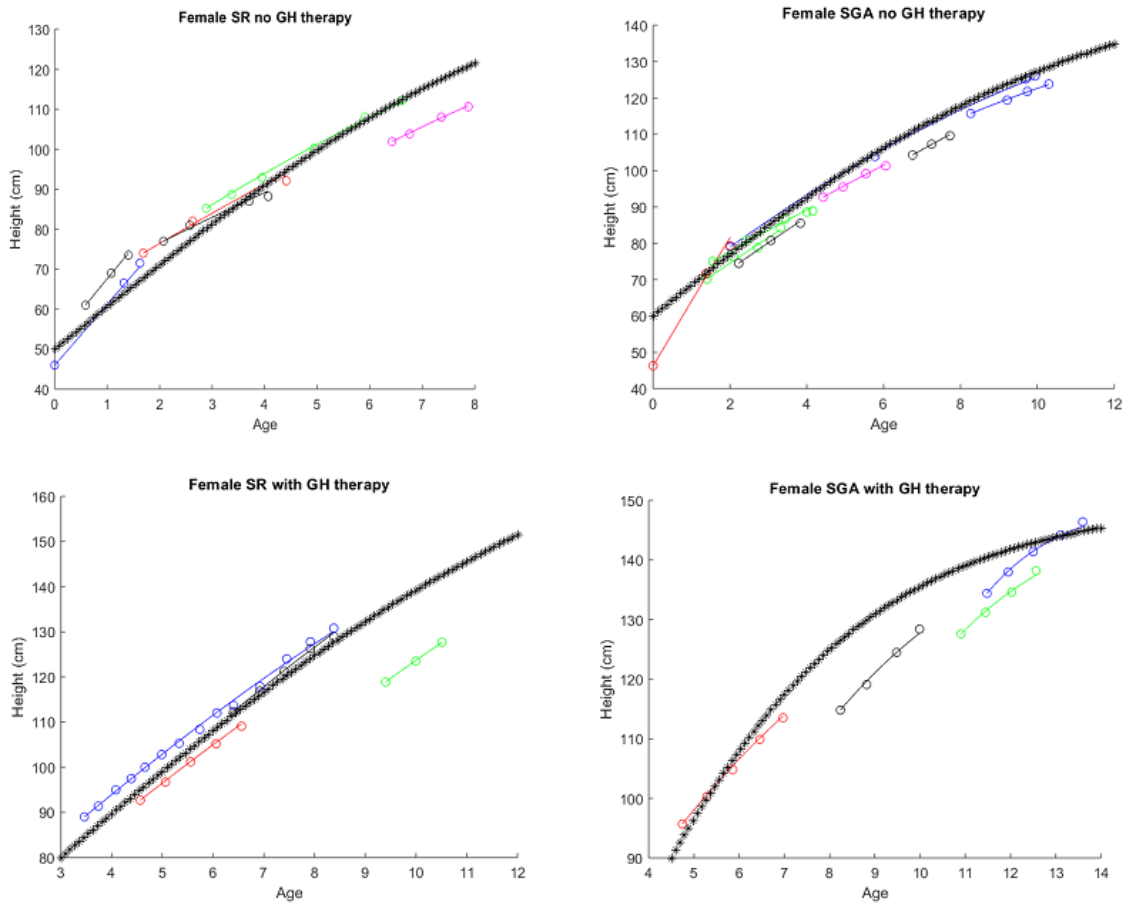
Results

SGA present higher neonatal weight ($p = 0.0035$) and length ($p = 0.0137$) than SRS, while differences on other baselines values are not significant (p-values > 0.05).
Without GH treatment, SGA and SRS patients have the same growth hope. With GH treatment, SRS increase in a significant way their growth hope, while SGA seem to improve only their growth velocity. This means that SGA will seem more advantaged by treatment. However, also SRS are advantaged by rhGH, although it is less evident. The trends could be compared also using the curves output of the model (see Figure 1). Circles are real data, straight lines are the single growth curves of each patient, stars are the median growth curve for each group (SGA or SRS females with and without GH treatment).

Conclusions

An early onset of the GH treatment could be useful in order to improve the response of both categories. The use of RBF-PSO model in clinical practice could help clinicians to see precociously a deviation from the expected curve, allowing further investigations and a rapid response to the problem.

Figure 1. *Reconstructed growth curves: circles are the real data, straight lines the curves of single patients from the RBF-PSO model and stars are the median expected patients for each group*



Bibliography

[1] E. L. Wakeling *et al.*, «Diagnosis and management of Silver-Russell syndrome: first international consensus statement», *Nat. Rev. Endocrinol.*, vol. 13, n. 2, pagg. 105–124, 2017.

[2] E. L. Wakeling *et al.*, «Epigenotype-phenotype correlations in Silver-Russell syndrome», *J. Med. Genet.*, vol. 47, n. 11, pagg. 760–768, nov. 2010.

[3] «Regione Piemonte - Sanità - Modifica della Nota AIFA 39», 01-mag-2017. [In linea]. Available at: http://www.regione.piemonte.it/sanita/cms2/farmaceutica/comunicazioniinformative/173-varie/2697-modifica-della-nota-aifa-39. [Consultato: 01-mag-2017].

[4] E. Perracchione e I. Stura, «RBF kernel method and its applications to clinical data», *Dolomites Res. Notes Approx.*, vol. 9, n. Special Issue, pagg. 13–18, 2016.

[5] E. Perracchione e I. Stura, «An RBF-PSO based approach for modeling prostate cancer», presentato al AIP Conference Proceedings, 2016, vol. 1738.

[6] G. Migliaretti *et al.*, «Long-term response to recombinant human growth hormone treatment: a new predictive mathematical method», *J. Endocrinol. Invest.*, vol. 41, n. 7, pagg. 839–848, lug. 2018.

NEXT GENERATION SEQUENCING E ANALISI DI CORRELAZIONE GENOTIPO FENOTIPO.

Gentilini D.[1,2,3], Oliveri A.[1], Fazia T.[1], Pini A.[4], Marelli S.[4], Di Blasio A.M.[3], Spada E.[1], Bernardinelli L.[1]

[1] *Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy.*

[2] *Istituto Auxologico Italiano IRCCS, Bioinformatics and Statistical Genomics Unit, Cusano Milanino, Milano, Italy*

[3] *Istituto Auxologico Italiano IRCCS, Molecular Biology Laboratory, Cusano Milanino, Milano, Italy*

[4] *Rare Disease Center, Marfan Clinic, Cardiology department, ASST-FBF-Sacco, Milano, Italy.*

Introduzione

La Sindrome di Marfan descrive una complessa patologia ereditaria (autosomica dominante) a carico del tessuto connettivo causata, nella maggior parte dei casi, da una mutazione a livello del gene FBN1, che codifica per la fibrillina-1, o del gene TGFβR2 che codifica per il recettore di tipo II del fattore di crescita TGFβ [1]. Poichè tutti gli organi contengono tessuto connettivo, le manifestazioni cliniche della MFS possono colpire diverse parti dell'organismo: vasi sanguigni, cuore, ossa, legamenti, occhi e polmoni [2]. La malattia ha una trasmissione autosomica dominante, tuttavia, in un quarto dei casi circa, sembra manifestarsi in modo sporadico senza che vi siano altri casi in famiglia. Spesso il paziente è ignaro di essere portatore di tale sindrome ed il primo sintomo, in alcuni casi la dissecazione acuta dell'aorta, può essere fatale. È quindi è molto importante individuare il prima possibile i portatori MFS poiché il trattamento in fase precoce delle alterazioni cardiovascolari prolunga notevolmente la speranza di vita di tali pazienti. La sindrome di Marfan è caratterizzata da penetranza completa ed espressività variabile, la malattia può in alcuni casi decorrere in modo completamente asintomatico mentre in altri può manifestarsi con un fenotipo molto grave. Le ragioni che sono alla base di quest'ampia variabilità devono ancora essere pienamente chiarite. I pazienti affetti da MFS presentano una serie di caratteristiche somatiche distintive: gambe e braccia molto lunghe, falangi lunghe ed affusolate con pollice allungato, scoliosi più o meno accentuata, i piedi piatti, miopia e possono presentare lussazione del cristallino. Tra i segni più problematici associati alla sindrome di Marfan si hanno il prolasso della mitrale e la dissecazione dell'aorta che soventemente porta a morte precoce. Sebbene i tratti somatici associati alla sindrome di Marfan siano particolarmente distintivi esistono alcune difficoltà nell'inquadramento clinico di tali pazienti: • Molti soggetti affetti mostrano un fenotipo lieve e difficilmente caratterizzabile • Esistono altre condizioni sindromiche come la sindrome di Loeys-Dietz (LDS), la sindrome di Ehlers-Danlos (EDS), la sindrome di Beals (BS), la sindrome di Shprintzen-Goldberg (SGS), che pur coinvolgendo altri geni sono caratterizzate dal manifestare un fenotipo del tutto simile al fenotipo Marfan tanto che sono state definite sindromi Marfan-Like [3] Al fine di inquadrare correttamente i pazienti affetti da sindrome di Marfan sono stati proposti dei criteri clinici noti con il nome di "Nosologia di Ghent". Secondo questi criteri, affinchè un paziente venga considerato affetto da sindrome di Marfan esso deve presentare un certo numero di caratteristiche cliniche considerate unitamente alla storia familiare e all'analisi genetica. Recentemente è stata inoltre proposta una revisione di tale nosologia che fra i vari cambiamenti ha previsto anche l'introduzione di un systemic score (SS), ovvero di un punteggio che varia tra zero e ventuno e che riassume le caratteristiche clinico-morfologiche dei vari apparati. Secondo questo nuovo approccio clinico un individuo è considerato affetto da Sindrome di Marfan se il suo SS è maggiore o uguale a 7. Benchè nella maggior parte dei casi la sindrome di Marfan è causata da mutazioni del gene FBN1 la presenza della mutazione non è tuttavia sufficiente a spiegare l'estrema variabilità dei fenotipi. L'approccio finora utilizzato nella maggior parte degli studi genetici è stato quello di analizzare le mutazioni presenti nel gene FBN1 cercando una correlazione fra quest'ultime e i diversi fenotipi clinici. Alcuni studi hanno cercato di identificare se mutazioni che colpiscono particolari domini della proteina possono dar luogo a fenotipi più gravi, altri hanno cercato correlazioni tra la severità del fenotipo e il tipo di mutazione genetica (missenso, splicing frameshift) [4]. Sebbene l'interesse clinico sia molto forte in tal senso, non sono emersi ad oggi evidenze

significative che consentano di poter predire il decorso della malattia o il fenotipo a partire da dati genetici. È inoltre importante considerare che molti individui risultanti positivi ai criteri di Ghent non presentano mutazioni in corrispondenza del gene FBN1. Si potrebbe quindi supporre che anche mutazioni in altri geni siano coinvolte nel processo che determina il decorso clinico della patologia. Inoltre, individui portatori della stessa mutazione, siano essi all'interno della stessa famiglia oppure non correlati, possono variare sia il momento dell'insorgenza dei sintomi, sia il coinvolgimento degli organi e sia la severità, confermando un'estrema espressività associata alla patologia.

Obiettivi dello studio

Lo scopo di questo studio è di far luce su alcune problematiche che risultano tuttora aperte: i) Cercare di capire quanto la componente genetica possa spiegare l'estrema variabilità fenotipica associata alla sindrome di Marfan, ii) cercare una correlazione tra genotipo e fenotipo aumentando il quadro di indagine molecolare, iii) considerando tutta la variabilità genetica e non solo le mutazioni causative aumentare la percentuale di variabilità fenotipica che siamo in grado di spiegare usando la genetica. L'idea originale alla base dello studio è di estendere l'analisi genetica anche a geni coinvolti direttamente e indirettamente nella genesi della matrice extra cellulare ed a geni che interagiscono con la fibrillina 1. (*FBN1*, *TGFBR1, TGFBR2, COL1A1, COL1A2, COL3A1, COL5A1, COL5A2, MYH11, NOTCH1* e *ACTA2*). Lo studio ha esteso l'analisi di correlazione genotipo-fenotipo includendo non solo le mutazioni genetiche presenti in altri geni e potenzialmente importanti per il loro effetto sul fenotipo clinico, ma ha anche considerato le varianti geniche polimorfiche SNPs. Tali varianti geniche in quanto polimorfiche non sono al momento considerate nell'ambito dell'analisi molecolare che si focalizza prevalentemente sulla ricerca di mutazioni; è opportuno tuttavia considerare che unitamente a mutazioni causative anche le varianti polimorfiche potrebbero giocare un ruolo a vario grado nella modulazione dei tratti fenotipici e sarebbe quindi opportuno valutarne l'effetto.

Metodi

Lo studio ha previsto la raccolta di due tipologie di dati: dati fenotipici e dati genotipici. Per ogni soggetto sono state registrate diverse variabili, tra cui: • variabili anagrafiche • variabili cliniche • variabili relative ad interventi chirurgici subiti dall'individuo a causa della MFS e presenza di protesi cardiache • fattori di rischio cardiovascolari (abitudine al fumo, patologie) • caratteristiche morfologiche tipiche della Sindrome di Marfan come aracnodattilia, segno del polso e del pollice, lussazione cristallino, presenza di aneurismi aortici, dissecazione aortica, difetti scheletrici, ectasia della dura. Per garantire una buona qualità dei dati fenotipici si è deciso di eliminare dallo studio tutti quei soggetti con un eccesso di dati mancanti e/o informazioni in contrasto tra loro. La raccolta dei dati genotipici ha previsto l'utilizzo di dati NGS. A partire da diversi files formato vcf che contengono l'elenco delle varianti genetiche identificate per ogni soggetto, attraverso una funzione informatica, implementata attraverso il software R, è stato prodotto un database genotipico in formato "pedegree file", cioè un unico database contenente i dati genetici di tutti i soggetti in modo integrato. Un controllo di qualità è stato eseguito anche sul database genotipico allo scopo di eliminare le varianti geniche affette da errori di sequenziamento.

Si tratta di uno studio osservazionale condotto in una popolazione costituita esclusivamente da soggetti affetti. Sono stati arruolati pazienti con sospetto clinico di sindrome di Marfan che si sono sottoposti ad un'approfondita visita specialistica. I dati fenotipici registrati durante la visita specialistica (in gran parte presenza assenza del segno clinico) unitamente ai dati genetici ottenuti sono stati analizzati utilizzando un approccio caso-controllo. Dopo la fase di controllo di qualità dei dati è stata condotta un'analisi descrittiva sui tratti fenotipici e sulla variabilità genetica. L'analisi delle componenti principali PCA è stata applicata per ridurre la complessità dei fattori che caratterizzano il fenotipo. L'analisi statistica ha previsto di creare un modello in cui l'effetto delle varianti genetiche rare e quello delle varianti comuni è stato stimato con metodi appropriati. Per le varianti comuni è stato misurato l'effetto sui tratti clinici usando un comune test di

associazione mentre per le varianti rare è stato impiegato un'analisi burden [5]. I due effetti sono stati quindi integrati con un modello di analisi basato su equazioni strutturali SEM allo scopo di valutare la percentuale di variabilità fenotipica spiegata dall'intera componente genetica e non solo dalla mutazione potenzialmente causale.

Risultati

L'analisi NGS degli 11 geni candidati in 181 campioni non correlati ha identificato 361 (28%) varianti comuni (MAF> 0,05), 128 varianti a bassa frequenza (10%) (0,01 <MAF <0,05) e 784 varianti rare (62%) (MAF <0,01). Tra queste 1273 varianti genetiche 283 esoniche (22,2%), 903 introniche (70,9%), 3 varianti di splicing (0,2%) e 84 erano localizzate nelle regioni UTR (6,6%). Le varianti esoniche sono risultate per lo più missense (47,7%) e sinonime (42,4%) mentre le varianti stop e indels sono risultate rispettivamente del 4,2% e del 5,6%. Sulla base di standard e linee guida tutte le varianti sono state classificate anche come benigne (87,8%) patogene (2,6%) (incluse anche probabili-patogene) e sconosciute (9,6%). L'analisi di associazione delle varianti comuni è stata studiata tramite una regressione logistica con ogni segno clinico a sua volta come risultato e ogni variante genetica con un MAF> 0,05 come covariata, aggiustando per età e il numero di varianti rare benigne, patogene e sconosciute. Dopo una correzione multipla, uno SNP del gene COL1A1, rs2075557, è risultato associato a scoliosi (OR = 0,2533, valore p aggiustato = 0,04), inoltre altri 4 SNP in COL1A1 sono risultati anche associati a scoliosi, sebbene a livello marginale: rs2696247 (OR = 0,266, valore p regolato = 0,07), rs2075554 (OR = 0,267, valore p regolato = 0,102), rs2277632 (OR = 0,295, valore p regolato = 0,121), rs2586494 (OR = 0,305, valore p regolato = 0,141).
L'effetto di varianti rare su tratti fenotipici è stato testato seguendo due strategie. La prima strategia ha studiato il loro impatto complessivo sui fenotipi clinici. Questa analisi ha mostrato che il numero complessivo di varianti genetiche rare è significativamente associato al numero di manifestazioni cliniche (RR = 1.018, valore p = 0.018). Sulla base di questo risultato, è stato utilizzato lo stesso modello di regressione per indagare il ruolo delle varianti rare considerando anche la loro classificazione clinica (cioè patogenica, sconosciuta e benigna). Solo il numero di varianti rare patogeniche e sconosciute è risultato significativamente associato al numero di manifestazioni cliniche (RR = 1.533, valore p = 9.87e-06, RR = 1.045, valore p = 0.024 rispettivamente). Per quanto riguarda la seconda strategia, l'associazione tra tratti fenotipici e varianti genetiche rare è stata esplorata a livello genico tramite SKAT-O e usando l'età come covariata. Questa analisi ha confermato che varianti rare patogene e sconosciute hanno un impatto significativo su diverse manifestazioni cliniche mentre varianti rare benigne non erano significativamente associate con nessuno dei tratti fenotipici considerati. Inoltre, per diversi tratti fenotipici è stato osservato anche un contributo poligenico di varianti rare. L'effetto combinato di varianti comuni e rare su ciascun segno clinico è stato esplorato con SEM. Per ciascun tratto fenotipico, i risultati ottenuti dall'analisi delle varianti comuni sono stati raccolti in Poligenic Risk Scores specifici del segno. PRS e risultati significativi ottenuti dalla rara analisi delle varianti sono stati infine utilizzati come variabili esogene nel SEM.

Conclusioni

Questa studio ha mostrato che diversi tratti fenotipici caratterizzanti la sindrome di Marfan e quadro clinico marfan-Like sono significativamente influenzati sia da varianti genetiche comuni che rare. Lo studio ha dimostrato che la maggior parte della variabilità genetica che viene identificata da analisi NGS è rara e che nella maggior parte dei casi l'azione delle varianti rare sul fenotipo è poligenica. Impiegando il modello di analisi proposto in questo studio siamo stati in grado di aumentare la percentuale di varianza del fenotipo che siamo in grado di spiegare partendo dai soli dati di genetica migliorando la correlazione genotipo-fenotipo che fino ad oggi è stata inconcludente.

Bibliografia

[1] Pepe G, Giusti B, Evangelisti L et al. Fibrillin-1 (FBN1) gene frameshift mutations in Marfan patients: genotype-phenotype correlation. Clinical Genetics, 2001 59, 444–450

[2] Dietz HC, Pyeritz RE. Mutations in the human gene for fibrillin-1 (FBN1) in the Marfan syndrome and related disorders. *Hum Mol Genet*. 1995;4 Spec No:1799-1809. http://www.ncbi.nlm.nih.gov/pubmed/8541880.

[3] Child AH. Non-cardiac manifestations of Marfan syndrome. *Ann Cardiothorac Surg*. 2017;6(6):599-609. doi:10.21037/acs.2017.10.02

[4] Faivre L, Collod-Beroud G, Callewaert B et al. Clinical and mutation-type analysis from an international series of 198 probands with a pathogenic FBN1 exons 24-32 mutation. *Eur J Hum Genet*. 2009;17(4):491-501. doi:10.1038/ejhg.2008.207

[5] Ionita-Laza I, Lee S, Makarov V, Buxbaum JD et al. Sequence Kernel Association Tests for the Combined Effect of Rare and Common Variants. *Am J Hum Genet*. 2013;92(6):841-853. doi:10.1016/j.ajhg.2013.04.015

APPLICAZIONE DI UN INDICATORE DI DIPENDENZA NON LINEARE NELL'ANALISI STATISTICA DI TRACCIATI ELETTROCARDIACI.

Di Rocco A, Vestri A.R., Nofroni I., Di Traglia M., Centurion Aznaran C*., Pandozi C*.

Dipartimento di Sanità Pubblica e Malattie Infettive, Università degli Studi di Roma " Sapienza"

**Divisione di Elettrofisiologia Cardiaca, Ospedale S. Filippo Neri, Roma*

Introduzione

Nell'analisi statistica dei fenomeni quantitativi dal punto di vista dinamico, univariati oppure multivariati (scalari o vettoriali), è fondamentale valutare la dipendenza tra le misure dello stesso fenomeno, o tra lo stesso e fenomeni differenti, in più momenti temporali diversi. Quando si prende in considerazione l'aspetto temporale di qualsiasi fenomeno, i dati ad esso relativi sono chiamati, nella letteratura statistica, serie storica (univariata, se si tratta di un singolo fenomeno e multivariata se i fenomeni sono più di uno). Le serie storiche, al fine di poterne analizzare la dipendenza o interdipendenza, sono rappresentate in un particolare spazio matematico denominato spazio delle fasi. L'analisi di tale spazio permette di costruire modelli matematici in grado di leggere le informazione sulla struttura di dipendenza ed interdipendenza nella stessa serie o tra le differenti serie che compongono il sistema oggetto di analisi. In ogni caso bisogna utilizzare misure di dipendenza tra fenomeni quantitativi. Nella letteratura statistica esistono numerose misure della dipendenza tra fenomeni quantitativi (r di Pearson, correlazione di Spearman, correlazione di Kendall, ecc.) ma esse sono utilizzabili solo quando la relazione è di tipo lineare (r di Pearson) oppure quando essa è monotona (correlazione di Spearman, correlazione di Kendall). Quando le relazioni tra fenomeni quantitativi sono non lineari e non monotone, le misure menzionate falliscono. Ossia, possono segnalare assenza di relazione anche quando essa è, invece, presente e molto intensa. L'importanza di misurare correttamente l'intensità di una relazione non lineare tra due o più fenomeni, risiede nel fatto che tale misura viene, poi, utilizzata nelle analisi statistiche multivariate come, ad esempio, l'ACP (Analisi in Componenti Principali, Regressione, Analisi di Serie Storiche e Territoriali, ecc..) con lo scopo di costruire modelli matematico-statistici con finalità di predizione e controllo. Esistono degli indici di Correlazione Curvilinea ma sono basati su modelli di interpolazione dei dati (spesso modelli polinomiali), ossia sono modello-dipendenti e, quindi, possono essere distorti. L'indice che proponiamo è non-parametrico e differisce dai MDI (Mutual Information Index) come l'entropia di Shannon o la divergenza di Kullback-Leibler, o il CHI quadro di Pearson in quanto non si basa sulle frequenze marginali o sulle congiunte ma sul pattern dei punti campione nel reticolo che ricopre l'area (o lo spazio) generata dallo scatterplot dei dati nel piano cartesiano (o nello spazio).

Obiettivi

L'obiettivo di questo lavoro è l'applicazione di un indicatore di dipendenza non lineare tra fenomeni quantitativi (M. Di Traglia, 2012) [1], a tracciati elettrocardiaci al fine di trarre informazioni su eventuali anomalie nei differenti cicli di cui i tracciati stessi sono composti. Le informazioni che si otterranno, attraverso l'analisi non lineare, permetteranno di stimare le relazioni tra i differenti tracciati e migliorare la diagnostica sulla presenza di eventuali patologie, rispetto all'analisi lineare. L'indicatore che si intende applicare è basato sull'Integrale di Correlazione (Grassberger & Procaccia, 1983) [2], il quale è stato introdotto come strumento per lo studio dello spazio delle fasi di un sistema dinamico di tipo non lineare. In letteratura, l'integrale di Correlazione ha trovato applicazione prevalentemente nell'ambito dell'analisi statistica dei sistemi dinamici spazialmente estesi che presentano relazioni non lineari e dinamiche di natura caotica. L'Integrale di Correlazione si basa sul concetto di "ricoprimento" di uno spazio metrico (o metrizzabile) attraverso reticoli regolari come una funzione del rapporto tra il numero di celle che contengono punti

(osservazioni) ed il totale delle celle che ricoprono l'intero spazio. Tale rapporto diminuisce all'aumentare del numero di celle del reticolo e tende a zero quando il numero delle celle tende all'infinito oppure, allo stesso modo, quando la dimensione della celle tende a zero. Il tendere a zero della dimensione delle celle, coinvolge problematiche di natura teorica che, in questa sede, non verranno considerate in quanto non interessano i casi pratici.

Metodi

L'indice di dipendenza–associazione tra caratteri quantitativi, (che chiameremo ISSO -Index of Structural Self Organitation), è costruito tenendo conto dell'area sottesa alla curva dell' Integrale di Correlazione. La sua espressione formale è data da:

$$ISSO = 1 - A^{-1} \int_{\min(l)}^{\max(l)} C(l)dl$$

Dove l rappresenta l'ampiezza delle celle, C(l) l'Integrale di Correlazione in funzione dell'ampiezza delle celle al cui interno figura almeno 1 punto, A indica l'intera area (nella quale è racchiusa la curva C) ed è rappresentata dal prodotto tra il numero di celle che ricoprono l'area per la loro ampiezza l (si dimostra che è costante in quanto all'aumentare del numero delle celle diminuisce la loro ampiezza). Recentemente è stato applicato l'Integrale di Correlazione per lo studio della dipendenza non lineare in genetica (Y. Ann Chen et.al.2010)[3] ma con una metodologia completamente differente.

Risultati

Tale indicatore è stato applicato alla dinamica dei tracciati elettrocardiaci nelle 12 derivazioni, trattate come una serie storica vettoriale. La rappresentazione nello spazio delle fasi ha consentito di utilizzare l'Analisi In componenti principali (ACP) per costruire degli indicatori delle frequenze ed ampiezze dei segnali. Attraverso l'ISSO sono state stimate le funzioni di auto correlazione ed autocorrelazione incrociate e confrontate con quelle che si sarebbero ottenute attraverso l'analisi lineare. I risultati evidenziano che alcuni di questi tracciati a particolari ritardi presentano pattern di correlazione completamente differenti se analizzati linearmente. Il significato clinico di tali differenze è ancora in fase di valutazione in quanto necessita di ulteriori dati su pazienti con particolari patologie confrontati con pazienti senza patologie cardiache.

Conclusioni

Durante la fase sperimentale, sono state effettuate delle simulazioni numeriche sulle differenti tipologie di dipendenza statistica (lineare, non lineare monotona, non lineare non monotona) alle quali è stato applicato l'indicatore ISSO. Nelle simulazioni, nel caso non lineare non monotono, la misura della intensità del legame con ISSO mostra una sostanziale differenza con i risultati ottenuti applicando, alle stesse simulazioni, gli indicatori classici. Tali differenze diminuiscono se il legame è non lineare monotono (confrontabile con la correlazione tra ranghi) e coincidono se il legame è lineare. Lo stesso comportamento ha avuto nell'applicazione a dati reali sui tracciati ECG presi in considerazione. L'indicatore di dipendenza non lineare (ISSO) applicato ai tracciati elettrocardiaci ha consentito di quantificare l'informazione sulle intensità delle relazioni non lineari tra i differenti cicli dei quali è composto ed a differenti istanti di tempo. L'applicazione effettuata su dati reali, ha riguardato, lo studio dei tracciati elettrocardiaci in pazienti affetti da alcune patologie ed è ancora in fase di elaborazione. Uno studio su un campione più ampio di casi permetterebbe di raggiungere conclusioni generalizzabili in ambito medico - epidemiologico.

Bibliografia

[1] Chen Y. A., Jonas S. Almeida, Adam J. Richards, et.al. (2010); A Nonparametric Approach to Detect Nonlinear Correlation in Gene Expression; *Journal of Computational and Graphical Statistics, Volume 19, Number 3. (EN) H. Yang, S. Bukkapatnam , R. Komanduri, Spatiotemporal representation of cardiac vectorcardiogram (VCG) signals, in BioMedical Engineering OnLine, 2012, DOI:10.1186/1475-925X-11-16, PMC 3439290, PMID 22463593.*

[2] Di Traglia M. (2012); Cellular Automata and Random Field: statistical analysis of complex space-time systems; *Lecture Notes in Computer Science – © Springer-Verlag Online ISBN 978-3-642-33350-7. https://doi.org/10.1007/978-3-642-33350-7_68*

[3] Grassberger, P., and Procaccia, I. (1983), "Characterization of Strange Attractors," *Physical Review Letters*, 50, 346–349. [553,554].

[4] Segreti L., Lavalle C., M. Pandozi C., et al,(2019); Characterization of multi-component potentials inside the Koch"s Triangle with ultra-high density mapping system *EP Europace, Volume 21, Issue Supplement_2, March 2019, Pages ii251–ii531;Conference: EHRA 2019.*

[5] Pandozi C., Ficili S., Centurion Aznaran C., et. Al. (2018); Reasons for successful clinical outcome following pulmonary vein isolation despite lack of persistent LA-PV conduction block. *Official Journal of world Society oh Arrithmias; Pacing and Clinical Electrophysiology; Online ISSN:1540-8159.*

BIG DATA, OBSERVATIONAL RESEARCH, AND P-VALUE: A RECIPE FOR FALSE POSITIVE FINDINGS? A STUDY OF SIMULATED AND REAL PROSPECTIVE COHORTS

Giovanni Veronesi[1], Guido Grassi[2], Giordano Savelli[3], Piero Quatto[4], Antonella Zambon[5]

[1] *Centro Ricerche in Epidemiologia e Medicina Preventiva, Dipartimento di Medicina e Chirurgia, Università degli Studi dell'Insubria, Varese;*

[2] *Clinica Medica, Dipartimento di Medicina e Chirurgia, Università di Milano-Bicocca, Milano;*

[3] *UO Medicina Nucleare, Fondazione Poliambulanza Istituto Ospedaliero, Brescia;*

[4] *Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa, Università di Milano-Bicocca, Milano;*

[5] *Dipartimento di Statistica e Metodi Quantitativi, Università di Milano-Bicocca, Milano.*

Background

The documented decline in participation rates observed in epidemiological studies, or the adoption of self-selected samples such as in studies based on Mobile-research platforms, may exacerbate the risk of selection bias generally present in observational research. In combination with a large sample size, selection bias may generate a number of false positive results, as standard inference methods may fail to adequately control the false discovery rate. The "empirical calibration of p-value" method relies on a set of variables with no known association with the study outcome ("negative controls") to obtain a distribution of effect sizes under the null hypothesis ("empirical null distribution") in the available data [1-2]. Once estimated from the data, the parameters of the null empirical distribution are used to calibrate the test statistic and the p-value for the exposure of interest, assuming that bias arises from the same distribution (aka "exchangeability" assumption). We investigate whether the method is able to preserve type-I error in presence of selection bias and in the context of survival analysis, using simulated and real-life prospective cohorts.

Methods

In simulations, we conceived selection bias as an underlying mechanism Z driving participation in the study sample used to make inference on the true relationship between an exposure X and an outcome Y. In other words, the vector of data (xj, yj) for the j-th individual in the population of interest is observed with probability p_j which depends on Z: p_j = Prob[missing(x_j, y_j) =1| z_j)]. The values of p_j were chosen to mimic cohort studies with 50% participation rate under 2 distinct selection mechanisms, in which specific groups of participants (e.g. healthier or sicker) are more likely to be included in the sample. We simulated a total number of 8 scenarios, due to the combination of sample size (n=5000 or 20000), amount of selection bias (mild or moderate) and two selection mechanisms. The same scenarios were considered after having relaxed the exchangeability assumption, for a total number of 16 simulation settings. We estimated the log of the hazard ratio $\widehat{\beta}_X$ for the association between X and Y from Cox regression models on the available data after having applied the selection bias mechanisms. We computed rejection rate, as the number of p-values below 0.05 on the total number of replications (R=200); bias, as the average over R of the difference between the numerator in the t-statistic for $\hat{\beta}_X$ and the true value β_X without and with empirical calibration; and standard deviation, as the average over R of the denominator of the t-statistic for $\hat{\beta}_X$ without and with empirical calibration.
We also apply the calibration method to a real-life application on predictors of cancer mortality using data from 6976 participants to four population-based cohorts (three MONICA-Brianza surveys and the PAMELA study) with a median follow-up of 17 years. At baseline, a fasting serum blood sample was drawn from each participant and n=34 markers were measured on the fresh specimens in a centralized laboratory. To investigate the association between blood markers and cancer mortality in the general population, we

provided the list of available markers to a clinical oncologist (SG), initially blinded about the objective of the research. After a review of the relevant literature, SG indicated n=21 negative controls, and n=13 positive associations. We first estimated the hazard ratio of cancer mortality for 1 standard deviation-increase for each of the 34 markers from Cox regression models, adjusting for age and gender. Then, to investigate type-I error, we applied the empirical calibration of p-value method to negative controls using a leave-one-out cross-validation approach [2]. Simulations and analyses were performed using standard procedures in the SAS Software 9.4 release.

Results

The standard inference without calibration of p-value had a type-I error rate above the nominal 5% level in 15 out of the 16 simulated settings. With the largest sample size, a null association with hazard ratio=1.05 had 42.5% chances of being statistically significant. Conversely, calibration of p-value preserved the 5% nominal level in all the simulation settings, with as low as 10 negative controls and under both selection bias mechanisms. Moreover, when “exchangeability” assumption is verified, calibration reduced bias by 80%-90% as compared to the raw estimator. In the real-life application, out of n=21 blood markers with no association with cancer mortality according to literature, n=15 (71%) had a p-value less than 0.05 in Cox models adjusted for age and gender. After calibration, only 1 (4.8%) remained statistically significant.

Conclusions

In observational research with contemporary presence of large sample sizes and selection bias, the standard application of traditional inference tools may fail to adequately control for type-I error. The generation of an elevated number of false positive discoveries to be further screened for biological plausibility, clinical relevance and external validity negatively may sensibly slow down time from discovery to real-setting implementation, and negatively affect the impact of large and costly observational research studies. The empirical calibration of p-value with a reasonable number of negative controls constitutes a robust method that can be applied from available data to preserve type-I error as close as possible to the desired nominal level. Our real-data application provided some guidance for practitioners, in particular on the choice of negative controls.

References

[1] Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Stat Med*. 2014;33(2):209-18.

[2] Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Robust empirical calibration of p-values using observational data. *Stat Med*. 2016;35(22):3883-8.

BIVARIATE SPLINE MODELS TO ASSESS THE JOINT EFFECT OF INTENSITY AND DURATION OF ALCOHOL DRINKING AND CANCER OF THE ORAL CAVITY: A FOCUS ON A NOVEL APPROACH

Di Credico Gioia[1], Polesel Jerry[2], Dal Maso Luigino[2], La Vecchia Carlo[3], Pauli Francesco[1], Torelli Nicola[1], Edefonti Valeria[3] and the INHANCE consortium investigators[1]

1 Department of Economics, Business, Mathematics and Statistics, University of Trieste, Trieste, Italy.

2 Cancer Epidemiology Unit, Centro di Riferimento Oncologico di Aviano (CRO) IRCCS, Aviano, Italy.

3 Branch of Medical Statistics, Biometry and Epidemiology "G. A. Maccacaro", Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milano, Italy.

Introduction

When modeling the relationship between a response and some continuous covariates, assuming linearity may be too restrictive in many contexts. Naive solutions to overcome this limitation, such as categorisation of the predictor, have well-known drawbacks. Polynomial regression can be a flexible solution when visual inspection of the variables is available. However, in complex scenarios this approach can be unfeasible and selecting the polynomial degree may be challenging. Indeed, the higher the degree, the higher the risk of overfitting. In addition, polynomial bases are highly sensible to outliers, and the identification of outliers in multivariate contexts poses extra challenges itself.
A viable alternative is represented by spline functions. They are defined as piecewise polynomials with a fixed degree, whose joint points are called knots. Splines are highly flexible and have proven to be excellent approximation tools. Indeed, varying the number and position of knots may lead to extremely different shapes.
The two factors that affect the most the flexibility of spline functions are the number of distinct knots and the polynomial degree. To manage the flexibility of regression splines, a classical approach consists in using an optimising criterion with a suitable penalisation to control the roughness of the function. In this case, the number and position of the knots is not as crucial and potential knots are evenly-spaced on the predictor range or are placed on the quantiles of the predictor. Then, the optimal number of knots is selected as the minimiser of some criteria through cross-validation or, alternatively, a penalisation term on the spline coefficients is added. In the literature, the latter option is usually preferred since the former setting may require high computational effort with complex models. The penalisation term usually acts on the curvature of the function leading to shrunk spline coefficients. The term that calibrates the amount of penalisation can be selected using cross-validation criteria, Akaike Information Criterion (AIC), Generalized Cross-validation (GCV) or their derivations [1,2].
In epidemiological studies, the number and position of knots usually have an important meaning. Therefore, in these applications, special attention should be posed to techniques that allow to choose the number and position of knots. Among others, we mention the use of variable selection to choose basis function, or samplers that allow for varying dimension of the parameter space. Variable selection is usually applied to a set of knots to choose from, thus number and positions of knots are determined and fixed. There are several examples of previous epidemiological applications that go in this direction. However, this approach provides only a partial view on the problem. Transdimensional methods allow to estimate both the number and the position of the knots and therefore they should be preferred. Greater methodological complexity, several hours of computing, and possible convergence issues are frequent drawbacks of these approaches, especially when the dimension of the parameter space is high.

[1] See the full list of authors at the end of the manuscript.

Here, we will follow one of the most recent approaches to variable selection in a Bayesian context. Estimating the positions of the knots is not an easy task and, for a fixed degree, regression coefficients and locations of knots have to be estimated simultaneously, turning the standard estimation procedure into a non-linear optimisation problem.

Aims

The aim of the present work is to: 1. introduce a two-step Bayesian procedure within the semiparametric generalised linear model framework to be applied in epidemiological studies where the effect of a continuous exposure on risk is under investigation; 2. show how this framework is applied in a bivariate context where the aim is to modeling the joint effect of intensity and duration of alcohol drinking in cancer of the oral cavity.

Methods

Model. The present work assumes the following model:

$$E[y_i] = g^{-1}(\eta_i), \qquad \eta_i = z_i\alpha + f(x_i), \quad \text{for } i=1,\ldots,n,$$

where Y is the dependent variable, g is the link function and η is the linear predictor. Furthermore, Z is the covariate vector that enters linearly in the model, α is the vector of regression coefficients and X is a continuous variable affecting the response through a smooth function f: $R \rightarrow R$, described with a spline with few knots.
We restrict our analysis to those situations in which a low number of knots can be adequate and their positions are directly interpretable and of specific interest for the analysis. A typical example arises when truncated power basis of order one is used. In this case, positions of knots represent change points in the slope. One of the main drawbacks of the truncated power basis representation is that the basis is not orthogonal, which can lead to numerical instability and slow convergence of the optimisation algorithm. Keeping a low number of knots alleviates the issue [1]. Let then:

$$f(x)=\beta_0 +\beta_1 x+\sum_k \gamma_k(x-\xi_k)_+, \quad k=1, \ldots, K$$

where ξ_k is the position of the k-th knot, K is the total number of knots and the last term is the truncated linear function.
For a given number of knots, parameters estimation reduces to maximum likelihood optimisation with derivations of Newton-Raphson method.
The usual approach is, therefore, to: 1. choose the knot locations using standard criteria (such as quantiles of the predictor distribution, uniformly distributed knots on the range of the independent variables and user-defined knots following *a priori* information [1]; 2. estimate models with a different number and location of knots and compare them through standard criteria, such as AIC, or GCV. This procedure often results in an insufficient ability to discriminate among all competing models.
Free-knot regression splines. A possible extension of the previous approach is to estimate knot locations together with the other regression coefficients.
Within a maximum likelihood approach, exploration of the objective function surface could locate local maxima, and this may lead to solutions that strongly depend on the starting values. A Bayesian specification of the model and exploration of the posterior distribution, possibly by using Markov Chain Monte Carlo simulations, could be much more effective.

1. No variable selection approach. A tricky approach to estimate both number and location of knots is to estimate several models with free knot locations and an increasing but fixed number of knots. A reasonable constraint of ordering of knots should be included in the prior structure. We assume consistently that:

$\xi_k \sim \text{Unif}(\min(x), \max(x))$, subject to $\xi_k \leq \xi_{k+1}$, for k = 1,...,K.

Diffuse priors on the regression and spline coefficients are chosen. In particular,

α i.i.d. $\sim N(0,\sigma_\alpha)$ and β i.i.d. $\sim N(0,\sigma_\beta)$

where both $\sigma\alpha$ and $\sigma\beta$ are selected such that the prior distribution is weakly informative. We will refer to this model as the no variable selection (NVS) model.
Models with an increasing number of knots are compared on the basis of diagnostic tools such as trace plots and Rhat to check convergence of parameters. Information criteria are used to choose the best model among the estimated ones. The main drawback of this procedure is that a large number of models have to be considered, with huge computational effort in high dimensional problems.

2. Stochastic search variable selection approach. Checks on simulated data show that the NVS approach performs well (i.e. reasonable estimates and convergence of chains for knot-related parameters) only if the number of specified knots is lower or equal to the true one. This prompted us to consider a two-step procedure:

- select the optimal number of knots considering a large, possibly overparameterised model;

- fit the final model on a restricted set of knots by simultaneously estimating knot locations, regression and spline coefficients.

In the first step, we estimate a model having more knots than reasonably warranted. This leads to an overparameterised model where the posterior of some knot locations are expected to concentrate at the limits of the predictor range.
To assess convergence of the spline parameters, our advice is to run several chains and look at the results of each chain separately. Indeed, overparameterising the model may lead to chains that converge at different points. As an example, suppose that the true number of knots is two and we simulate two chains to fit the model with 5 ordered knots. It can happen that in the first chain the first and the second knot parameters, say ξ_1 and ξ_2, converge on the values of the true knots, while in the second chain the second and third knots parameters, ξ_2 and ξ_3, converge on the right values. Looking at the posterior results distinctly for each chain would let us to properly recognize the presence of two knots.
Since each knot location is uniquely linked to a spline coefficient, we evaluate the potential presence of a knot based on the analysis of the posterior distribution of the associated coefficient.
The concept underlying the proposed methodology is to perform variable selection on the basis functions, for this purpose we employ spike-and-slab priors.
Several versions of this approach have been proposed in the literature [3] but, generally speaking, prior distributions for the regression coefficients are defined with a spike component, usually highly concentrated around zero, and a diffused slab part. This is the case of the *stochastic search variable selection approach (SSVS)* that defines a mixture distribution for each parameter that has to be selected [3]. This type of methodology gives us the opportunity to evaluate the presence of a variable through the marginal posterior distribution of the mixing proportion.
Starting from the NVS model specification, we set a prior distribution on each spline parameter γ_k such that

$\pi(\gamma_k|\lambda_k) = \lambda_k N(0, \sigma_{sl}) + (1 - \lambda_k)N(0, \sigma_{sp})$,

where the mixing proportion $\lambda_k \sim$ Beta(a,b), with a = b. Standard deviations of the two mixture components, σ_{sl} and σ_{sp}, are chosen to be respectively large and small. Appropriate values have to be evaluated taking into account the unit of measurement of the dependent and independent variables.
Our method adapts the modified *SSVS* approach by assuming λ_k to be dependent on the knot location ξ_k. The prior distributions of the ordered knots remain defined as Uniform on the support of the variable X and independent from both the mixing proportion λ and the coefficient γ. Each coefficient γ_k, conditioned on the mixing parameter λ_k follows the same mixture distribution of two components specified in the SSVS approach described above, while each element of the mixing proportion vector λ is now defined as:

$\lambda_k|\xi_k \sim$ Beta(a,b_k),

where a is a positive but very small value and b_k : [min(x); max(x)] ⟶ [a; 1 + a] is a U-shaped even function of the knot location which returns values close to 1 + a when the knot is near the boundaries of the variable, while it is almost uniform and close to a elsewhere. In practice, the prior for the mixing parameter swings between a Beta U-shaped distribution (when the knot location is on plausible values) and a Beta distribution highly concentrated on zero (when the knot is close to the boundaries).
All the other prior distributions remain defined as in the previous model specification.
Bivariate extension of the SSVSξ model. In this section we apply the proposed *SSVSξ* methodology to pooled data derived from the International Head and Neck Cancer Epidemiology Consortium [4] to assess the joint effect of intensity and duration of alcohol drinking on the risk of cancer of the oral cavity. We therefore introduce the bivariate extension of the *SSVSξ* model.
In detail, we specify a semiparametric logistic model for the Bernoulli-distributed random variable Y that is explained by confounding factors Z (which enter linearly in the model) and variables X (alcohol intensity) and W (alcohol duration) which enter in the model as a bivariate linear spline function represented through the truncated linear basis:

$$f(x,w) = \beta_0 + \beta_1 x + \beta_2 w + \beta_3 xw + \sum_{kx=1}^{K_x} \gamma_i^{(x)} \left(x - \xi_{kx}^{(x)}\right)_+ + \sum_{ky=1}^{K_w} \gamma_j^{(w)} \left(w - \xi_j^{(w)}\right)_+ + \sum_{kx=1}^{K_x} \gamma_i^{(xw)} \left(x - \xi_{kx}^{(x)}\right)_+ w + \sum_{ky=1}^{K_w} \gamma_j^{(wx)} \left(w - \xi_{ky}^{(w)}\right)_+ x + \sum_{i=1}^{K_x}\sum_{j=1}^{K_w} \gamma_{kx,ky}^{(xw)} \left(x - \xi_{kx}^{(x)}\right)_+ \left(w - \xi_{ky}^{(w\text{å})}\right)_+$$

Prior distributions on the knots positions are defined as before as Uniform distributions on the range of the related predictor, subject to ordered constraint. Similarly, priors on the regression coefficients α and on the spline coefficients β are defined as weakly informative T distributions with parameters (3,10) and (3, 2.5). In addition we have to specify prior distributions on the spline coefficients γ and on the mixing parameters λ and we choose the following ones:

π(γkx|λkx)=λkxN(0,100)+(1−λkx)N(0,0.1),

π(γkw|λkw)=λkwN(0,100)+(1−λkw)N(0,0.1),

π(γ2,kx|λ2,kx)=λ2,kxN(0,100)+(1−λ2,kx)N(0,0.1),

π(γ2,kw|λ2,kw)=λ2,kwN(0,100)+(1−λ2,kw)N(0,0.1),

π(γ3,kx,kw |λ3,kx,kw) = λ3,kx,kw N(0, 100) + (1 − λ3,kx,kw)N(0, 0.1),

and:

$\lambda_{kx}|\xi_{kx} \sim$Beta(0.5,$b_{kx}$),

$\lambda_{kw}|\xi_{kw} \sim$Beta(0.5,$b_{kw}$),

$\lambda_{2,kx}|\xi_{kx} \sim$Beta(0.5,$b_{kx}$),

$\lambda_{2,kw}|\xi_{kw} \sim$Beta(0.5,$b_{kw}$),

$\lambda 3_{,kx,kw}|\xi_{kx,kw} \sim$ Beta(0.5, min (b_{kx} , b_{kw})),

where b_{kx} : [min (x); max (x)] → [0.5; 1.5] and b_{kw} : [min (w); max (w)] → [0.5; 1.5].

Due to the high number of parameters that have to be estimated and the dimension of the dataset, we chose to test the bivariate extension of the *SSVSξ* methodology fixing the number of estimated knots to two for both risk factors. We run 10 chains with 2,000 iterations each. Initial values for the knot location parameters are chosen uniformly spread on the linked predictor range, while regression and spline parameters are initialised at 0. The mixing parameters are initialised at 0.9 to support complete exploration of the posterior distribution.

The INHANCE Consortium

The INHANCE Consortium was established in 2004 to elucidate the aetiology of head and neck cancer through the collection of large-scale epidemiological studies all over the world [4]. Within data version 1.5, there are 33 case-control studies which provide information on alcohol intensity and duration. All of them provide information on cancer of the oral cavity in current drinkers. The total sample size includes 4839 cancer cases and 25,871 controls.
Calculations were carried out using the open-source Stan program [5] within the open-source R program [6].

Results

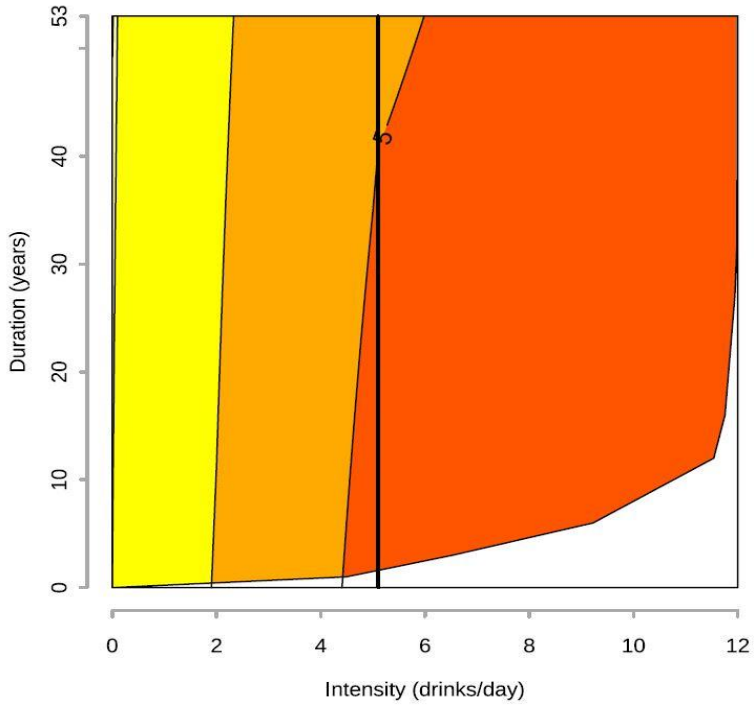
Figure 1 shows the bi-dimensional and three-dimensional representations of oral cancer risk according to different combinations of intensity and duration of alcohol drinking. After model selection, the model showed 1 knot for intensity (around 5 drinks/day) and no knots for duration. At the highest levels (from 4 drinks/day onward, and any duration), the odds ratio reached 5.

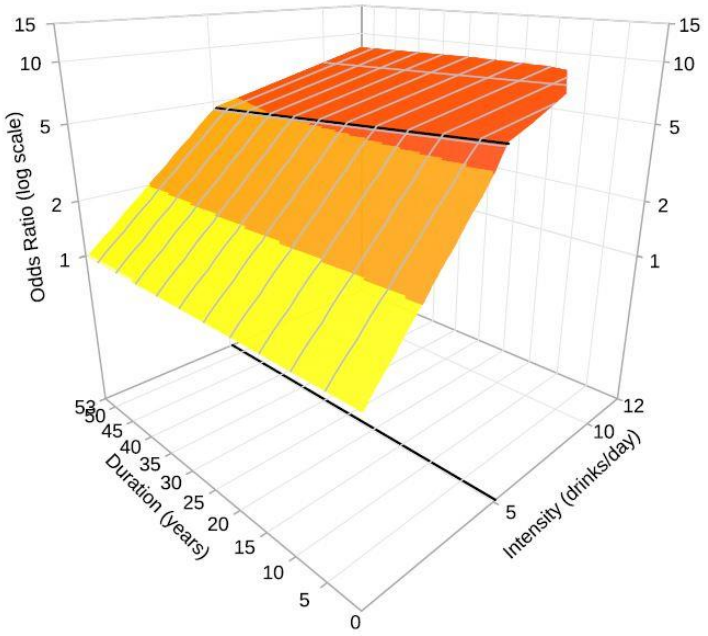
Conclusions

The proposed methodology aims to solve the problem of estimating the number and positions of knots in semiparametric regression models with linear splines. A well-known variable selection technique has been adapted in order to estimate the presence or absence of knots in possible overparameterised models. Once the number of knots is selected, the appropriate model can be fitted with the preferred technique. Moreover, the method gives us a first guess on the knot locations through the inspection of the marginal posterior distributions of the knots location parameters. This can be useful in the initialisation step of algorithms with difficulties in exploring entirely the parameter space, especially in high dimensional problems.
In terms of computational complexity, this methodology requires a higher number of parameters to be estimated, as compared to one model as specified in the *NVS* approach. Consequently, it is more time consuming. On the other hand, this approach requires the estimate of only one model to select the number

on knots, while the *NVS* approach needs estimating a possibly large number of models, especially in the bivariate case, and, moreover, is based on information criteria that have been criticised in the statistical literature.
Lastly, in order to compute the WAIC or LOO criteria used in the *NVS* approach we need to include additional steps in the simulations. This has two main consequences: a negligible increment in the sampling in terms of time, but a remarkable increase of the memory needed to store the simulations.
The methodology is designed for situations in which the number of expected slope changes is limited. This is reasonable in many epidemiological studies. When we expect an high number of knots, this methodology may be not appropriate, but despite this, the knots corresponding to the most evident slope changes are correctly identified in a simulation study, thus supporting the application of this methodology.
The methodology was also tested on the semiparametric logistic model with bivariate spline on real data. In this case, the model structure is much more complex as compared to the initial model presentation. However, changes in the slope of the surface are estimated together with the other parameters and this represents the main innovative result of our methodology.

Figure 1. *Perspective and contour plots representing the risk of oral cancer as related to alcohol drinking intensity and duration. The surface is estimated through the overparameterised SSVSξ model with two knots on intensity and two knots on duration. On the grid, black thicker lines represent knot locations. Dark grey lines in contour plots (upper panel) indicate iso-risk curves at defined levels of risk.*





Full list of authors from the INHANCE consortium:

Gioia Di Credico[1], Jerry Polesel[2], Luigino Dal Maso[2], Francesco Pauli[1], Nicola Torelli[1], Diego Serraino[2], Eva Negri[3], Daniele Luce[4], Isabelle Stucker[5], Keitaro Matsuo[6], Paul Brennan[7], Marta Vilensky[8], Leticia Fernandez[9], Maria Paula Curado[10], Ana Menezes[11], Alexander W. Daudt[12], Rosalina Koifman[13], Victor Wunsch-Filho[14], Ivana Holcatova[15], Wolfgang Ahrens[16,17], Pagona Lagiou[18], Lorenzo Simonato[19], Lorenzo Richiardi[20], Claire Healy[21], Kristina Kjaerheim[22], David I. Conway[23], Tatiana V. Macfarlane[24], Peter Thomson[25], Antonio Agudo[26], Ariana Znaori[7], Leonardo F. Boaventura Rios[27], Tatiana N. Toporcov[14], Silvia Franceschi[2], Rolando Herrero[7], Joshua Muscat[28], Andrew F. Olshan[29], Jose P. Zevallos[30], Carlo La Vecchia[31], Deborah M. Winn[32], Erich M. Sturgis[33], Guojun Li[33], Eleonora Fabianova[34], Jolanda Lissowska[35], Dana Mates[36], Peter Rudnai[37], Oxana Shangina[38], Beata Swiatkowska[39], Kirsten Moysich[40], Zuo-Feng Zhang[41], Hal Morgenstern[42], Fabio Levi[43], Elaine Smith[44], Philip Lazarus[45], Cristina Bosetti[46], Werner Garavello[47], Karl Kelsey[48], Michael McClean[49], Heribert Ramroth[50], Chu Chen[51], Stephen M. Schwartz[51], Thomas L. Vaughan[51], Tongzhang Zheng[52], Gwenn Menvielle[53], Stefania Boccia[54,55], Gabriella Cadoni[56,57], Richard B. Hayes[58], Mark Purdue[32], Maura Gillison[59], Stimson Schantz[60], Guo-Pei Yu[61], Hermann Brenner[62,63,64], Gypsyamber D'Souza[65], Neil D. Gross[66], Shu-Chun Chuang[67], Paolo Boffetta[68], Mia Hashibe[69], Yuan-Chin Amy Lee[69], Valeria Edefonti[31]

[1]Department of Economics, Business, Mathematics and Statistics, University of Trieste, Trieste, Italy; [2]Cancer Epidemiology Unit, Centro di Riferimento Oncologico di Aviano (CRO) IRCCS, Aviano, Italy; [5]Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milano, Italy; [4]Université Rennes, INSERM, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail) - UMR_S 1085, Pointe-à-Pitre, France; [5]Inserm, Center for Research in Epidemiology and Population Health (CESP), Cancer and Environment team, Université Paris-Sud, Université Paris-Saclay, Villejuif, France; [6]Aichi Cancer Center Research Institute, Nagoya, Japan; [7]International Agency for Research on Cancer, Lyon, France; [8]Institute of Oncology Angel H. Roffo, University of Buenos Aires, Argentina; [9]Institute of Oncology and Radiobiology, Havana, Cuba; [10]Epidemiology - CIPE/ACCAMARGO, Sao Paulo, Brazil; [11]Universidade Federal de Pelotas, Pelotas, Brazil; [12]Hospital Moinhos de Vento, Porto Alegre, Brazil; [13]Escola

Nacional de Saude Publica, Fundacao Oswaldo Cruz, Rio de Janeiro, Brazil; [14]University of São Paulo School of Public Health, São Paulo, Brazil; [15]Institute of Hygiene & Epidemiology, 1st Faculty of Medicine, Charles University, Prague, Czech Republic; [16]Leibniz Institute for Prevention Research and Epidemiology, BIPS, Bremen, Germany; [17]University of Bremen, Faculty of Mathematics and Computer Science, Bremen, Germany; [18]Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece; [19]Department of Cardiovascular and Thoracic Sciences and Public Health, University of Padova, Padova, Italy; [20]Department of Medical Sciences, University of Turin, Turin, Italy; [21]Trinity College School of Dental Science, Dublin, Ireland; [22]Cancer Registry of Norway, Oslo, Norway; [23]School of Medicine, Dentistry, and Nursing, University of Glasgow, Glasgow, UK; [24]School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK and School of Medicine, University of Dundee, Dundee, UK; [25]University of Hong Kong, Hong Kong, China; [26]Institut Catala Oncologia, Barcelona, Spain; [27]University of São Paulo School of Medicine, São Paulo, Brazil; [28]Penn State College of Medicine, Hershey, PA, USA; [29]University of North Carolina School of Public Health, Chapel Hill, NC, USA; [30]Division of Head and Neck Surgical Oncology in the Department of Otolaryngology/Head and Neck Surgery at Washington University School of Medicine, USA; [31]Branch of Medical Statistics, Biometry and Epidemiology "G. A. Maccacaro", Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milano, Italy; [32]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA; [33]UT- M.D. Anderson Cancer Center, Houston, Texas, USA; [34]Regional Authority of Public Health in Banska Bystrica, Slovakia; [35]The M. Skasodowska-Curie Memorial Cancer Center and Institute of Oncology, Dept. of Cancer Epidemiology and Prevention, Warsaw, Poland; [36]National Institute of Public Health, Bucharest, Romania; [37]National Institute of Environmental Health to National Public Health Institute, Budapest, Hungary; [38]Cancer Research Centre, Moscow, Russia; [39]Institute of Occupational Medicine, Lodz, Poland; [40]Roswell Park Cancer Institute, Buffalo, NY, USA; [41]UCLA School of Public Health, Los Angeles, CA, USA; [42]Departments of Epidemiology and Environmental Health Sciences, School of Public Health and Department of Urology, Medical School, University of Michigan, Ann Arbor, MI, USA; [43]Institut Universitaire de Medecine Sociale et Preventive (IUMSP), Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland; [44]College of Public Health, University of Iowa, Iowa City, IA, USA; [45]Washington State University College of Pharmacy and Pharmaceutical Sciences, Spokane, WA, USA; [46]Department of Oncology, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy; [47]Department of Otorhinolaryngology, School of Medicine and Surgery, University of Milano - Bicocca, Monza, Italy; [48]Brown University, Providence, RI, USA; [49]Boston University School of Public Health, Boston, MA, USA; [50]University of Heidelberg, Germany; [51]Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [52]Department of Epidemiology, Brown University School of Public Health, Providence, RI, USA; [53]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique IPLESP, Department of Social Epidemiology, F75012, Paris, France; [54]Department of Woman and Child Health and Public Health - Public Health Area, Fondazione Policlinico Universitario A. Gemelli IRCCS, Roma, Italia; [55]Sezione di Igiene, Istituto di Sanità Pubblica, Università Cattolica del Sacro Cuore, Roma, Italia; [56]Dipartimento di Scienze dell'Invecchiamento, Neurologiche, Ortopediche e della Testa-Collo, Fondazione Policlinico Universitario A. Gemelli IRCCS, Roma, Italia; [57]Istituto di Clinica Otorinolaringoiatrica, Università Cattolica del Sacro Cuore, Roma, Italia; [58]Division of Epidemiology, New York University School Of Medicine, NY, New York, USA; [59]"Thoracic/Head and Neck Medical Oncology", The University of Texas MD Anderson Cancer Center, TX, USA; [60]New York Eye and Ear Infirmary, New York, NY, USA; [61]Medical Informatics Center, Peking University, China; [62]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; [63]Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany; [64]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany; [65]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; [66]Department of Head and Neck Surgery, Division of Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [67]Institute of Population Health Sciences, National Health Research Institutes, Taiwan; [68]The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA; [69]Division of Public Health, Department of Family & Preventive Medicine, University of Utah School of Medicine and Huntsman Cancer Institute, Salt Lake City, UT, USA.

References

[1] Ruppert D., Wand M. P., and Carroll R. J., Semiparametric regression. Cambridge University Press, 2003.

[2] Wood S. N., Generalized additive models: an introduction with R. Chapman and Hall, 2006.

[3] O'Hara R. B., Sillanpää M. J., A review of Bayesian variable selection methods: what, how and which. Bayesian Analysis, 2009; 4(1), 85–117.

[4] Winn D. M., Lee Y. C., Hashibe M., and the INHANCE consortium, The INHANCE consortium: toward a better understanding of the causes and mechanisms of head and neck cancer. Oral Dis., 2015 Sep; 21(6):685-93.

[5] Stan Development Team, Stan Modeling Language Users Guide and Reference Manual, Version 2.17.0. 2017; http://mc-stan.org.

[6] R Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2018.

RELATIONSHIP AMONG HOSPITAL VOLUME AND MORTALITY IN PATIENTS AFFECTED BY LUNG CANCER: EVIDENCE FROM AN ITALIAN REAL-WORLD INVESTIGATION

Ieva Francesca[1,2,3], Rea Federico[1,4], Barni Sandro[5], Merlino Luca[6], Apolone Giovanni[7], Corrao Giovanni[1,4]

[1] *National Centre for Healthcare Research & Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy*

[2] *MOX – Modeling and Scientific Computing lab, Department of Mathematics, Politecnico di Milano, Milan, Italy*

[3] *CADS – Center for Analysis Decisions and Society, Human Technopole, Milano, Italy*

[4] *Laboratory of Healthcare Research & Pharmacoepidemiology, Unit of Biostatistics, Epidemiology and Public Health, Department of*

Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

[5] *U.O. Oncologia Medica, ASST Bergamo Ovest, Ospedale di Treviglio, Italy*

[6] *Epidemiologic Observatory, Lombardy Regional Health Service, Milan, Italy.*

[7] *Fondazione IRCCS - Istituto Nazionale dei Tumori, Milan, Italy.*

Introduction

Lung cancer is one of the most common malignancies and the leading cause of cancer death, affecting an estimated 2.1 million people and contributing to 1.8 million deaths globally [1], with a 5-year overall survival from diagnosis of 19% [2]. Among the treatment options, surgical therapy aims to decrease morbidity and mortality [3], and almost 70% and 10% of patients with stage I/II and III/IV undergo surgery, respectively [4]. Although it has been postulated that patients might benefit from the centralization in high-volume specialized centres [5,6], conflicting results have been reported on the relationship between hospital volume and outcomes among patients with lung cancer.

Objectives

The aim of the present work is to generate new evidence for this ongoing debate, conducting a population-based investigation on patients hospitalized for surgery of lung cancer.

Methods

The data used for the present study were retrieved from the healthcare utilization databases of Lombardy. For each patient, we linked the hospitalization diagnosis and procedures dataset (coded according to the ICD-9-CM classification system), the outpatient drug prescriptions one (coded according to the ATC classification system) and the outpatient visits (including visits in specialist ambulatories and diagnostic laboratories) via a single identification code. In order to preserve privacy, each identification code was automatically converted into an anonymous code. The inverse process was prevented by deletion of the conversion table.
The target population consisted of Lombardy residents, aged 18 years or older, who were hospitalized for surgery of lung cancer during 2012-2014. The first hospital admission during this period was defined as index hospitalization. Cohort member who were beneficiaries of the National Health Service from less than 3 years before the index hospitalization were excluded in order to allow the patient's characterization through their previous contacts with the National Health Service. Exclusion was extended to other three categories of patients: (i) who had a diagnosis of lung cancer between the third year before and the year previous the index hospitalization, (ii) who had a diagnosis of other cancer within 3 years from the index hospitalization, and (iii) who had malignant neoplasm within 3 years from the index hospitalization. The remaining patients were included into the final cohort whose members accumulated person-years of follow-up from the surgery

date of the index hospitalization until death (outcome) or censoring (i.e., emigration or 30 June 2018), whichever came first.
For each hospital we recovered the annual number of lung cancer surgeries from the database of hospital discharge records over the period of patients' enrolment (2012-2014). With the aim to identify the 10% most specialized centres, high volume was defined according to the 90th percentile of the annual surgeries (i.e., >95 procedures per year), accounting for almost 40% of cohort members. Conversely, low volume was defined according to the median of this distribution (i.e., ≤10 procedures per year), accounting for about 5% of cohort members. Thus, three categories of hospital volume were identified: low (≤10 procedures per year), intermediate (11-95) and high (>95). Baseline characteristics included gender, age, resection volume and use of neoadjuvant chemotherapy within six months before the index hospitalization. Furthermore, the clinical status was assessed by the so-called Multisource Comorbidity Score (MCS) [7]. A multilevel approach was adopted with the subjects as the lower level and the hospital as the higher one [8]. Multilevel logistic regression was fitted to assess the variance of the random effect and to evaluate the impact of the hospital, as a grouping factor, on the 3-year survival [9]. According to the estimates of regression parameters and random effect variance, survival probabilities were calculated in three scenarios, setting the hospital effect to: (i) – two-fold the standard deviation of the random effect (worst scenario); (ii) 0 (standard); (iii) + two-fold the standard deviation of the random effect (best scenario). Cox frailty model was fitted to estimate the hazard ratio, and its 95% confidence interval, of all-cause death in relation to hospital volume, using the lowest category (≤10 procedures per year) as reference. Adjustments were made for the aforementioned baseline covariates.

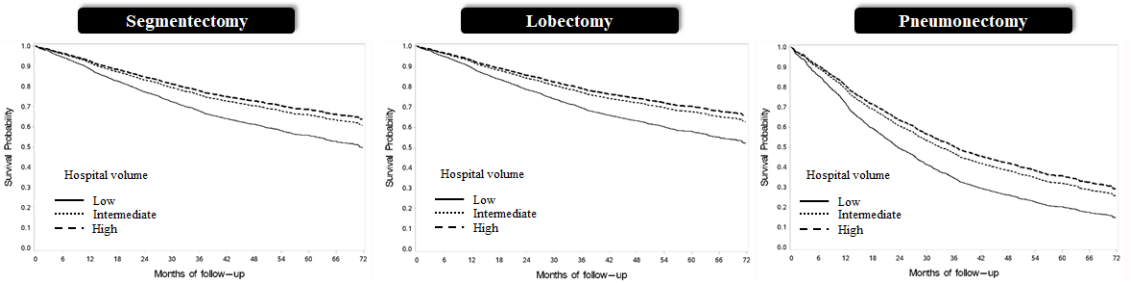
Results

The proportion of male patients and age showed a reduction as hospital volume increased. Moreover, the segmentectomy was more performed at hospitals with low volume whereas the lobectomy among those with intermediate and high volume.
Taking into account the patient's characteristics, the survival probability was strongly affected by the hospital. An average patient (male gender, 65 years, MCS<3, no neoadjuvant chemotherapy) who underwent to segmentectomy at a hospital in the lower tail of the random effect distribution (worst scenario) showed a similar survival than a patient who underwent to pneumonectomy at a hospital in the upper tail (best scenario) (0.64 vs 0.59, respectively). Patients undergone pneumonectomy showed the lowest survival probabilities, ranging between 0.30 (worst scenario) to 0.59 (best scenario).
Compared to patients discharged from a low-volume hospital, the risk exhibited a marked progressive reduction as hospital volume increased to the intermediate (-29%) and high categories (-40%).
Figure 1 shows the survival curves of an average patient (male gender, 65 years, MCS<3, no neoadjuvant chemotherapy) according to hospital and resection volume based on the Cox model. The figure highlighted the worse outcome of patients underwent pneumonectomy and who were discharged from a low-volume hospital.

Conclusions

Our real-world investigation provides two main findings. First, among patients underwent lung cancer surgery, the hospital has a noteworthy impact on the mortality, since the 3-year survival probabilities in the worst and in the best scenario differed for almost 30 percentage points. Second, the procedural volume appears to be a hospital parameter that strongly affects the subject's survival. Indeed, patients discharged from a hospital with high volume exhibited a 40% reduction in the risk of death with respect to those underwent surgery at a facility with low volume.

Figure 1. Kaplan-Meier survival curves for all-cause death according to the type of surgery and the hospital volume



Bibliografia

[1] https://www.who.int/news-room/fact-sheets/detail/cancer [Accessed on 7 May 2019]

[2] Howlader N, Noone AM, Krapcho M, et al. SEER Cancer Statistics Review (CSR), 1975-2014. Bethesda: National Cancer Institute, 2017

[3] Lackey A, Donington JS. Surgical management of lung cancer. Semin Intervent Radiol. 2013;30:133-40

[4] Miller KD, Siegel RL, Lin CC et al. A. Cancer treatment and survivorship statistics, 2016. CA Cancer J Clin. 2016;66:271-89

[5] Epstein AM. Volume and outcome--it is time to move ahead. N Engl J Med. 2002;346:1161-4

[6] Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Intern Med. 2002;137:511-20

[7] Corrao G, Rea F, Di Martino M et al. Developing and validating a novel multisource comorbidity score from administrative data: a large population-based cohort study from Italy. BMJ Open. 2017;7:e019503

[8] Rice N, Leyland A. Multilevel models: applications to health data. J Health Serv Res Policy. 1996;1:154-64

[9] Grieco, N., Ieva, F., Paganoni, A.M. Performance assessment using mixed effects models: a case study on coronary patient care. IMA Journal of Management Mathematics. 2012;23:117-131

EVALUATING THE RELATIONSHIP BETWEEN SOCIOECONOMIC INDICATORS AND PREMATURE MORTALITY IN ITALIAN MUNICIPALITIES: A COMPARISON OF STANDARD AND NEW STATISTICAL APPROACHES

Alicandro Gianfranco[1], Bertuccio Paola[2], Carlo La Vecchia[3], Frova Luisa[1]

[1] *Direzione centrale per le statistiche sociali e il censimento della popolazione, Istituto Nazionale di Statistica (Istat), Roma, Italia*

[2] *Dipartimento di Scienze Biomediche e Cliniche L. Sacco, Università degli Studi di Milano, Milano, Italia*

[3] *Dipartimento di Scienze Cliniche e di Comunità, Università degli Studi di Milano, Milano, Italia*

Introduction

Important geographic differences in mortality have been reported in Italy, and the uneven distribution of socioeconomic indicators across different areas may explain part of this variability [1]. However, evaluating the relationship between territorial socioeconomic indicators and mortality is challenging, due to the wide range of indicators available and to the high degree of correlation among them. In this context, standard statistical approaches are of limited use in identifying the territorial indicators more related to mortality.

Aims

In this work, we explored standard and new statistical approaches to evaluate the relationship between a set of selected territorial socioeconomic indicators and premature mortality in Italian municipalities.

Methods

We obtained the age-standardized mortality rates (ASMRs) per 10,000 person-years, at age 0-74 years, for Italian municipalities from a record-linkage cohort study, based on the 2011 Italian census, and the mortality archives (2012-2016). To have reliable estimates of the ASMR, we selected the municipalities with at least 5000 inhabitants registered in the 2011 census. Socioeconomic indicators were retrieved from the *8milaCensus* database, a national archive based on the information collected in the 2011 census that provides a set of 99 indicators measured at municipality level and grouped in eight different dimensions (population, foreigners' integration, families, housing conditions, education and training, labour market, mobility, social and material vulnerability). We selected 12 indicators, including foreigners' unemployment (INT1), young single-parent families (FAM1), non-owned houses (HOU1), illiteracy (EDU1), early exit from education and training (EDU2), people with less than high school diploma (EDU3), young people not studying and not working (OCC1), unemployment (OCC2), people working in low-qualified activities (OCC3), people living in economic disadvantage (VUL1), people living in overcrowded houses (VUL2), and young people out of the labour market, education and training (VUL3). Socioeconomic indicators are expressed as percentage of resident population or families living in the municipality. When the distribution of the indicator was skewed, we applied a natural logarithm or square root transformation and we standardized all the transformed indicators before analysis.
We first fitted simple and mutually adjusted linear regression models including the natural logarithm of the ASMR as dependent variable and the socioeconomic indicators as predictors. Beta coefficients were estimated for an increment equal to one standard deviation of the predictor. Since most indicators are highly correlated the estimates obtained through this standard approach can be biased due to multicollinearity. To overcome this issue, we applied a classification approach. We then explored the correlation structure among the indicators and applied a principal component analysis (PCA) to obtain a smaller number of unobservable and uncorrelated components while maintaining as much information as possible of the original variables

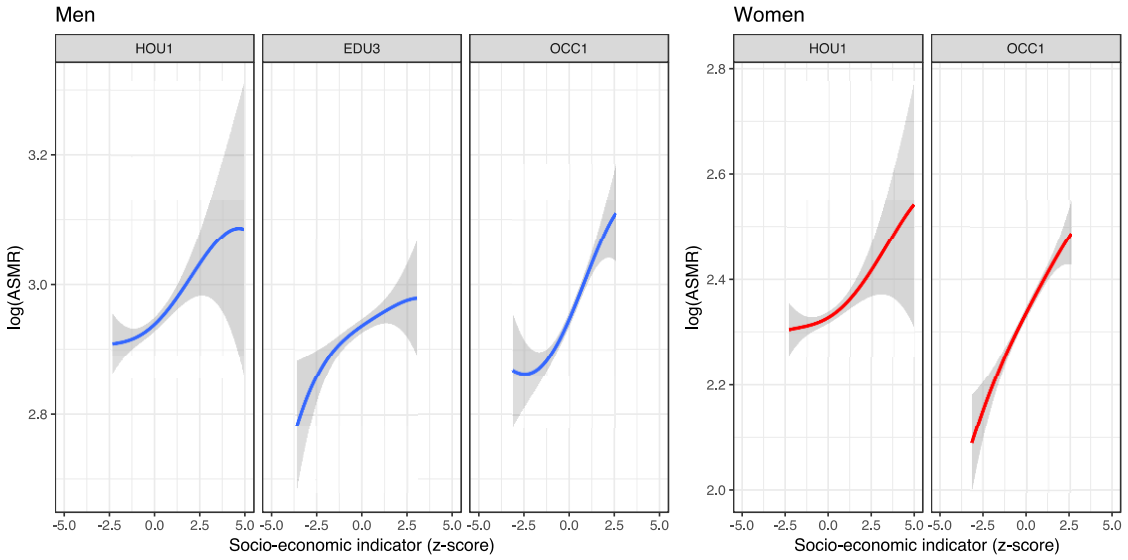
[2]. We chose the number of components on the basis of two criteria: factor eigenvalue >1 and the interpretability. We run a varimax rotation to provide an easier interpretability, and we obtained a set of rotated factor loadings. These loadings indicate the variance in the original indicator explained by each component, and the indicators with the highest loadings (≥0.63) on a given component were used to interpret and name the extracted components. For each component score, municipalities were grouped into ten categories according to deciles of the component. We then fitted a linear regression model including the deciles of each component score as covariates, using the category below the 10th percentile as reference. However, this approach does not allow to evaluate the individual contribution of each indicator to the overall effect of the component.
Finally, we tried to apply to our data the Bayesian Kernel machine regression (BKMR), a recently proposed non-parametric approach originally developed to estimate the health effect of exposures to multi-pollutant and chemical mixtures [3]. This technique uses a kernel function to estimate the multivariable exposure-response function in a flexible way that allows to model nonlinear and non-additive relationships. In addition, the technique can incorporate a hierarchical variable selection that addresses the multicollinearity by simultaneously conducting variable selection on correlated groups of exposures as well as on the individual exposures within each group. We fitted a BKMR with hierarchical variable selection with groups based on PCA results. For each municipality i =1,…, n, the BKMR is given by $\log(ASMR_i) = h(INT1_i, FAM1_i, HOU1_i, EDU1_i, EDU2i, EDU3_i, OCC1_i, OCC2_i, OCC3_i, VUL1_i, VUL2_i, VUL3_i) + e_i$, where h() denotes the exposure-response Gaussian Kernel function to be estimated and the residual e_i ,~ $N(0,\sigma^2)$ are assumed to be independent and identically distributed with a common variance. The model was fitted using Markov chain Monte Carlo.
We carried out the BKMR using the R package “bkmr”.

Results

We considered 2390 out of 8092 Italian municipalities with a minimum number of 5000 inhabitants covering 82.6% of the total Italian population. Over the observation period, 596.144 deaths were registered (367,871 men and 228,273 women). A high variability emerged in premature mortality across municipalities with ASMRs ranging between 8.3 and 33.5 per 10,000 person-years in men and 3.9 and 22.5 in women.
Median values (range) of the socioeconomic indicators were*:* 43% (13.7-83) for INT1, 0.9% (0.1-3.7) for FAM1, 24.9% (10.8-56.8) for HOU1, 0.7% (0.1-8.0) for EDU1, 14.3% (2.1-45.3) for EDU2, 48% (16.5-74.7) for EDU3, 18.1% (5.9-51) for OCC1, 8.3% (1.3-37.2) for OCC2, 16% (5.9-60.1) for OCC3, 1.4% (0.3-17.9) for VUL1, 1% (0.7-10.8) for VUL2 and 9.7% (7.5-33.7) for VUL3.
High positive correlations were observed among most indicators of education, labour market and social and material vulnerability (Spearman’s r coefficients ranging from 0.7 to 0.9), while INT1, FAM1, HOU1 and VUL2 showed lower correlations.
In simple linear regression models, each indicator showed a significant positive association with premature mortality in both sexes, with beta coefficients ranging from about 0.04 to 0.08. When we applied mutually adjusted linear regression models, the associations became weaker and mostly not significant, whereas a few indicators changed the direction of the association, likely due to multicollinearity. Specifically, OCC1 remained positively associated (beta= 0.06 for men and 0.08 for women). FAM1, HOU1 and EDU3 (only in men) showed weaker associations, with beta coefficients around 0.01-0.03 among men and 0.01-0.02 among women.
Through the PCA, we identified three main components accounting for 78% of the total variance of the original variables. The first component had the highest factor loadings on INT1, OCC1, OCC2, VUL1 and VUL3, the second component on EDU1, EDU2, EDU3 and OCC3, while the third component on FAM1, HOU1 and VUL2. We labelled the three components: “Unemployment and economic disadvantage”, “Low education and occupation in low-qualified activities” and “Households and living conditions”, respectively.

In the multivariable linear regression model, all the three component scores were positively associated with premature mortality. The beta coefficients for the category in the highest decile compared to the lowest one were: 0.17, 0.14 and 0.16 in men and 0.18, 0.18 and 0.11 in women, for the 1st 2nd and 3rd component, respectively.
BKMR selected three indicators associated with the overall mortality in men (HOU1, EDU3 and OCC1) and two in women (HOU1 and OCC1). The univariate relationships between selected socioeconomic indicators and premature mortality are shown in the Figure 1. In men, the estimated difference in log(ASMR) between the 10th and 90th percentile of each selected indicators, where all the other indicators are fixed at their median value was: 0.07 (95% CI: 0.04-0.09) for HOU1, 0.05 (95% CI: 0.03-0.08) for EDU3 and 0.16 (95% CI: 0.13-0.19) for OCC1. Corresponding figures in women were 0.06 (95% CI: 0.03-0.08) for HOU1 and 0.17 (95% CI: 0.14-0.19) for OCC1. An additive effect emerged between HOU1 and OCC1 (in both sexes), while an interaction effect was observed between OCC1 and EDU3 in men.

***Figure 1**. Univariate exposure-response function (95% credible intervals) between selected socioeconomic indicators and the age-standardized mortality rate (0-74 years) in Italian municipalities while fixing other indicators at median values. The function was estimated by Bayesian Kernel machine regression including the following indicators: foreigners' unemployment (INT1), young single-parent families (FAM1), non-owned houses (HOU1), illiteracy (EDU1), early exit from education and training (EDU2), people with less than high school diploma (EDU3), young people not studying and not working (OCC1), unemployment (OCC2), people working in low-qualified activities (OCC3), people living in economic disadvantage (VUL1), people living in overcrowded houses (VUL2), and young people out of the labour market, education and training (VUL3).*



Conclusions

The results of the different approaches are generally consistent but the hierarchical variable selection of the BKMR overcomes multicollinearity that arises in multivariable linear regression. In addition, it provides a set of summary output that describe the overall effect of socioeconomic indicators, the individual contribution of each indicator as well as non-linear, additive and interactive effects. All these results can be useful in understanding the relationship between territorial socioeconomic indicators and geographic inequalities in premature mortality. This is the first time that BKMR is used in this context and further work is needed to explore all potential utility of this technique.

References

[1] Petrelli A., Di Napoli A., Sebastiani G., Rossi A. et al. Italian Atlas of mortality inequalities by education level. Epidemiol Prev., 2019;43(1S1):1-120.

[2] Johnson R., Wichern D. Applied mutlivariate statistical analysis. Upper Saddler River (NJ): Prentice Hall; 2002.

[3] Bobb J.F., Valeri L., Claus Henn B. et al. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. Biostatistics, 2015;16(3):493-508.

THE HARTUNG, KNAPP, SIDIK JONKMAN ADJUSTMENT FOR CONFIDENCE INTERVALS IN PRESENCE OF HIGH ETEROGENITY: AN APPLICATION TO A META-ANALYSIS ON THE ASSOCIATION BETWEEN RENIN-ANGIOTENSIN SYSTEM INHIBITORS AND RISK OF ALZHEIMER'S DISEASE.

Scotti Lorenza[1], Soranna Davide[2], Bilo Grzegorz[2,3], Parati Gianfranco[2,3], Zambon Antonella[1]

[1] *Dipartimento di Statistica e Metodi Quantitativi, Università degli studi di Milano-Bicocca, Milano*

[2] *Istituto Auxologico Italiano, IRCCS, Milano*

[3] *Dipartimento di Medicina e Chirurgia, Università degli studi di Milano-Bicocca, Milano*

Introduction

In presence of heterogeneity between studies, the DerSimonian and Laird (DL) approach is commonly used in meta-analyses, to obtain pooled estimates accounting for between study variability [1]. The corresponding confidence intervals are usually calculated assuming the asymptotic normal distribution of the pooled estimate (or its logarithmic transformation in case of association measures), however, when the number of included studies is small, the sample size is largely variable between studies and/or the extent of between studies variability is expected to be moderate or substantial, this method may lead to over precise confidence intervals [2]. The consequence is the identification of too many statistically significant results [2] and therefore, methods to obtain suitable confidence intervals under these conditions are needed. Hartung, Knapp, Sidik and Jonkman proposed a simple method (HKSJ method) to provide adjusted confidence intervals in this situation [3].

Aim

To compare the results obtained by the two approaches used to estimate confidence intervals obtained under the normal distribution assumption and using the HKJS method applied to a meta-analysis aimed to evaluate the association between use of renin-angiotensin system inhibitors (RAS inhibitors) and risk of Alzheimer's disease (AD).

Materials and methods

A systematic MEDLINE search of observational studies investigating the association between use of ARBs and risk of AD up to 31st March 2019 was performed. The following keywords were included both as text and Mesh terms in the search strategy: "Alzheimer's disease" AND "Antihypertensive" OR "Angiotensin-converting Enzyme Inhibitors" OR "Angiotensin II type 1 receptor blockers", as well as the single active ingredients belonging these two antihypertensive drug classes.
The studies were included if they: (i) were original observational studies on humans; (ii) were written in English; (iii) considered as exposure angiotensin-converting enzyme Inhibitors (ACE- inhibitors) or angiotensin-receptor blockers (ARB) and, as reference, users of any other antihypertensive drug; (iv) included incident AD cases defined according to standardized criteria or diagnosed by qualified physicians and (v) reported crude or adjusted estimates of the association estimates [hazard ratios, relative risks or odds ratios) and the corresponding variability measures (95% confidence intervals, standard error) or enough data for their calculation.
Several information were extracted from the studies, including: first author, year of publication, study design, country, sample size, mean sample age, gender, association estimates and the corresponding variability measure. When the studies reported multiple estimates with different levels of adjustment, only the most adjusted effect estimates were extracted.

Statistical analysis. Fixed effects and, in case of between study heterogeneity, random effects pooled relative risks (pooled RRs) were calculated separately for the association between the use of each RAS inhibitors drug class (ACE inhibitors and ARB) and risk of AD compared to use of any other antihypertensive drugs. Between-study heterogeneity was assessed using the I^2 index (the percentage between study variability due to heterogeneity). The presence of heterogeneity was defined as values of I^2 index greater than 50% [4]. When random effect estimates are reported, classical and HKJS-adjusted confidence intervals were calculated.

Calculation of confidence intervals.

For each pooled estimate, 95% confidence intervals (95%CI) were calculated according to the normal distribution assumption and to the HKSJ method.

According to classical method, 95%CI for the pooled estimates were calculated as:

$$95\%CI: e^{(\ln(pooled\ RR) \pm z_{1-\alpha/2} * se)}$$

Where pooled RR is the usual DL random pooled estimate and se the corresponding standard error.

In order to determine the adjusted interval according to the HKJS method, the following quantity should be calculated:

$$q = \frac{1}{k-1} \sum_{i=1}^{k} w_i \left(y_i - pooledRR_{fe}\right)^2$$

Where k is the number of included studies, $pooledRR_{fe}$ is the pooled estimate calculate using the inverse variance method (fixed effect pooled estimate), w_i are the study specific weights (inverse of variance) and y_i represents the logarithm of the association estimate of the i-th study.
Once q is computed, the confidence interval for the random effect pooled estimate adjusted using the HKJS method is calculated as:

$$HKJS - 95\%CI: e^{(\ln(pooled\ RR) \pm t_{(k-1),(1-\alpha/2)} * \sqrt{q} * se)}$$

where $t_{(k-1);(1-\alpha/2)}$ is the $(1 - \alpha/2)$-quantile of the Student-t distribution with $(k - 1)$ degrees of freedom [5].
To compare the two methods, it was evaluated whether both confidence intervals included the value assumed under the null hypothesis of absence of association.
All analyses were performed using R software version 3.6.0.

Results

The literature search allowed the identification of 587 studies. After the application of the inclusion criteria, 5 studies were included in the meta-analysis [6-10].
Regarding the study design, all included studies were cohort studies. Two studies were performed in the US, 2 in Asia and 1 in Europe. The average age of studies' participant varies from 65 [10] to 82 years old [8] and the proportion of males varied from 33% [8] to 98% [6]. All studies considered both ACE-inhibitors and ARBs as exposure.
The table below shows the pooled RR for the association between ACE-inhibitors or ARB and AD, the corresponding confidence intervals calculated with both methods under study, the range of the sample size of the included studies, the number of estimates pooled for each comparison and the I^2 index.

Exposure	Sample size range	N	I^2	Pooled RR	95%CI	HKJS-95%CI
ACE-inhibitors	781-105,187	5	87%	0.93	(0.80-1.09)	(0.76-1.15)
ARBs		5	88%	0.71	(0.58-0.95)	(0.49-1.04)

As shown in table, the sample size of the included studies is largely variable ranging from 781 [7] to 105,187 [6] subjects. Moreover, heterogeneity is substantial, in fact the I^2 indexes are equal to 87% and 88% respectively for ACE-inhibitors ARB. Compared to users of other antihypertensive drugs, ACE-inhibitors' users have a non-significant 7% reduced AD. As expected, the confidence intervals calculated using the HKJS adjustment are slightly wider than conventional confidence intervals but the results in terms of statistical significance are comparable. The pooled RR for the association between use of ARBs and AD is 0.71, the confidence interval based on the normal distribution ranges from 0.58 to 0.95, and that obtained after HKJS's adjustment from 0.49 to 1.04 making the association not statistically significant.

Conclusions

In meta-analyses pooling few studies with variable sample size and affected by high heterogeneity, the use of methods to adjust confidence intervals may lead to different results in terms of statistical significance compared to the methods commonly used for their calculation. The results of this meta-analysis showed a non-statistically significant reduced risk of AD regardless the application of any correction for the calculation of confidence intervals. Regarding the association between ARBs and AD, a non-statistically significant reduce risk was observed when the confidence intervals are calculated using the HKJS adjustment. Although recent studies showed that ARBs may have neuroprotective properties through several biological mechanisms such as decreasing oxidative stress in cerebral microvessels and antinflammatory effects in the central nervous system [11,12]; however, the protective effect of ARBs is not confirmed by large randomized clinical trials on this issue [13-15] suggesting that the effect observed in observational studies might be due to flaws in the study conduction. In conclusion, in meta-analyses based on few studies and in presence of high heterogeneity, suitable methods should be applied in the confidence interval calculation to reduce the number of potential false positive associations.

References

[1] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7: 177–188

[2] IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol. 2014 18; 14:25

[3] Sidik K, Jonkman JN. A simple confidence interval for meta-analysis. Stat Med. 2002; 21:3153-9

[4] Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003; 327:557-60

[5] Röver C, Knapp G, Friede T. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. BMC Med Res Methodol. 2015; 15:99.

[6] Li NC, Lee A, Whitmer RA, Kivipelto M, et al. Use of angiotensin receptor blockers and risk of dementia in a predominantly male population: prospective cohort analysis. BMJ. 2010 Jan 12; 340:b5465.

[7] Yasar S, Xia J, Yao W, et al. Antihypertensive drugs decrease risk of Alzheimer disease: Ginkgo Evaluation of Memory Study. Neurology. 2013;81:896-903.

[8] Davies NM, Kehoe PG, Ben-Shlomo Y, et al. Associations of anti-hypertensive treatments with Alzheimer's disease, vascular dementia, and other dementias. J Alzheimers Dis. 2011;26:699-708.

[9] Hwang D, Kim S, Choi H, et al. Calcium-Channel Blockers and Dementia Risk in Older Adults - National Health Insurance Service - Senior Cohort (2002-2013). Circ J. 2016;80:2336-2342.

[10] Kuan YC, Huang KW, Yen DJ, et al. Angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers reduced dementia risk in patients with diabetes mellitus and hypertension. Int J Cardiol. 2016;220:462-6.

[11] Kehoe PG, Miners S, Love S. Angiotensins in Alzheimer's disease - friend or foe? Trends Neurosci. 2009;32:619-28.

[12] Mogi M, Li JM, Tsukuda K, et al. Telmisartan prevented cognitive decline partly due to PPAR-gamma activation. Biochem Biophys Res Commun. 2008;375:446-9

[13] Lithell H, Hansson L, Skoog I, et al. The Study on Cognition and Prognosis in the Elderly (SCOPE): principal results of a randomized double-blind intervention trial. J Hypertens. 2003;21:875-86.

[14] Diener HC, Sacco RL, Yusuf S, et al. Effects of aspirin plus extended-release dipyridamole versus clopidogrel and telmisartan on disability and cognitive function after recurrent stroke in patients with ischaemic stroke in the Prevention Regimen for Effectively Avoiding Second Strokes (PRoFESS) trial: a double-blind, active and placebo-controlled study. Lancet Neurol. 2008;7:875-84.

[15] Anderson C, Teo K, Gao P, et al. Renin-angiotensin system blockade and cognitive function in patients at high risk of cardiovascular disease: analysis of data from the ONTARGET and TRANSCEND studies. Lancet Neurol. 2011;10:43-53.

LA VALIDAZIONE DI UN MODELLO DI PREDIZIONE LINEARE: PRINCIPALI METODI E PROPOSTA DI UN INDICE ALTERNATIVO

Signoriello Simona[1], Vittorio Simeon[1], Chiodini Paolo[1]

[1] Dipartimento di Salute Mentale e Fisica e Medicina Preventiva, Università degli studi della Campania Luigi Vanvitelli

Introduzione

In ambito medico i modelli predittivi, che possono includere fattori di rischio, sistemi di stadiazione, caratteristiche del paziente, modelli statistici e algoritmi informatici, sono sviluppati con l'obiettivo di standardizzare e aumentare l'accuratezza della predizione di un determinato outcome. In questo contesto, sebbene diverse tecniche siano utilizzate per sviluppare i modelli, tutte utilizzano un campione di soggetti nel quale l'outcome è conosciuto ed è messo in relazione con le caratteristiche dei soggetti. Un problema comune nell'applicazione di un modello predittivo è che l'accuratezza predittiva degrada dal campione nel quale il modello è stato sviluppato all'applicazione seguente. Obiettivo della validazione è dunque quello di determinare se il modello si dimostra appropriato per lo scopo per il quale è stato sviluppato. La validazione è quindi un passo necessario per la verifica dell'applicabilità dei modelli predittivi. Ad esempio in ambito farmacogenetico validare un modello che ha l'obiettivo di predire la dose farmacologica sulla base delle caratteristiche fisiche sociali e genetiche del paziente, significa valutare se in un nuovo campione di soggetti tale modello è adeguato nel predire la dose del farmaco per il singolo individuo, confrontando i valori predetti dal modello con quelli osservati per ogni singolo soggetto.
Diversi metodi sono stati proposti in letteratura per valutare l'accuratezza di un modello in cui la variabile di outcome è di tipo continuo. Tra questi i più utilizzati sono il t test per dati appaiati, il coefficiente di determinazione R^2 e l'errore medio assoluto (MAE). Tuttavia, queste misure singolarmente non bastano per stabilire se il modello è accurato ma ognuna ha una sua specificità nella valutazione dell'accuratezza.

Obiettivi

Scopo di questo lavoro è di valutare i principali metodi proposti in lettera per la validazione in termini di accuratezza di un modello che predice una variabile quantitativa continua e proporre un nuovo indice. I diversi indici analizzati e il nuovo indice verranno confrontati simulando differenti scenari.

Metodi

Diversi metodi sono stati proposti in letteratura e utilizzati in differenti campi di studio. In molti casi la scelta della tecnica è limitata dal potenziale uso del modello, dal tipo di dati che il modello genera e dalla disponibilità di dati reali. L'dea di base è valutare se il modello predice accuratamente una determinata variabile quantitativa continua (supposta per comodità gaussiana con media μ e varianza σ^2) tramite un modello di regressione lineare semplice del tipo$Y=\alpha+\beta*Yp+\varepsilon$ dove Y sono i valori osservati e Yp quelli predetti dal modello che si vuole validare ed ε è il temine di errore con distribuzione $\mathcal{N}(0,\sigma^2)$.
Le tecniche di validazione comprendono metodi grafici, test statistici e misure di devianza.
Dal punto di vista grafico l'approccio più semplice è quello di costruire uno scatter plot ponendo sull'asse delle ascisse la variabile Yp e sull'asse delle ordinate la variabile Y e disegnando la retta Y=Yp che rappresenta la condizione di perfetta predizione. Questo grafico mostra in maniera diretta la bontà di adattamento dei dati dalla retta Y=Yp evidenziando la presenza di un eventuale bias. Un approccio grafico alternativo è quello proposto da Bland e Altman per descrivere la concordanza tra due misure quantitative [1]. Il grafico di Bland e Altman è uno scatter plot in cui sull'asse delle ordinate sono mostrate le differenze appaiate tra le due misure (Y-Yp) e sull'asse delle ascisse la media di queste ultime ((Y+Yp)/2). Dal punto di vista inferenziale, una prima valutazione della performance del modello è testare se esiste una differenza

statisticamente significativa tra i valori osservati e quelli predetti utilizzando il t test per dati appaiati. Tuttavia, questo approccio non esclude un eventuale bias nella predizione anche in presenza di un risultato non statisticamente significativo.
Un'altra misura utilizzata in letteratura è l'errore medio assoluto (MAE) definito come:

$$MAE = \frac{\sum|y - y_p|}{n}$$

che quantifica in termini assoluti l'errore sistematico della predizione rispetto ai valori osservati.
La validazione di un modello può essere visto anche come la valutazione della concordanza tra una quantità, una osservata (Y) e una predetta (Yp). Le osservazioni appaiate sono spesso impropriamente valutate utilizzando il coefficiente di correlazione lineare di Pearson (r) [2-6]. Questa misura tuttavia non consente di valutare se è verificata la condizione Y=Yp. In alternativa è possibile utilizzare il coefficiente di concordanza di Lin [7] calcolato come:

$$r_{c=}r * \frac{2s_y s_{y_p}}{s_y^2 + s_{y_p}^2 + \left(\bar{y} - \overline{y_p}\right)^2}$$

Il coefficiente di Lin è una modifica del coefficiente di correlazione di Pearson ottenuto moltiplicando il coefficiente r per un fattore di correzione che misura quanto la retta di regressione (Y=a+b*Yp) devia dalla retta di perfetta predizione (Y=Yp).
L'analisi di regressione dei valori osservati verso i predetti fornisce informazioni utili per valutare sia la presenza di un errore sistematico sia la bontà di adattamento. Il test composto sui parametri α e β permette di testare l'ipotesi nulla congiunta α=0 e β=1 evidenziando la presenza di eventuali bias nella predizione. Tuttavia, non consente di valutare la dispersione dei punti e pertanto al test va affiancata una misura di devianza per valutare la bontà dell'adattamento.
La misura più utilizzata in letteratura è il coefficiente di determinazione R^2 calcolato come:

$$R^2 = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(y - \bar{y})^2}$$

dove $\hat{y}$ sono i valori stimati dalla retta di regressione e $\bar{y}$ è la media della variabile Y. In ambito di validazione tuttavia, il problema fondamentale è che R^2 misura la bontà di adattamento rispetto alla retta di regressione (Y=a+b*Yp) e non alla retta di perfetta predizione (Y=Yp).
Una statistica adimensionale proposta da Loague e Green [8] è il modelling efficiency (EF) definito come:

$$EF = 1 - \frac{\sum(y - y_p)^2}{\sum(y - \bar{y})^2}$$

che può essere interpretato come una variabilità spiegata dalla retta perfetta Y=Yp. Tuttavia, questo indice può assumere valori tra $-\infty$ e 1 e quindi non conforme con le dimensioni di R^2.
Viene qui proposto un coefficiente di determinazione che consideri oltre alla variabilità dei punti attorno alla retta Y=a+b*Yp anche lo scostamento di quest'ultima dalla retta di perfetta predizione Y=Yp.
Abbiamo pertanto calcolato un inflation error (IE) come il rapporto tra le due devianze residue:

$$IE = \frac{\sum(y - y_p)^2}{\sum(y - \hat{y})^2}$$

dove il numeratore è la devianza residua relativa alla retta di perfetta predizione Y=Yp e il denominatore è la devianza residua relativa alla retta di regressione Y=a+b*Yp. Tale rapporto risulta sempre ≥1 ed è pari a 1 quando la retta di regressione stimata coincide con la condizione di perfetta predizione.
Scomponendo la devianza totale come somma della devianza residua e devianza spiegata dal modello, possiamo riscrivere l'R^2 come:

$$R^2 = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(\hat{y} - \bar{y})^2 + \sum(y - \hat{y})^2}$$

e ricalcolare R^2 moltiplicando la devianza residua per il rapporto IE nel seguente modo:

$$R^2_{IE} = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(\hat{y} - \bar{y})^2 + \sum(y - \hat{y})^2 * IE} = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(\hat{y} - \bar{y})^2 + \sum(y - y_p)^2}$$

tale quantità assume valori compresi tra 0 ed 1 e sarà inferiore a R2 se la retta di regressione stimata non coincide con la retta di perfetta predizione. È possibile inoltre calcolare la devianza residua relativa alla retta di perfetta predizione Y=Yp anche dalla varianza di errore di un modello di regressione con fissati a=0 e b=1 e calcolare così R201 che approssima R2IE.

Una simulazione è stata effettuata per verificare rispetto a diversi scenari la variabilità di alcuni degli indici presentati. In particolare sono stati ipotizzati 8 diversi scenari totali in base a 4 diverse combinazioni di α e β e 2 valori della deviazione standard dell'errore σ. In particolare sono stati considerati i due valori di σ=1 e σ=0.5 che corrispondono nella situazione di α=0 e β=1 a valori di R2 pari a 0.5 e 0.8 rispettivamente. Le 4 combinazioni dei valori α e β sono state: α=0 e β=1, α=1 e β=1, α=0 e β=2, α=1 e β=2. La variabile Yp è stata simulata da una N (0,1) e la dimensione del campione posta uguale a 200. Gli indici considerati nella simulazione sono stati MAE, r, rc, R2, EF, R2IE, R201 simulando per ciascun scenario 10.000 campioni. Per ciascun indicatore sono riportati mediana e intervallo interquartile.

Risultati

I risultati della simulazione sono riportati in Tabella1. La variabilità degli indicatori proposti è risultata coerente con i diversi scenari analizzati. Valori crescenti di MAE sono stati osservati nelle simulazioni in cui lo scostamento della retta stimata dalla retta di perfetta predizione era maggiore. I valori di r_c sono risultati inferiori ai valori di r. È stato osservato un incremento dell'IE direttamente proporzionale ai parametri α, β e inversamente proporzionale a σ. Nelle simulazioni con α=0 e β=1 i valori di R^2, R^2_{01}, R^2_{IE} e EF sono risultati approssimabili tra loro e il valore di IE è risultato vicino al valore atteso di 1. Inoltre, in tutti i diversi scenari R^2_{01} e R^2_{IE} sono risultati perfettamente sovrapponibili. Come atteso, il valore di R^2 è risultato dipendente dal valore del coefficiente angolare stimato mentre è risultato invariato al variare del valore dell'intercetta. La differenza tra R^2_{IE} e R^2 appare maggiore all'aumentare della distanza tra la retta stimata e la bisettrice. Tuttavia, mentre sono stati osservati valori di EF minori di zero in tutte le simulazioni il valore di R^2_{IE} è risultato tra 0 e 1.

Tabella1. Risultati della simulazione. I dati sono riportati come mediana e intervallo interquartile (Q1; Q3)

	α=0, β=1, σ=1		**α=1, β=1, σ=1**		**α=0, β=2, σ=1**		**α=1, β=2, σ=1**	
MAE	0.798	(0.770;0.827)	1.166	(1.128;1.205)	1.085	(1.050;1.120)	1.398	(1.360;1.436)
r	0.712	(0.691;0.733)	0.734	(0.716;0.753)	0.881	(0.873;0.889)	0.900	(0.892;0.907)
r_c	0.671	(0.653;0.689)	0.537	(0.518;0.555)	0.649	(0.640;0.659)	0.580	(0.570;0.589)
R^2	0.508	(0.477;0.537)	0.539	(0.512;0.567)	0.777	(0.761;0.791)	0.809	(0.797;0.822)
EF	0.504	(0.473;0.533)	0.072	(0.007;0.131)	0.581	(0.571;0.590)	0.415	(0.395;0.434)
IE	1.007	(1.003;1.014)	2.021	(1.905;2.144)	1.876	(1.772;1.987)	3.071	(2.884;3.282)
R^2_{IE}	0.506	(0.476;0.535)	0.367	(0.343;0.391)	0.649	(0.641;0.657)	0.580	(0.570;0.590)
R^2_{01}	0.507	(0.477;0.536)	0.368	(0.344;0.393)	0.651	(0.642;0.659)	0.582	(0.571;0.591)
	α=0, β=1, σ=0.5		**α=1, β=1, σ=0.5**		**α=0, β=2, σ=0.5**		**α=1, β=2, σ=0.5**	
MAE	0.399	(0.385;0.413)	1.009	(0.985;1.032)	0.911	(0.890;0.932)	1.230	(1.208;1.251)
r	0.889	(0.881;0.897)	0.883	(0.875;0.891)	0.972	(0.970;0.973)	0.971	(0.969;0.973)
r_c	0.882	(0.874;0.889)	0.584	(0.572;0.595)	0.763	(0.758;0.768)	0.640	(0.634;0.646)
R^2	0.790	(0.776;0.804)	0.780	(0.765;0.795)	0.944	(0.940;0.948)	0.942	(0.939;0.946)
EF	0.788	(0.774;0.802)	-0.113	(-0.176;-0.051)	0.707	(0.701;0.712)	0.470	(0.458;0.481)
IE	1.007	(1.003;1.014)	5.052	(4.735;5.412)	5.230	(4.887;5.574)	9.223	(8.622;9.844)
R^2_{IE}	0.789	(0.774;0.802)	0.412	(0.396;0.428)	0.763	(0.759;0.767)	0.640	(0.635;0.645)
R^2_{01}	0.789	(0.775;0.803)	0.413	(0.397;0.429)	0.764	(0.760;0.767)	0.641	(0.636;0.646)

Conclusioni

Nella validazione di un modello predittivo con outcome continuo sono diverse le metodologie grafiche e analitiche proposte in letteratura. Tuttavia, alcune delle misure di accuratezza proposte e valutate in questo lavoro hanno presentato dei limiti. In questo contesto è necessario valutare sia la presenza di un eventuale bias sistematico tra la variabile osservata e quella predetta sia l'associazione tra le variabili. L'indice di concordanza di Lin (r_c) pur unendo entrambe le valutazioni è definito in un contesto di concordanza nel quale non esiste una differenziazione tra valori osservati (Y) e predetti (Yp). Tra le misure basate sulla varianza spiegata l'indice EF risulta essere caratterizzato dalla mancanza di un limite inferiore che ne limita l'interpretazione. L'indicatore proposto R^2_{IE} nasce con l'obiettivo di aggiustare l'indice R^2 per lo scostamento dalla retta di perfetta predizione, ma mantenendo un range di valori compreso tra 0 e 1. La variabilità di questo indice è risultata coerente con le simulazioni valutate e quindi è, a nostro parere, una valida alternativa nel processo di validazione di un modello predittivo con outcome continuo.

Bibliografia

[1] Bland J.M., Altman D.G. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307-10.

[2] Cuervo-arango J., Newcombe J.R. Repeatability of preovulatory follicular diameter and uterine edema pattern in two consecutive cycles in the mare and how they are influenced by ovulation inductors. Theriogenology. 2008;69(6):681-7.

[3] Vyt P., Maes D., Rijsselaere T., et al. Motility assessment of porcine spermatozoa: a comparison of methods. ReprodDomest Anim. 2004;39(6):447-53.

[4] Nagy P, Solti L, Kulcsár M, et al. Progesterone determination in equine plasma using different immunoassays. Acta Vet Hung. 1998;46(4):501-13.

[5] Christensen P, Hansen C, Liboriussen T, et al.. Implementation of flow cytometry for quality control in four Danish bull studs. AnimReprod Sci. 2005;85(3-4):201-8.

[6] Colazo MG, Ambrose DJ, Kastelic JP, Small JA. Comparison of 2 enzyme immunoassays and a radioimmunoassay for measurement of progesterone concentrations in bovine plasma, skim milk, and whole milk. Can J Vet Res. 2008;72(1):32-6.

[7] Lin L.I. A concordance correlation coefficient to evaluate reproducibility. Biometrics. 1989;45(1):255-68

[8] Loague K., Green R.E. Statistical and graphical methods for evaluating solute transport models: Overview and application. J. Contam. Hydrol. 1991; 7:51-73

COMPARISON OF APPROACHES TO MODEL SEQUENTIAL MULTIPLE MEDIATORS

Zugna Daniela[1], Fasanelli Francesca[1], Popovic Maja [1], Richiardi Lorenzo[1]

Unit of Cancer Epidemiology, Department of Medical Sciences, University of Turin

Introduction

Mediation analysis aims to disentangle the effect of an exposure on the outcome into a direct effect and an indirect effect through an intermediate variable on the causal pathway between the exposure and the outcome. An extension of standard mediation analysis involving a single mediator to multiple mediators is required when there is a confounder of the mediator-outcome association affected by the exposure and/or when the research question focuses on multiple mechanisms.

Aim

To compare different approaches to model sequential multiple mediators through an application to real data from a birth cohort study.

Methods

We used data from the NINFEA birth cohort to estimate the amount of the effect of maternal depression/anxiety in pregnancy on the risk of wheezing between 6 and 18 months of life explained by adverse reproductive outcomes (low birth weight, preterm birth, caesarean section) and by lower respiratory infections in the first 6 months of life. We assumed that the adverse reproductive outcomes (M_1) affect the respiratory infections in the first 6 months (M_2). We estimated the direct and indirect effects through M_1 and then through M_1 and M_2 jointly by using different approaches: the weighting approach[1], the odds ratio weighting approach[2], the imputation approach[3] and the extension of imputation-based approach[4].

Results

The analysis was performed on complete data of 4797 participants. The total effect, in terms of relative risk, of maternal depression/anxiety in pregnancy on the risk of wheezing was 1.64 (95% CI:1.31-2.05). By the extension of imputation-based approach, the natural direct effect was 1.57 (95% CI: 1.26-1.97), the natural indirect effect through M_1 and M_2 was 1.05 (95% CI: 1.01-1.09), the natural direct effect through M_1 was 1.00 (95% CI: 0.99-1.00) and the partial indirect effect through M_2 was 1.05 (95% CI: 1.01-1.09). Similar results were obtained by the other approaches although these did not provide the partial indirect effect through M_2.

Conclusions

Although the validity of the described approaches depends on different assumptions, they give similar results: most of the effect of maternal depression/anxiety in pregnancy on the infant wheezing is direct, while only lower respiratory infections in the first 6 months of infant life seem to have a weak mediating role.

References

[1] VanderWeele TJ, Vansteelandt S. Mediation analysis with multiple mediators. Epidemiol Methods. 2014;2:95-115.

[2] Tchetgen Tchetgen EJ. Inverse odds ratio-weighted estimation for causal mediation analysis. Stat Med. 2013;32:4567-80.

[3] Vansteelandt S, Bekaert M, Lange T.Imputation strategies for the estimation of natural direct and indirect effects. Epidemiol Methods. 2012;1:131-158.

[4] Steen J, Loeys T, Moerkerke B, and Vansteelandt S. Flexible Mediation Analysis With Multiple Mediators. Am J Epidemiol. 2017;186:184-193.

Acknowledgement

LifeCycle (European Union's Horizon 2020 Research and Innovation Programme)



Atti | Sessione Giovani

THE VALIDATION OF A HIGHLY SKEWED SURROGATE FOR A SURVIVAL ENDPOINT: NEW APPROACHES

Capitoli Giulia[1], Rotolo Federico[2], Rebora Paola[1], Valsecchi Maria Grazia[1], Galimberti Stefania[1]

[1] *Center of Biostatistics for Clinical Epidemiology, School of Medicine and Surgery, University of Milano - Bicocca, Italy*

[2] *Biostatistics and Data Management Unit, Innate Pharma, Marseille, France*

Introduction

The selection of the primary "endpoint" is a very important step in the design of clinical trials. Typically, the goal of a clinical trial is to assess the effect of treatment on this endpoint. It often appears, however, that the most sensitive and relevant clinical endpoint, which will be called the "true" endpoint, might be difficult to assess and to overcome this problem, a seemingly attractive solution is to replace the true endpoint with another one, which is measured earlier, more conveniently or more frequently. Such "replacement" endpoint is termed "surrogate" endpoint and has the purpose of evaluating the effect of a specific treatment for a specific disease. Once a candidate surrogate is identified according to some specific properties, several formal methods are available for the validation of the surrogate endpoint, depending on the number of the trials performed.
The first formal statistical approach dates back to 1989, when Prentice proposed a definition of surrogate endpoint and four criteria to validate it. The most important criterion among these is called "The Prentice's Criterion", which implies that "the full effect of treatment upon the true endpoint is captured by the surrogate" (Prentice 1989). More recently, a meta-analytic approach to the validation of a surrogate endpoint was proposed for a multi-trial context [2-3]. It consists in estimating associations at two different levels: the association between the surrogate and the clinical endpoint, called the "individual-level association", and the association between the effect of treatment on the surrogate and on the clinical endpoint, called "trial-level association". A good surrogate is one that has biological/clinical plausibility and is shown, statistically, to have strong individual-level and trial-level associations with the true endpoint.
In this context, no specific method exists to deal with a continuous variable with a spike in zero as surrogate endpoint for a time-to-event outcome.
The motivating clinical question is to assess whether early minimal residual disease (MRD) is a surrogate for event-free-survival in childhood B-lineage acute lymphoblastic leukemia. MRD, which quantifies at the molecular level the number of circulating leukemic cells, is a skewed marker with a high proportion of zero values.

Objective

The main goal of the present work is developed methods to deal with this zero-inflated distribution for the surrogate endpoint and we investigated their adequacy.

Methods

We explored through simulations the appropriateness of different parametric forms for the surrogate endpoint, considering both fully continuous and a mixed discrete-continuous distribution. We assessed the Beta and Gamma (with inverse and log links) distributions and the hurdle gamma model that accounts for two treatment effects for the surrogate endpoint. Additionally, by modelling the zeros of the surrogate as small values below the limit of detection, we evaluated the possibility to refer to the standard validation setting of a survival surrogate (with left censored observations) and a survival true endpoint. Our proposals

were also compared to the simplified approach, we used to address the validation of MRD, that considers the continuous surrogate as categorical ordinal endpoint.

Results

The estimation of the individual-level parameter is unbiased and precise, regardless of the parametric distribution, in all scenarios and for low and high values of association. As for the trial-level surrogacy, the parametric solutions are the most promising. However, there is loss in accuracy and underestimation of the trial-level association when the distribution is heavily skewed.

Conclusions

When validating a continuous surrogate for a survival true endpoint, the presence of a zero-inflated highly skewed surrogate distribution can be managed without recurring to methods that reduce the surrogate to an ordinal variable.

Bibliography

[1] Prentice, R.L. Surrogate endpoints in clinical trials: definitions and operational criteria. Statistics in Medicine, 1989; 8: 431–440.

[2] Buyse, Marc, et al. "The validation of surrogate endpoints in meta-analyses of randomized experiments." Biostatistics, 2000; 1.1: 49-67.

[3] Burzykowski, Tomasz, et al. "Validation of surrogate end points in multiple randomized clinical trials with failure time end points." Journal of the Royal Statistical Society: Series C (Applied Statistics), 2001; 50.4: 405-422.

BAYESIAN MODEL AVERAGING TO BUILD A COMORBIDITY SCORE USING HEALTHCARE UTILIZATION DATABASES

Bucci Andrea[1], Faragalli Andrea[1], Ferrante Luigi[1], Skrami Edlira[1], Gesuita Rosaria[1], Dajko Marianxhela[1], Carle Flavia[1]

[1] *Centro di Epidemiologia, Biostatistica e Informatica medica, Università Politecnica delle Marche, Ancona*

Introduction

Evaluating the presence of multiple illnesses in a patient, i.e. comorbidity, can be particularly useful to predict health outcomes and categorize the severity of the presence of multiple conditions. In a first attempt to classify the severity of the presence of several conditions, Charlson et al. [1] introduced a prognostic index for comorbid conditions that might somehow influence the risk of 1-year mortality in patients. The authors relied on a list of potential prognostic variables whose adjusted relative risk was used to create a weighted index. More recently, Corrao et al. [2] developed a comorbidity score, the *Multisource Comorbidity Score* (MCS), predictive of mortality, hospitalisation and high health costs, that relied on multiple source information from the administrative Italian National Health System (NHS) databases. The MCS had the objective to obtain a better control of the confounding factors in epidemiological studies and dispose a tool for the stratification of the population on the basis of its prognostic fragility [2].
The list of diseases and conditions used in the work of Corrao et al. [2] to predict the desired outcome had been selected from a large set of diseases/conditions utilized in the already Charlson's Comorbidity Score (CCS), Elixhauser's Index (EI), and MCS through a Least absolute shrinkage and selection operator (LASSO) regression. Nonetheless, this method, used to automatically select a set of explicative variables, did not account for the uncertainty in the model. This aspect could be crucial in the definition of a score on the presence of multiple diseases or conditions. For this reason, we chose the list of covariates defining the comorbidity score by applying a Bayesian model Averaging (BMA) approach [3,4]. BMA allowed us to implement variable selection and coefficients estimation simultaneously, without discarding any kind of information.

Objective

This work aimed at understanding whether avoiding the *a priori* discard of potentially relevant conditions could lead to a better control of confounding factors in terms of mortality prediction.

Methods

Our database included all subjects resident in Marche Region who received NHS assistance in 2015-2016; data were collected from regional healthcare utilization databases that include information on i) demographic characteristics, ii) hospital discharge records with information on primary diagnosis up to six coexisting conditions, coded in accordance with the International Classification of Diseases 9th Revision (ICD-9) Clinical Modification classification system (http://icd9.chrisendres.com/), iii) drug prescriptions which have been reimbursed by NHS, identified by the Anatomical Therapeutic Chemical (ATC) classification [5,6].

Candidate predictors selection. With the purpose to keep the model feasible, we selected a list of candidate covariates already entailed in the comorbidity indices used as reference in this work, i.e. CCS, EI and MCS. Thus, the final set of potential predictors included 46 conditions/diseases identified by hospital discharges records, drug prescriptions or both.

Score development. The list of diseases/conditions capable of predicting 1-year mortality was selected by Bayesian Model Averaging (BMA). This methodology allows the researcher to consider all the uncertainty in the selection of the model to estimate the coefficients, which are used to assign a score to each selected covariate. Due to this specification, a large number of potential predictors can be included in the set of covariates and their effect in the estimate can be evaluated in a more accurate way.
Given a set of *k* possible predictors, there are 2^k possible combinations of the *k* covariates. The idea behind BMA is that the relevance of a potential model can be determined by its a *posteriori* probability for each of the 2^k possible models.
Once specified the *a priori* probability and determined the likelihood of each model, the *a posteriori* probability of the single model is simply the probability that it is the true model. In this way, each of the 2^k models is considered in the information set of the estimation process and no information is discarded. This allows the researcher to have an improved predictive accuracy respect to the one-shot selection provided by other techniques (e.g. LASSO).
In order to identify which variables should contribute to the definition of the score, we performed BMA on a training sample of 417 396 randomly selected individuals. The considered subjects, at baseline (in 1st January 2017), were 50 years or older, NHS beneficiaries and resident in Marche Region from at least 2 years resident in one of the ten participant Italian regions from at least two years (i.e., were recorded as beneficiaries of the regional NHS before the year 2015). The conditions/diseases were retrieved from the database of Marche Region and entered in the model as binary variables, where 1 indicated the presence of the condition in the patient at least once within 2 years before the baseline (2015-2016) and 0 otherwise. The set of predictors further included the gender and the age of the subject (in 1st January 2017).
We applied BMA by fitting a survival model based on Cox regression to predict 1-year mortality [7]. Using BMA procedure, we obtained a posterior probability of being in the true predictive model for each covariate that can be used to define a final score.
The definition of the score of each disease/condition for the determination of the score was the combination of the estimated coefficient of the disease/condition and its posterior probability of inclusion in the true predictive model. The coefficients estimated from the Cox-BMA were weighted according to the posterior probability and converted into scores by multiplying the products by 10 and rounding them to the higher whole number, then for each subject the scores were sequentially summed to produce a total aggregate score.
The total aggregate score obtained through the application of the above presented methodology was categorised in 5 categories of aggregate score, 0-4, 5-9, 10-14, 15-19, and ≥20.
The distribution of the score was evaluated by stratifying the population according to the age and the gender. The new score was validated on a testing sample of 208 698 subjects, with the same inclusion/exclusion criteria of the training sample, in comparison with MCS, CCS and EI. We used the receiver operating characteristic (ROC) curve and the area below it (AUC) with relative 95% Confidence Interval (95% CI) to assess the discriminatory power of our score. MBCS was further validated by evaluating the net proportion of subjects correctly reclassified, through net reclassification improvement (NRI), among those who experienced or not the outcome.

Results

As shown in Table 1, the disease/conditions mostly determining the score were metastatic cancer, kidney dialysis and non-metastatic cancer. The list of diseases not contributing to the total score, i.e. those with zero a posteriori probability of being included in the true predictive model, entailed hypothyroidism, obesity, depression, drug dependence, anxiety, coronary disease, peptic ulcer, Crohn's disease, kidney transplant, rheumatological condition, glaucoma, hypertension, pain, and hyperlipidaemia.
By analysing the MBCS distribution in the testing sample among categories, several results emerged. On the one hand, 72.9% of subjects belonged to the lowest category of MCS (0-4), 12.8% to the second

category (5-10), 8.5% to the third (10-14), 3.1% to the fourth (15-19), and 2.7% to the highest category (≥20).
On the other hand, when the sample was stratified by age and gender, it could be observe that the subjects with the lowest MBCS category changed from 89.6% to 46.41% in females when we compared subjects with age comprised between 50 and 59 and age greater than 80, and from 88.6% to 43.6% in males. Conversely, for the same age groups, the percentage of NHS beneficiaries belonging to the highest MBCS varied from 0.8% to 6.0% in females, and from 0.8% to 7.5% in males.
The novel score showed the best discriminatory power among the compared scores. In fact, the AUC (95% CI) of MBCS, MCS, CCS and EI were respectively 0.81 (0.80-0.82), 0.79 (0.78-0.81, p-value <0.001), 0.68 (0.67-0.69, p-value < 0.001), and 0.67 (0.66-0.68, p-value < 0.001). Furthermore, the net reclassification of 1-year mortality, assessed by NRI, improved in our score, specifically by 4.92% (95% CI 63-72, p-value < 0.01), 19.37% (95% CI 67-76, p-value < 0.001), and 19.65% (95% CI 73-81, p-value < 0.001) when compared with MCS, CCS and EI, respectively.

Table 1. *Definition of weights of MBCS for each disease/condition*

Disease/condition	**BMA coefficient (SE*)**	**P(β ≠0)[§]**	**Weight**
Metastatic cancer	1.73 (0.05)	100.0	18
Kidney dialysis	0.91 (0.16)	100.0	10
Cancer	0.87 (0.04)	100.0	9
Psychoses	0.82 (0.04)	100.0	9
Liver disease	0.9 (0.08)	100.0	9
Malignancy medication	0.78 (0.05)	100.0	8
Heart failure	0.67 (0.02)	100.0	7
Dementia	0.52 (0.05)	100.0	6
Alcohol abuse	1.05 (0.29)	55.9	6
Weight loss	0.49 (0.11)	100.0	5
Disorders of fluid, electrolyte and acid-base balance	0.41 (0.06)	100.0	5
Anaemias	0.43 (0.04)	100.0	5
Hemiplegia and hemopoiesis	0.49 (0.12)	99.4	5
Parkinson's disease	0.49 (0.04)	100.0	5
HIV	1.53 (0.45)	23.6	4

Lymphoma	0.33 (0.09)	94.2	4
Diabetes	0.34 (0.02)	100.0	4
Other neurological disease	0.32 (0.06)	100.0	4
Cerebrovascular disease	0.37 (0.03)	100.0	4
Kidney disease	0.34 (0.04)	100.0	4
Epilepsy	0.25 (0.03)	100.0	3
Acute myocardial infarction	0.28 (0.06)	100.0	3
Vascular diseases	0.26 (0.02)	100.0	3
Gout	0.13 (0.02)	100.0	2
Arrhythmia	0.15 (0.03)	100.0	2
Transplant (any kind, kidney excluded)	0.47 (0.15)	32.7	2
Tuberculosis	0.32 (0.15)	1.6	1
Coagulation defects	0.39 (0.17)	2.7	1
Bipolar disorder	0.47 (0.22)	1.5	1
Valvular diseases	0.16 (0.06)	20.1	1
Chronic pulmonary diseases	0.06 (0.02)	11.8	1
Cystic fibrosis	0.16 (0.06)	8.0	1

* Standard error

§P(β ≠0) is the posterior inclusion probability that the coefficient is non-zero.

Conclusions

This work introduced a novel comorbidity score, MBCS, based on multiple administrative databases regarding NHS beneficiaries. In a first analysis, it has been showed that MBCS is capable of accurately stratifying the population in accordance with their 1-year risk of death. When compared with other prognostic scores in terms of discriminatory power, MBCS exhibited the overall best performance. Moreover, it improved the net 1-year mortality reclassification by all other scores.
Interestingly, the list of covariates affecting 1-year mortality in MBCS was slightly different from MCS. For example, when the estimation of the coefficients was made through BMA-Cox, the alcohol abuse exhibited a mitigated effect on 1-year mortality. Conversely, kidney dialysis seemed to worsen the one-year survival probability in a more impacting way. Moreover, the effect of tuberculosis in our score was strongly affected by its inclusion probability. In fact, even if the size of the coefficient was comparable with other relevant diseases/conditions, the probability that this covariate was relevant to explain 1-year mortality was

particularly low (i.e. 0.0016). This means that it is highly unlikely that this covariate affects 1-year mortality also when the set of covariates changes.
It may be also noticed that some diseases/conditions that necessitate a long-lasting treatment, such as hypertension, glaucoma, Crohn's disease, or rheumatological conditions, did not positively impact on 1-year mortality or show a zero-inclusion probability in the true model. Moreover, in some cases (e.g. hyperlipidaemia, glaucoma and hypertension), the estimated coefficients showed a negative effect on mortality, and were not included in the set of potential sources of the score.
It should be further noticed that our score, even if based on the same list of diseases/conditions used in MCS, CCS and EI, relied on a totally new methodology for such literature. In fact, the implementation of BMA to estimate the survival regression coefficients and to define the weights of the disease/condition guaranteed that none of the information was discarded. This does not happen, for example, in other techniques like LASSO, and helps defining more robust weights based both on the estimated conditional value of the coefficient and the probability of being included in the true model.
Finally, as for MCS, the usage of multiple databases was a semi-novelty for the literature of comorbidity scores and helped reducing possible distortions in the attribution of the disease/condition to each NHS beneficiary.
The future extensions of the score could entail a larger list of diseases and the evaluation of MBCS predictive accuracy on other outcomes including 5-year all-cause mortality, 1-year and 5-year hospital admissions for all causes.

Bibliography

[1] Charlson M. E., Pompei P. et al., A New Method of Classifying Prognostic Comorbidity in longitudinal studies: Development and Validation. Journal of Chronic Diseases , 1987, 40(5): 373-383

[2] Corrao G., Rea F., Di Martino M. et al., Developing and validating a novel multisource comorbidity score from administrative data: a large population-based cohort study from Italy. BMJ Open, Dec. 2017, 7 (12)

[3] Hoeting J.A., Madigan D., Raftery A. E. et al., Bayesian Model Averaging: A Tutorial. Statistical Science, 1999, 14 (4): 382-417

[4] Raftery A.E., Bayesian model selection in structural equation models. In Bollen, K.A., & Long, J.S. Testing structural equation models. Newbury Park, CA: Sage, 163:180

[5] Huber C.A., Thomas D. S. et al., Identifying patients with chronic conditions using pharmacy data in Switzerland: an updated mapping approach to the classification of medications. BMC Public Health, 2013, 13: 1030

[6 Halfon P., Eggli Y. et al., Disease identification based on ambulatory drugs dispensation and in-hospital ICD-10 diagnoses: a comparison. BMC Health Services Research, 2013, 13:453

[7] Watts K. L., Bayesian Methods and Computation for Large Observational Datasets. Doctoral Dissertation, Harvard University, 201

A META-ANALYTIC APPROACH TO ESTIMATE THE PROPORTION OF LONG-TERM SURVIVORS IN IMMUNE CHECKPOINT INHIBITORS CLINICAL TRIALS

Peveri Giulia[1], Oriecuia Chiara[1,2], Pagan Eleonora[2], Conforti Fabio[3], Bagnardi Vincenzo[2], Specchia Claudia[1]

[1] *Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy*

[2] *Department of Statistics and Quantitative Methods, University of Milan-Bicocca, Milan, Italy*

[3] *Division of Medical Oncology for Melanoma and Sarcoma, European Institute of Oncology, Milan, Italy*

Introduction

Ipilimumab, an anti-CTLA4 anticancer immunotherapy approved in March 2011 by the FDA and in July of the same year by the EMA, was the first agent demonstrated to increase overall survival (OS) in patients treated for advanced melanoma. A pooled analysis of trials testing Ipilimumab for the treatment of unresectable or metastatic melanoma, based on individual patient's data (IPD), was published in 2015 by Schadendorf et al [1] and widely cited thereafter: in this study, survival data on 1861 Ipilimumab-treated patients coming from 12 different studies (two phase III, five phase II, three phase I/II and two retrospective studies) were considered. The OS Kaplan-Meier (KM) curve reached a plateau starting at approximately 3 years (with a follow-up as long as 10 years for some patients), suggesting the presence of a proportion of long-term survivors of approximately 21%, empirically derived from the pooled survival curve. However, authors did not use in their analyses meta-analytic methods in order to take into account heterogeneity between trials, neither tried to formally estimate the proportion of long-term survivors using suitable statistical models.

Objectives

Our aim is to propose a novel method, based on cure models [2], to estimate the proportion of long-term survivors from published data, using a meta-analytical approach.
As a motivating example, we applied our method to the data extracted from the clinical trials included in the original pooled analysis on Ipilimumab [1].

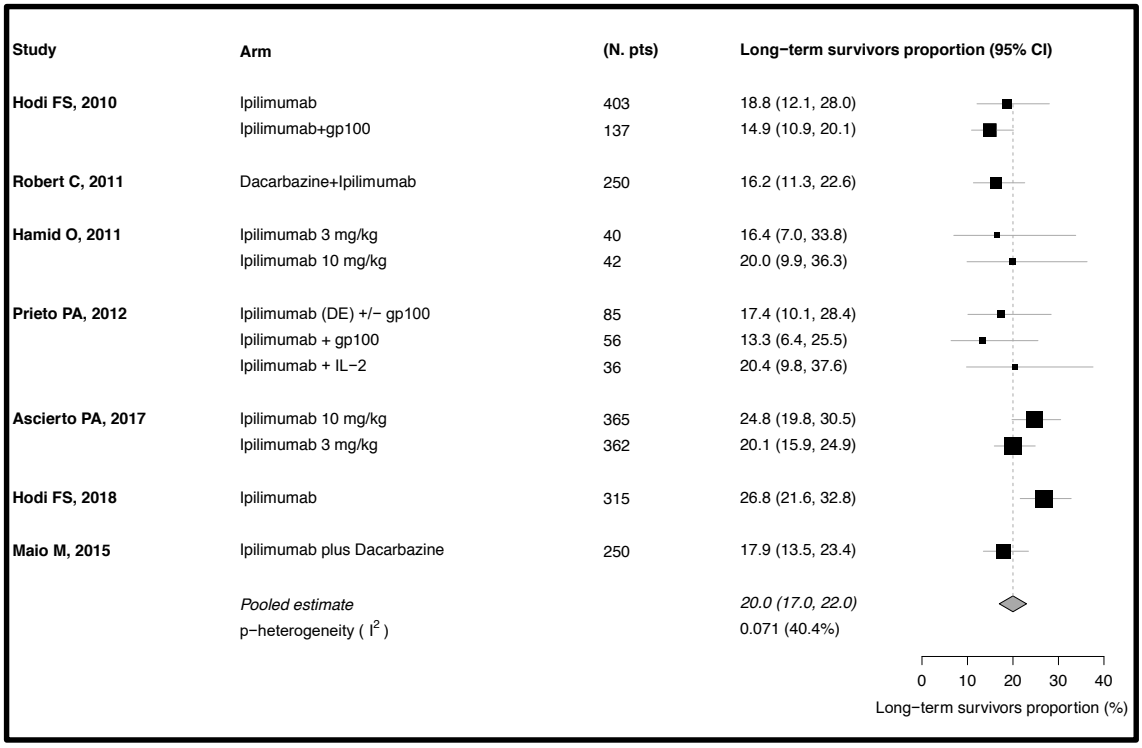
Methods

Cure models allow both the estimation of the fraction of long-term survivors and the investigation of what covariates are associated with either short-term or long-term effects. An important subclass of cure models is represented by mixture models, which explicitly consider two subgroups of patients: those who are susceptible to treatment, and those who are not. They are formally described through the mixture of a logistic regression model (to model the fraction of cured patients) and a parametric or semi-parametric survival function to model the hazard of event in not cured patients [2].
The estimate of the proportion of long-term survivors based on cure models in a meta-analytical framework requires the impractical task of obtaining IPD directly from trials investigators. On the other hand, IPD could be reconstructed from published Kaplan-Meier survival curves using the algorithm proposed by Guyot et al, which requires a digitized version of the curve and the number of patients at risk at two or more time-points of the follow-up [3]. We therefore digitized all the published OS KM Ipilimumab arms from the 12 studies included in the original pooled analysis [1], and, using the algorithm, extracted the pseudo-IPD.
We then estimated the fraction of long-term survivors based for each individual curve using survival mixture cure models. Only studies with at least one event observed after 3 years of follow-up were considered. Finally, we applied a two-stage random-effects model to obtain the pooled estimate of the long-term survivor

proportion. Heterogeneity between trials was assessed through Cochran's Q-test, and quantified through the I^2 statistic, which represents the percentage of variation across studies attributable to heterogeneity rather than chance. All analyses (including IPD reconstruction) were conducted using R 3.5.1 and SAS 9.4.

Results

The pooled OS KM curve obtained using pseudo-IPD was very similar to the curve presented in the original pooled analysis based on actual IPD: the median OS calculated from our data was 11.7 months (the original being 11.4 months), and the 3-year OS was 22%, identical to the original.
Seven out of 12 included studies reported at least one event after 3 years of follow-up and were therefore considered in the estimation of the proportion of long-term survivors based on cure-models.
The meta-analytic pooled estimate of the long-term survivor proportion was 20% (95% CI: 17%-22%). The heterogeneity between study-specific estimates, quantified by I^2, was 40%.
The figure below shows the study-specific estimated proportion of long-term survivors, along with the pooled estimate.



Conclusions

Not surprisingly, our meta-analytic pooled estimate of long-survivor proportion of 20% was close to the empirical 21% proportion, based only on the visual examination of the shape of the pooled OS curve, reported by Schadendorf et al. However, we were able to provide an interval estimate of the long-term survivor proportion, taking into account at the same time the between study heterogeneity.
To our knowledge, this is the first application of cure models in a meta-analytical framework.
We proved the usefulness of pseudo-IPD reconstruction method to extract novel measure of interest from published KM curves, such as the long-survivor fraction.

Bibliography

[1] Schadendorf D., Hodi F.S., Robert C. et al. Pooled analysis of long-term survival data from phase II and phase III trials of ipilimumab in unresectable or metastatic melanoma. J Clin Oncol., 2015; 33:1889–94

[2] Othus M., Barlogie B., LeBlanc M.L. at al. Cure models as a useful statistical tool for analyzing survival. Clin Cancer Res., 2012; 18:3731–6

[3] Guyot P., Ades A.E., Ouwens M.J.N.M., et al., Enhanced secondary analysis of survival data: Reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol., 2012; 12:9

STATISTICAL ISSUES ARISING IN SYSTEMATIC REVIEWS AND META-ANALYSES OF IMMUNE CHECKPOINT INHIBITORS CLINICAL TRIALS. AN APPLICATION TO MELANOMA.

Pagan Eleonora[1], Oriecuia Chiara[1,2], Peveri Giulia[2], Conforti Fabio[3], Specchia Claudia[2], Bagnardi Vincenzo[1]

[1] *Department of Statistics and Quantitative Methods, University of Milan-Bicocca, Milan, Italy*

[2] *Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy*

[3] *Division of Medical Oncology for Melanoma and Sarcoma, European Institute of Oncology, Milan, Italy*

Introduction

Immune checkpoint inhibitors (ICIs) are a class of immunotherapies that represents the new frontier of cancer treatment.
Due to their biological mechanism of action, ICIs may not show early benefit [1]; moreover, by breaking immune-tolerance against tumor-associated antigens, ICIs prolonged tumor control and lead to a non-negligible fraction of patients achieving a prolonged survival, suggesting that many patients could be potentially be cured of their disease [2]. Both these phenomena lead to survival curves that, when compared to standard chemotherapy (CT), are not proportional [3]. However, the hazard ratio (HR) estimated from a Cox proportional hazard (PH) model is routinely used to empirically quantify the between-group difference, leading to underpowered and biased estimates of treatment effect [4].
Despite alternative methodologies to take into account non-proportionality of hazards, such as the use of restricted mean survival times (RMST) [5], have been proposed, they are rarely used and presented in published RCTs testing the efficacy of ICI.

Aims

The aim of the present study was to derive and reanalyze overall survival (OS) and progression-free survival (PFS) individual patient data (IPD) from published RCTs in order to test the proportional hazard (PH) assumption and to estimate the treatment effect in terms of RMST ratio. We considered RCTs testing the efficacy of four ICI agents for metastatic melanoma (Ipilimumab, Tremelimumab, Nivolumab and Pembrolizumab), in combination or not with standard CT versus control arms receiving only CT.

Methods

A systematic review of all phase II and III RCTs testing anti-CTLA4 or anti-PD1 drugs, in patients with advanced melanoma, was conducted.
The test of PH assumption and the estimate of RMST in a meta-analytical framework necessitate the, often impractical, task of obtaining IPD directly from trials investigators. On the other hand, IPD could be reconstructed from published Kaplan-Meier survival curves using the algorithm proposed by Guyot et al, which requires a digitized version of the curves and the number of patients at risk at two or more time-points of the follow-up [6]. All the OS and PFS Kaplan-Meier curves presented in the included studies were therefore digitized, and, using the algorithm, the pseudo-IPD were extracted.
The PH assumption was tested for each included trial for both OS and PFS using the Kolmogorov-Smirnov type supremum test [7].
Study-specific RMST ratio were then computed. Briefly, RMST is a measure of treatment effect on the scale of the time to event. RMST is valid under any distribution of the survival in the treatment groups, and it is obtained integrating the observed survival function between 0 and some prespecified timepoint t*. RMST can be interpreted as the mean survival time up to t*; the ratio of RMSTs quantifies the relative gain in life,

or progression-free, expectancy associated with the experimental treatment with respect to the standard treatment.
Heterogeneity between estimates was assessed with Cochrane's Q-test and the I^2 (percentage of the total observed variability do to heterogeneity) was also reported.
Single-study measures were pooled using a random-effects meta-analytic approach.
All analysis (including IPD reconstruction) were performed using R software v. 3.5.1 and SAS software v. 9.4.

Results

Seven RCTs were selected. All overall survival curves fulfilled the PH assumption. Conversely, for all the included trials, the progression-free survival curves significantly deviated from the proportional hazard assumption. The pooled RMST ratio for OS and PFS at 3 years were 1.26 (95% CI, 1.03-1.54) and 1.61 (95% CI, 1.00-2.60), respectively.
Significant heterogeneity between single study estimates was observed for both the outcomes.

Conclusions

We reported that, in RCTs investigating the ICI effects in advanced melanoma, overall survival curves satisfied the PH assumption. In this situation, the test of the null hypothesis based on RMST ratio has operating characteristics similar to the test based on the HR and the significance level and power of the two tests appear to be similar [5].
On the other hand, we showed in the same setting that all progression-free survival functions deviates from the PH assumption. In this situation, the RMST ratios are preferable than the HRs since they are free from assumptions and valid under any distribution of the time to event in the treatment groups.
The majority of immunotherapy trials are designed and analysed using a target HR for the treatment effect but the biologic mechanism of action of immunotherapies is unconventional. The design of future immunotherapy trials should be guided by the expectation of deviation from PH assumption [3] and other measures, instead of HR, should be considered in the design and analysis.

Bibliography

[1] Champiat S., Ferrara R., Massard C., et al., Hyperprogressive disease: recognizing a novel pattern to improve patient management. Nature Reviews Clinical Oncology, 2018;15:748–762

[2] Hamid O., C Robert C., Daud A., Five-year survival outcomes for patients with advanced melanoma treated with pembrolizumab in KEYNOTE-001. Ann Oncol., 2019; 30:582-588

[3] Alexander B.M., Schoenfeld J.D., Trippa L., Hazards of Hazard Ratios - Deviations from Model Assumptions in Immunotherapy. N Engl J Med., 2018;378:1158-9

[4] Mick R., Chen T.T., Statistical Challenges in the Design of Late-Stage Cancer Immunotherapy Studies. Cancer Immunol Res., 2015;3:1292-8

[5] Royston P., Parmar M.K.B., Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol., 2013;13:152

[6] Guyot P., Ades A.E., Ouwens M.J.N.M., et al., Enhanced secondary analysis of survival data: Reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol., 2012;12:9

[7] Lin D.Y., Wei L.J., Ying Z., Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals. Biometrika, 1993;80:557–72.

TEXT MINING E INEXACT MATCHING PER IDENTIFICARE DITTE CON ESPOSIZIONE AD AMIANTO.

Airoldi Chiara[1], Ferrante Daniela[1], Miligi Lucia[2], Migliore Enrica[3], Tumino Rosario[4], Romanelli Antonio[5], Mensi Carolina[6], Cavone Domenica[7], Romeo Elisa[8], Marinaccio Alessandro[9], Elisabetta Chellini[10], Magnani Corrado[1]

1 Università del Piemonte Orientale, Dipartimento di Medicina Traslazionale, Novara

2 Centro operativo regionale dei Tumori a bassa frazione eziologica della Toscana, Istituto per lo studio, la prevenzione e la rete oncologica, ISPRO, Firenze

3 Registro dei mesoteliomi maligni del Piemonte

4 Centro operativo regionale della Sicilia, Registro regionale siciliano dei mesoteliomi

5 Registro mesoteliomi della Regione Emilia-Romagna

6 Registro mesoteliomi Lombardia, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milano

7 Centro operativo regionale pugliese del Registro nazionale mesoteliomi

8 Centro operativo regionale Lazio

9 Dipartimento di Medicina, Epidemiologia, Igiene del Lavoro e Ambientale, INAIL, Roma

10 Centro operativo regionale dei mesoteliomi della Toscana, Istituto per lo studio, la prevenzione e la rete oncologica, ISPRO, Firenze

Introduzione

Il mesotelioma Maligno (MM) è un evento sentinella per la presenza di amianto. Valutare la storia lavorativa dei soggetti con MM aiuterebbe a identificare le ditte esposte e ricercare altri casi di patologie asbesto-correlate di origine professionale tramite metodo OCCAM.[1]
Il Registro Nazionale Mesoteliomi (RENAM) raccoglie tutti i casi di MM e il tipo di esposizione ad amianto. L'Istituto Nazionale di Previdenza Sociale (INPS) dispone di tutte le posizioni lavorative informatizzate dal 1974 in aziende con più di un dipendente. I due database contengono la ragione sociale (chiave) ma mentre INPS registra il reale nome delle aziende in RENAM è presente il nome riferito dall'intervistato.

Obiettivi

Creazione di un dizionario con codici azienda e ragione sociale delle ditte nelle quali si sono verificati casi di MM con un'esposizione ad amianto professionale.

Metodi

Sono stati acquisiti dall'archivio RENAM i periodi lavorativi con riconosciuta esposizione ad amianto (certa, probabile, possibile) e da INPS le ragioni sociali degli stessi soggetti. Si sono susseguite in modo iterato 3 fasi: i) analisi descrittiva della stringa ragione sociale- *text mining*, ii) pulizia del database e iii) *inexact matching*.
La ragione sociale nella fase i) è stata spezzata in subunità (parole) e si sono valutate le differenze di frequenze di parole tra RENAM e INPS e tra regioni. Similitudini tra parole sono state valutate tramite l'uso di *sounds-like operators* e funzioni di confronto.

Durante la fase ii) sono state rimosse le *stop words*, le parole di meno di due lettere e corretti alcuni termini (singolari-plurali, abbreviazioni, errori ortografici).
La fase iii) ha permesso di unire i due database usando la ragione sociale. L'*inexact matching* è avvenuto all'interno di ogni soggetto (almeno una parola uguale), tra tutti i record (almeno due parole uguali o una sola parola se stringhe di un solo elemento). Sono inoltre state separate le ragioni sociali non registrate da INPS (es agricoli, militari) e alcune aziende già note che presentavano più nomi (es FIAT).
I risultati ottenuti sono stati controllati in maniera manuale da esperti regionali.

Risultati

La procedura è stata creata per la regione Piemonte e validata/adattata sui dataset di Toscana, Emilia Romagna, Puglia, Sicilia e Lombardia; l'acquisizione dei dati per il Lazio è in corso d'opera. In tabella 1 sono riportati il numero di record e di casi (un soggetto può avere più di una esposizione lavorativa) per le varie regioni, e i risultati dell'*inexact matching*. Per il Piemonte la percentuale di successo della procedura è superiore al 57%, mentre le altre regioni hanno percentuali tra il 35 e il 50%.

Conclusioni

Risultati soddisfacenti si sono ottenuti per il Piemonte, per le altre regioni si ritengono necessari ulteriori approfondimenti. Questi risultati saranno a breve utilizzati per la lettura dei risultati di OCCAM [1].

Tabella 1. *Numero di righe e di casi di INPS e RENAM per ogni regione. Sono riportati N e % di inexact matching avvenuti rispetto a RENAM.*

Regione	INPS		RENAM		*INEXACT MATCHING*
	N righe	N casi	N righe	N casi	N righe (%righe)
Piemonte	4962	1998	2490	1104	1421 (57.07)
Toscana	2079	760	3466	906	1274 (36.76)
Emilia-Romagna	3589	1272	1367	891	591 (43.23)
Sicilia	2965	686	440	256	219 (49.77)
Lombardia	6734	2645	3965	2267	1897 (47.84)
Puglia	1626	527	440	335	204 (46.36)
Lazio	1295	464	496	301	--

Bibliografia

[1] Crosignani P, Massari S, Audisio R et al. The Italian surveillance system for occupational cancers: characteristics, initial results, and future prospects. Am J Ind Med, 2006 Sep; 49(9):791-8

THE INTEGRATION OF PRE-CODED INFORMATION WITH NARRATIVES DID NOT IMPROVE INJURY UNDERSTANDING: THE CASE OF ACCIDENTAL FALLS' SURVEILLANCE

Lorenzoni Giulia[1], Rampazzo Roberta[1], Buratin Alessia[1], Gregori Dario[1]

[1] Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, Vascular Sciences, and Public Health, University of Padova, Italy

Introduction

The surveillance of accidental falls in the hospital setting is crucial to better understand such phenomenon, which continues to represent a severe public health burden [1]. However, data collection systems often include both pre-coded information and narratives. Narratives would be a great source of additional information helping to better understand such phenomenon, but the manually revision of such information is unfeasible. Text mining would be a promising approach to exploit this information [2].

Objectives

The aim of the present study was to identify patterns of accidental falls integrating pre-coded information with free-text narratives using data from the incident reporting system of a Local Health Authority in the Northeast of Italy.

Methods

Each record reported detailed pre-coded information about patient and fall's characteristics, together with a narrative description of the fall. A topic model approach was employed to derive information from narratives, using a Latent Dirichlet Allocation (LDA) model. A Dirichlet prior was specified on the per-topic word distributions and on the per-document topic distribution. According to Griffiths and Styvers [3], the parameters were set equal to 0.1 and to 50 divided by the number of topics (50/k), respectively. The input data for the LDA was the document-term matrix built over the pre-processed narratives records, each row corresponding to a record and each column to a term. In addition, to reduce the sparsity of the document-term matrix, only terms that occurred in a minimum number of 5 documents were retained. Finally, we used the mean term inverse document frequency (TF-IDF) over the documents to reduce the sparsity of document-term matrix. TF-IDF allows for removing both less frequent and very common terms [4]. Multiple Correspondence Analysis (MCA) [5] was performed on pre-coded information and on both pre-coded and textual information all together. Finally, a *hierarchical classification on the principal components* (HCPC) [6] of the MCA was performed. Analyses were performed using R software [7].

Results

During the two-year period considered, 202 accidental falls were notified. In the MCA model without the topic variable, the first three dimensions explained the 19.72% of the overall data space, while with the topic variable the first three dimensions were found to explain the 17.4% of the overall data space.
Three clusters corresponding to three distinct profiles of unintentional falls were identified through the HCPC performed using only pre-coded information. The first one corresponded to 51 (27%) elimination-related falls. The second one grouped 139 (64%) unintentional falls occurred while the patient was in his hospital bedroom and was trying to get out from his hospital bed. The last cluster was the less represented (12, 8% of unintentional falls) corresponding to outpatients fell in the outpatient clinics without any fall risk with a resulting injury requiring a suture. HCPC with the topic variable provided overlapping results. Such results

were confirmed by the analysis of the within-cluster sum of square (WSS), which were 9.22 for the HCPC without the topic variable and 8.71 with the addition of the topic variable to the HCPC.

Conclusions

The HCPC analysis would be a powerful tool to identify patterns of accidental falls. However, information from narratives did not improve the identification of such patterns, suggesting that a data collection system including detailed pre-coded information would be adequate to monitor the phenomenon of accidental falls in such hospital setting.

References

[1] Morello RT, Barker AL, Watts JJ, et al., The extra resource burden of in-hospital falls: a cost of falls study. Med J Aust., 2015; 203:367

[2] Lucero RJ, Lindberg DS, Fehlberg EA, et al., A data-driven and practice-based approach to identify risk factors associated with hospital-acquired falls: Applying manual and semi-and fully-automated methods. Int J Med Inf., 2019; 122:63–9

[3] Griffiths TL, Steyvers M. Finding scientific topics. Proc Natl Acad Sci 2004; 101:5228–35.

[4] Leskovec J, Rajaraman A, Ullman JD. Mining of massive datasets. Cambridge university press 2014.

[5] Greenacre M, Blasius J. Multiple correspondence analysis and related methods. Chapman and Hall/CRC 2006

[6] Husson F, Josse J, Pages J. Principal component methods-hierarchical clustering-partitional clustering: why would we need to choose for visualizing data. Appl Math Dep 2010

[7] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing 2017. https://www.R-project.org/

ASSOCIATION BETWEEN RESPIRATORY DISEASES IN CHILDREN AND AIR POLLUTION PARTICULATE MATTER: A MULTI-POLLUTANTS APPROACH WITH RANDOM FOREST

Bottigliengo D.[1], Gallo E.[1], Bressan S.[2], Lanera C.[1], Lorenzoni G.[1], Ocagli H.[1], Berchialla P.[3], Baldi I.[1], Gregori D.[1]

[1] *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, and Vascular Sciences and Public Health, University of Padova, Padova, Italy*

[2] *Paediatric Emergency Unit, Department of Women's and Children's Health, University Hospital of Padova, Padova, Italy*

[3] *Department of Clinical and Biological Sciences, University of Torino, Torino, Italy*

Introduction

Respiratory diseases in children represent a considerable burden for public health [1]. Indeed, studying the potential causes of their occurrence is necessary to improve the quality of life in children and to reduce the associated costs for the health care system [2].
Exposure to air pollution is a known risk factor for the development of respiratory diseases among infants, such as asthma, bronchitis, pneumonia and chronic obstructive pulmonary disease [3,4]. Several studies have shown a positive association between pollutants, such as particulate matter, and risk of respiratory diseases using a single-pollutant modeling approach [5,6]. However, few studies addressed the issue from a multi-pollutants perspective [7–10]. Indeed, air pollution exposure is a complex mixture of different substances that can potentially impact on children's health. Thus, it can be of great interest studying how different pollutants can simultaneously affect the risk of developing a respiratory disease in subjects in the pediatric age and to examine their interactions [11].
Standard statistical approaches, such as regression models, have been mostly used to address such issues so far. Although they represent valuable tools to assess pollutant effect because of their interpretability, they have some limitations when dealing with highly correlated variables, a problem that is often encountered when several pollutants are evaluated at the same time. Machine Learning Techniques (MLTs) can provide an alternative solution to address the multi-pollutant issue. In fact, their flexibility may allow the detection of non-linearities and interactions among several variables that can be highly correlated to each other.

Objectives

The aims of the study are: (i) to assess if MLTs can provide more insights on the evaluation of the simultaneous effect of different types of particulate matter on respiratory diseases in children, (ii) to evaluate if MLTs can enhance the detection of respiratory diseases cases among infants. We believe that the use of such tools can help policy-makers during the implementation of strategies to prevent the occurrence of respiratory diseases in children.

Methods

In this study we use the data of the accesses to the pediatric emergency department of Padova, Italy, from 2007 to 2017. Emergency department visits for respiratory diseases were identified using the diagnosis provided by the physicians. Pollutants data were retrieved from the regional agency for environmental protection and prevention of Veneto (ARPAV). The daily concentrations levels of the following pollutants were collected from three monitoring stations near Padua: particulate matter with aerodynamic diameter less than 10 μm (PM10) and particulate matter with aerodynamic diameter less than 2.5 μm (PM2.5).

Moreover, daily levels of temperature, atmospheric pressure and humidity were also retrieved. Meteorogical data were linked to the pediatric emergency department visits using the closest monitoring station to the residence of the hospitalized subject.
Random Forest (RF) algorithm [12] was used to model the relationship between pollutants concentrations and the accesses for respiratory diseases. RF belongs to the family of tree-based methods, one of the most popular MLTs approaches used environmental data analysis [13]. A classification model was implemented to identify hospitalizations for respiratory diseases in contrast to accesses for other causes. In addition to pollutants exposures, the following variables were considered as additional predictors: daily mean concentrations of temperature, atmospheric pressure and humidity, the age and the sex of the hospitalized subject and the month of the year. Moreover, a different model accounting for lagged exposures to pollutants was also evaluated, considering simultaneously the exposure in the same day up to three days before. The models were tuned using data from 2007 to 2016 with 10-fold cross-validation to choose the optimal parameter, i.e. the number of variables to consider at each split when growing a tree in the forest. Downsampling of the minority class was implemented during model training to handle high class imbalance. The highest area under the ROC curve (AUC) was used as criterion to choose the best model. To assess the impact of the association between occurrence of respiratory diseases and pollutants concentrations, we used the total decrease in node impurity, measured by the Gini index, given by using the variable to perform the split and then averaging over all the trees. The trained algorithms were then used to recursively forecast the number of accesses for respiratory diseases on each day of 2017. That is, models were fitted subsequently adding a day of 2017 and they were used to predict the number of hospitalizations for respiratory diseases on the next day. Algorithms performances were evaluated by comparing the number of accesses for respiratory diseases observed and identified by the models using the median error between the daily predicted and observed count of accesses.

Results

A total of 293206 records were considered for the analysis. Among these, 12693 were identified as hospitalizations for respiratory disease. Median age, daily mean $PM_{2.5}$ and PM_{10} concentrations and daily mean temperature was higher for accesses for respiratory diseases, whereas daily mean pressure was similar between the groups.
Areas Under the ROC Curves (AUCs) on the training data are equal to 0.729 and 0.726 for model without and with lagged exposures to pollutants. Sensitivity values equal to 0.721 and 0.711 are observed in the compared models, whereas specificity is found to be 0.624 in the model without lagged exposures and 0.632 in the model with lagged exposure. Mean Gini index shows that age, daily mean temperature and $PM_{2.5}$ concentration were the three variables with highest association in the model without lagged exposure. The same pattern is also observed for the model with lagged exposures, with $PM_{2.5}$ variables showing a higher association than PM_{10} variables. In particular, the exposure 3 days before the access was found to be the most associated variables among the lagged ones.
The overall AUCs on the data from 2017 are equal to 0.769 and 0.750 for the model without and with the lagged variables, respectively. Sensitivity values equal to 0.770 and 0.749 are observed for the compared models, instead values equal to 0.654 and 0.666 are found for specificity. Comparison of observed and predicted number of hospitalizations for respiratory diseases in 2017 are graphically shown in Figure 1. Furthermore, the model without lagged variables shows a median error between the daily predicted and observed count of accesses equal to 24, whereas the error is 12 for the model with the lagged exposures.

Conclusions

In the present study, we evaluate simultaneously the association between different particulate matter exposures and hospitalizations for respiratory diseases in children using a machine learning algorithm.

Moreover, we also assess how MLTs can improve the detection of accesses for respiratory diseases. Regarding the predictive ability of the models, our results are in line with the findings of the previous studies whose aims were to estimate impact of pollutants on children's health outcomes [7], suggesting that signal recognition can be very challenging in such situations. The trained models show similar discrete performances, with higher sensitivity values denoting better ability in recognizing hospitalizations for respiratory diseases. The algorithms show higher predictive performances on the data from 2017. This is likely due to the recursive fitting of the algorithms, which subsequently used each day of 2017 to predict the number of accesses for the next day. Both models are prone to overestimate the number of hospitalizations for respiratory diseases, with the model with lagged variables showing a lower median error between the daily predicted and observed count of accesses. Nevertheless, the algorithms seem to be able to identify peaks of respiratory disease accesses, with higher predicted counts in the periods of the year known to have higher occurrence of such diseases.
Age and daily mean temperature are the variables with the strongest association with the outcome. PM2.5 is observed to have a stronger association than PM10 in both type of models. The lagged exposure model suggests that exposure 3 days before the access is the most associated among the lagged variables.
In summary, the developed MLT approaches show a discrete predictive ability and suggest that PM2.5 is more associated than PM10 with hospitalizations for respiratory diseases in children. Indeed, further research is needed to better evaluate the strength of pollutants' association and to better characterized how different pollutants interact with each other on the health outcome.

Bibliography

[1] World Health Organization, Inheriting a sustainable world?: atlas on children's health and the environment.; 2017; ISBN 978-92-4-151177-3.

[2] Landrigan P.J., Fuller R., Fisher S., et al., Pollution and children's health. Sci. Total Environ. 2019, 650, 2389–2394.

[3] Landrigan P.J., Fuller R., Acosta N.J.R., et al., The Lancet Commission on pollution and health. The Lancet 2018, 391, 462–512.

[4] Pavord I.D., Beasley R., Agusti A., et al., After asthma: redefining airways diseases. The Lancet 2018, 391, 350–400.

[5] Urman R., Eckel S., Deng H., et al., Risk effects of near-roadway pollutants and asthma status on bronchitic symptoms in children: Environ. Epidemiol. 2018, 2, e012.

[6] Ge E., Lai K., Xiao X., et al., Differential effects of size-specific particulate matter on emergency department visits for respiratory and cardiovascular diseases in Guangzhou, China. Environ. Pollut. 2018, 243, 336–345.

[7] Gass K., Klein M., Chang H.H., et al., Classification and regression trees for epidemiologic research: an air pollution example. Environ. Health 2014, 13.

[8] Sun Z., Tao Y., Li S., et al., Statistical strategies for constructing health risk models with multiple pollutants and their interactions: possible choices and comparisons. Environ. Health 2013, 12.

[9] Bobb J.F., Valeri L., Claus Henn B., et al., Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. Biostatistics 2015, 16, 493–508.

[10] Stingone J.A., Pandey O.P., Claudio L., et al., Using machine learning to identify air pollution exposure profiles associated with early cognitive skills among U.S. children. Environ. Pollut. 2017, 230, 730–740.

[11] Dominici F., Peng R.D., Barr C.D., et al., Protecting Human Health From Air Pollution: Shifting From a Single-pollutant to a Multipollutant Approach. Epidemiology 2010, 21, 187–194.

[12] Breiman L., Random Forests. Mach. Learn. 2001, 45, 5–32.

[13] Bellinger C., Mohomed Jabbar M.S., Zaïane O., et al., A systematic review of data mining and machine learning for air pollution epidemiology. BMC Public Health 2017, 17.

HANDLING MISSING DATA IN AIR POLLUTION STUDIES: A COMPARISON OF DIFFERENT APPROACHES BASED ON MULTIVARIATE TIME-SERIES MODELS

Gallo E.[1], Bottigliengo D.[1], Lanera C.[1], Lorenzoni G.[1], Ocagli H.[1], Zagolin L.[2], Marson G.[2], Berchialla P.[3], Baldi I.[1], Gregori D.[1]

[1] *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, and Vascular Sciences and Public Health, University of Padova, Padova, Italy*

[2] *Regional Agency for Environmental Prevention and Protection of the Veneto region (ARPAV)*

[3] *Department of Clinical and Biological Sciences, University of Torino, Torino, Italy*

Introduction

Exposure to air pollution is known to be associated with an increased risk of several diseases. In time-series air pollution studies, researchers often deal with data sets characterized by the existence of missing information that can occur for several reasons, such as interruptions or damages in the functioning of monitoring stations. When incomplete information on time-series data arise, single imputation methods, i.e. filling missing cells with one single value, is probably the most common approach among researchers. However, even if the imputation method is properly implemented, results from the analysis may be misleading because single imputation methods do not account for the sampling variability and the uncertainty that can arise if more than one model is used [1]. Thus, multiple imputation approaches have been proposed to replace missing values with two or more plausible values [2].
Multiple imputation approaches for time-series studies often assume that the data come from a multivariate normal distribution. Despite this assumption is only a naïve approximation to the true data distribution, many studies showed that this model works well in practice [3,4]. Moreover, the data can be transformed to make the normality assumption more plausible. A further assumption that must be made is that the data are missing at random (MAR), i.e. the pattern of missingness depends only on observed data.

Objectives

The aim of this study is to compare different multiple imputation approaches based on multivariate time-series models.

Methods

Environmental data collected by monitoring stations of the Regional Agency for Environmental Prevention and Protection of the Veneto region (ARPAV), Italy, from 2007 to 2018 were used for the analysis. Overall, the data from 4 monitoring stations were included in the analysis. The following information were retrieved: median daily values of the temperature in ^{o}C, relative humidity in percentage, atmospheric pressure in mbar, concentrations of particulate matter of less than 2.5 µm and 10 µm in aerodynamic diameter ($PM_{2.5}$ and PM_{10}) in µg/m^3, concentrations of nitrogen dioxide (NO_2) and ozone (O_3)in µg/m^3 and the total mm of rain fallen in a day.
Two imputation strategies were compared by specifying different relationships between the time and the data. In the first approach, the effect of time is modelled using flexible restricted cubic splines with different number of knots. In the second approach, the effect of time was modelled using lags and leads of the data. Moreover, the two approaches were also evaluated using ridge prior distributions on the covariance matrix of the model. Ridge prior information can be used to provide more stable imputation results, which are difficult to achieve if the data contain high proportion of missingness or if the variables are highly correlated.

Finally, we accounted for the clustered nature of the data by allowing for a different effect of the time in each monitoring station. The methods were compared both graphically, using density plots of observed and imputed data, and numerically, providing low-dimensional measures that assess the discrepancies between observed and imputed values.

Results

Overall, the distribution of imputed values obtained using the second approach were closer to the distribution of observed data. Regarding the first approach, the higher the number of knots the higher the plausibility of the imputed values. In both methods, the use of prior ridge led to more stable imputations, with less variability among imputed values.

Conclusions

The use of lags and leads in multivariate time-series models may provide imputed values that are closer to the observed ones. Furthermore, the specification of ridge prior distribution on the covariance matrix of the model decreases the variability of the results. These approaches can greatly improve the plausibility of imputed values in time-series studies and thus enhance the assessment of the effect of air pollution exposure on health outcomes.

Bibliography

[1] Hopke P.K., Liu C., Rubin D.B., Multiple Imputation for Multivariate Data with Missing and Below-Threshold Measurements: Time-Series Concentrations of Pollutants in the Arctic. *Biometrics* 2001, *57*, 22–33.

[2] Rubin D.B., *Multiple Imputation for Nonresponse in Surveys*; Wiley, 1987; ISBN 978-0-471-08705-2.

[3] Schafer J.L., *Analysis of Incomplete Multivariate Data*; Chapman and Hall/CRC, 1997;

[4] Schafer J.L., Olsen M.K., Multiple Imputation for Multivariate Missing-Data Problems: A Data Analyst's Perspective. *Multivar. Behav. Res.* 1998, *33*, 545–571.

DEVELOPMENT AND VALIDATION OF A NOVEL PRE-PREGNANCY SCORE PREDICTIVE OF PRETERM BIRTH IN PRIMIPAROUS WOMEN: A POPULATION-BASED COHORT STUDY

Merlo Ivan[1,2], Cantarutti Anna[1,2], Merlino Luca[3], Massi Michela Carlotta[1,4,5], Ieva Francesca[1,4,5], Corrao Giovanni[1,2]

[1] *National Centre for Healthcare Research & Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy*

[2] *Laboratory of Healthcare Research & Pharmacoepidemiology, Unit of Biostatistics, Epidemiology and Public Health, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy*

[3] *Operative Unit of Territorial Health Services, Region of Lombardy, Milan, Italy*

[4] *MOX – Modeling and Scientific Computing lab, Department of Mathematics, Politecnico di Milano, Milan, Italy*

[5] *CADS – Center for Analysis Decisions and Society, Human Technopole, Milan, Italy*

Introduction

Preterm birth (PTB), defined as any live birth occurring before 37 completed weeks of gestation, is the leading cause of perinatal morbidity and mortality, and second most common cause of death, after pneumonia, in children under 5 years of age [1, 2]. Moreover, preterm birth has also shown to increases the risk of medical and social disabilities in adulthood [3].
Worldwide, 11.1% of infants are born preterm every year; the rate is higher in low-income countries (11.8%), but the incidence is still significant in the most developed ones (5 - 9%) [4-6].
Multiple risk factors have been associated with preterm birth, including maternal age, socio-economic conditions, ethnic origin, multiple pregnancy, past reproductive history of preterm births, infections, inflammations and nutritional status [2, 6].
While many studies have evaluated the association between risk factors and preterm birth, there is a lack of data available on the prediction of preterm birth in relation to specific risk factors onset before pregnancy. Such evaluation has the potential to decrease the rate of preterm birth by facilitating monitoring and timely intervention based on individual women's risk factors. Moreover, that assessment could be useful to better understand the mechanisms that lead to preterm birth [7].

Objective

To develop and validate a novel comorbidity pre-pregnancy score predictive of preterm birth (preterm birth score - PTBS) in primiparous women using multiple source of information from the healthcare utilization databases of Lombardy region. The score intends to be calculable at the time of conception; therefore, it will not include information related to the gestational period. Only variables regarding the mother and collected in the four years preceding the beginning of gestation will be taken into account.

Methods

We conducted a population-based cohort study including 185,209 deliveries, with at least one infant born alive, that occurred between January 2012 and December 2017 in the Lombardy region of Italy. Only births from primiparous women with age at delivery between 15 and 55 and gestational age between 22 and 42 weeks were considered. A record linkage between regional healthcare utilization databases was conducted to trace the health and clinical history of the women in the four years prior to the start of gestation. The information available made it possible to outline the profile of the women respect to six areas: (i) drugs use,

(ii) exemption coverage, (iii) diagnoses received during hospitalization, (iv) inpatient interventions, (v) outpatient procedure and (vi) socio-demographic profile.
All information were evaluated according to their macro-categories (almost always defined by the first three digits of the specific code), in particular: the outpatients drug prescriptions were grouped based on the Anatomical Therapeutic Chemical (ATC) code, the diagnosed and procedures released from hospital were grouped based on the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM), the exemption based on their relative code, while the outpatient procedures using the relative branch code. Woman was considered exposed to one of these if, in the 4 years preceding the pregnancy, there was trace of at least one of the respective codes. With regard to the socio-demographics profile, we used the categorization retrieved from the Lombardy Certificates of Delivery Assistance (CeDAP) registry, which records data on maternal education level, marital status, employment status and gestational age (GA). In addition, information about the possible assisted medical conception was extracted from the CeDAP registry. A preterm birth was defined as a GA less than 37 weeks [2]. The cohort was randomly divided into training set (70%) and validation set (30%). With the aim of selecting conditions independently able to predict preterm birth, we proceed as follows. First, the least absolute shrinkage and selection operator (LASSO) method was applied, via logistic regression, on the training set to select the best predictors for the outcome. The choice of the LASSO tuning parameter was made using a 10-fold cross-validation. Second, the selected variables were used to build the PTBS. The coefficients estimated from the logistic LASSO model were used for assigning a score at each selected covariate. In particular, a score from 1 to 15 was assigned at each predictor. The total score was calculated as the sum of the individual variables scores.
The predicted absolute risk of preterm birth, given the total value of the score, was reported using a nomogram. The ROC curve of the PTBS and the respective AUC were calculated for both the training and the validation set. The predictive accuracy of the PTBS was furthermore investigated using the Net Reclassification Improvement (NRI) index, considering categories of predicted probability [8]. Three predicted risk categories were defined based on the average incidence of preterm birth and considering a clinical rationale. The NRI combines the net proportions of events and non-events correctly reclassified by the PTBS by evaluating the predicted probability among the women who switched from one risk category to another.

Results

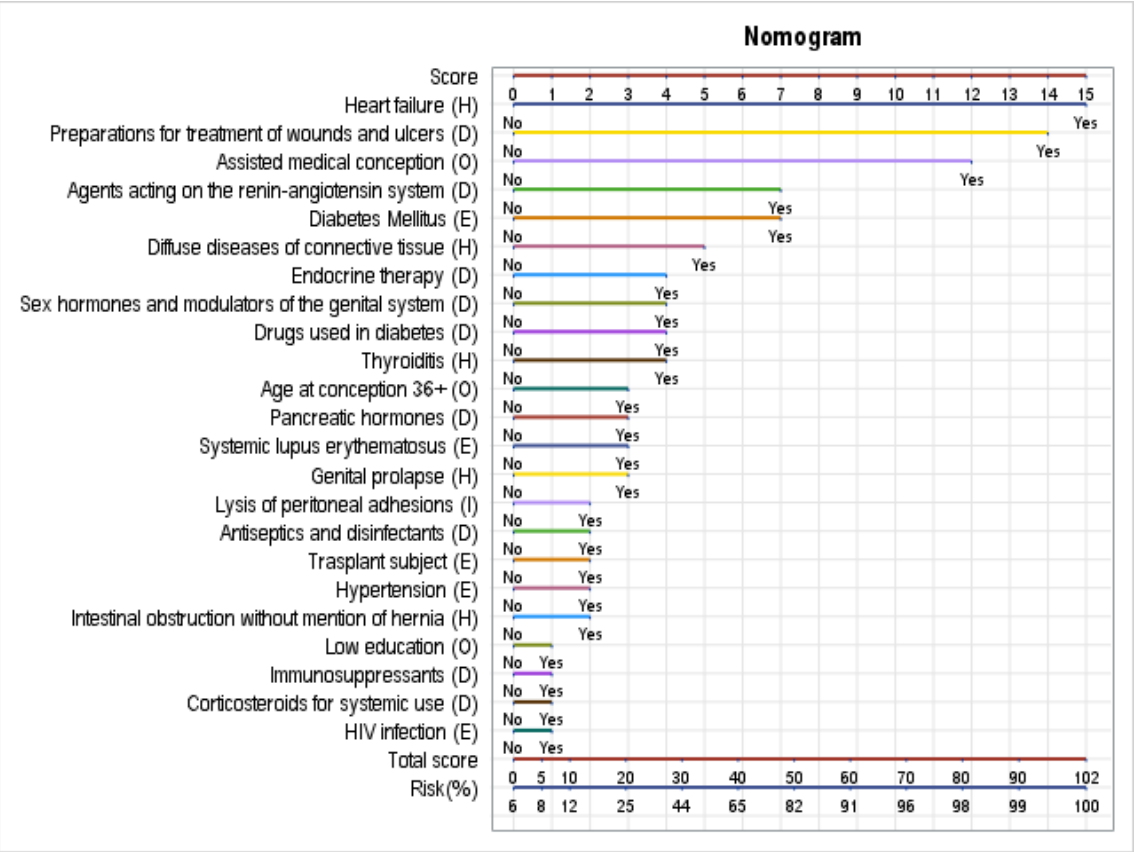
Of the 185,209 deliveries selected, 129.647 were used as training set and 55.562 as validation set. The percentage of preterm births in the training set was 7.4%.
A total of 1,767 macro-categories were found among the first five areas considered (73 drugs, 52 exemptions, 47 outpatient procedures, 862 diagnoses and 733 interventions during hospitalizations). The variables extracted from CeDAP registry were 5, including socio-demographic features and information regarding the type of conception.
The variables, which were included in the score, along with the points associated to each of them, are reported in **Figure 1**. Moreover, the nomogram shows the predicted absolute risk based on the total PTBS value. The preterm birth risk increases from to 5.5%, 8.3%, 12.3%, 25.0% and 44.2% in women having a total score of 0, 5, 10, 20 and 30, respectively. In the validation set, the following observed incidences were found: 6.1% for the women with score between 0 and 4; 10.2% for the women with score between 5 and 9; 18.5% for the women with score between 10 and 19; 24.4% for the women with score between 20 and 29; and 44.4% for the women with score equal or greater to 30.
The AUC value of the PTBS was 61% for both the training and validation set. The risk categories chosen for the NRI computation was 0 to 6%, 6 to 15% and 15% or more. The NRI value was 0.27 (44% of non-events correctly reclassified and 17% of events incorrectly reclassified) and 0.25 (44% of non-events

correctly reclassified and 19% of events incorrectly reclassified), respectively for the training and validation set.

Conclusions

Our study shows that the PTBS is able to stratify women according their absolute preterm birth risk despite its poor prediction accuracy. The selection mechanism allowed to highlight some variables significantly associated with preterm birth. However, the causal relationship between the variables and the outcome was not taken into account as well as the possible interactions between the different types of exposure.
Given the huge number of information, it is possible that other methodologies, as the machine learning techniques, will be more accurate, cause their better adaption to such situations, and could restrict the problems derived from variables selection methods.

***Figure 1**. Nomogram: the figure shows the variables included in the score. For each variable, a score from 1 to 15 was assigned. The maximum total score achievable, given by the sum of the variables individual score, was equal to 102. The nomogram gives, for every possible values of the total score, the predicted absolute risk.*



E: exemptions; **D**: drugs; **H**: diagnoses received during hospitalization; **I**: inpatient interventions; **O**: other information taken from the CeDAP registry.

References

[1] Phillips C., Velji Z., Hanly C. et al., Risk of recurrent spontaneous preterm birth: a systematic review and meta-analysis. BMJ Open, 2017; 7(6): e015402.

[2] Slattery M.M., Morrison J.J., Preterm delivery. The Lancet., 2002; 360: 1489-97.

[3] Moster D., Lie R.T., Markestad T., Long-term medical and social consequences of preterm birth. NEJM., 2008; 359(3): 262-73.

[4] Purisch S.E., Gyamfi-Bannerman C., Epidemiology of preterm birth. Seminars in Perinatology., 2017; 41(7): 387-391.

[5] Tucker J., McGuire W., Epidemiology of preterm birth. BMJ., 2004; 329(7467): 675–678.

[6] Goldenberg R.L., Culhane J.F., Iams J.D. et al., Epidemiology and causes of preterm birth. The Lancet., 2008; 371(9606): 75-84.

[7] Goldenberg R.L., Goepfert A.R., Ramsey P.S., Biochemical markers for the prediction of preterm birth. AJOG., 2005; 192(5): S36-S46.

[8] Pencina M.J., D'Agostino R.B.Sr., D'Agostino R.B.Jr. et al., Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Statist. Med., 2008; 27:157-17.

LA QUALITÀ DEI PERCORSI DI CURA NEI PAZIENTI CON DISTURBI MENTALI GRAVI IN ITALIA. Il Progetto QUADIM

Monzio Compagnoni Matteo[1,2], Franchi Matteo[1,2], Merlino Luca[3], Saponaro Alessio[4], Scondotto Salvatore[1,5], Gaddini Andrea[6], Di Fiandra Teresa[7], Lora Antonio[1,8], Corrao Giovanni[1,2]

[1] *Interuniversity Centre of Healthcare Research & Pharmacoepidemiology, located at the Laboratory of Healthcare Research & Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy*

[2] *Department of Statistics and Quantitative Methods, Division of Biostatistics, Epidemiology and Public Health, University of Milano-Bicocca, Milan, Italy*

[3] *Head, Epidemiologic Observatory, Regional Welfare Service, Lombardy Region, Milan, Italy*

[4] *Authority for Healthcare and Welfare, Emilia Romagna Regional Health Service, Bologna, Italy*

[5] *Epidemiologic Observatory, Sicily Regional Health Service, Palermo, Italy*

[6] *Roma 2 Local Health Authority, Lazio region, Rome, Italy*

[7] *Department of Health Planning, Italian Health Ministry, Rome, Italy*

[8] *Department of Mental Health, ASST di Lecco, Lecco, Italy*

Introduzione

In salute mentale la qualità delle cure erogate molto spesso non corrisponde agli standard previsti, e frequentemente i pazienti con disturbo mentale grave non beneficiano di percorsi di cura consistenti con le raccomandazioni basate sulle evidenze scientifiche in questo ambito [1-3]. Inoltre, è presente una elevata variabilità della qualità dei percorsi di cura erogati sia tra Regioni che, all'interno delle stesse, tra i vari Dipartimenti di Salute Mentale (DSM). Il progetto QUADIM (QUAlità dei percorsi di cura nei pazienti con Disturbo Mentali gravi), un ampio studio basato su dati real-world finanziato nel 2016 dal Ministero della Salute attraverso i programmi CCM, si inserisce dunque in questo contesto.

Obiettivo

L'obiettivo del progetto QUADIM è dunque quello di valutare la qualità dei percorsi di cura erogata ai pazienti con disturbo mentale grave utilizzando un set di indicatori di processo messo a punto da un panel di esperti (e approvati dalla Conferenza Stato Regioni nel 2014 [4]), in 4 Regioni italiane. Un altro obiettivo dello studio consiste inoltre nella validazione di tali indicatori di processo.

Metodi

Il progetto si basa sull'utilizzo degli archivi sanitari amministrativi delle 4 Regioni italiane partecipanti (Lombardia, Emilia-Roma, Lazio e Sicilia), che complessivamente coprono più di 25 milioni di cittadini assistiti dal Servizio Sanitario Nazionale (SSN), circa il 40% dell'intera popolazione italiana.
Tramite il flusso SISM (Sistemi Informativi per il monitoraggio e la tutela della Salute Mentale) sono stati identificati tutti gli individui beneficiari del SSN che nel biennio 2015-2016 avevano età superiore ai 18 anni, erano residenti in una delle Regioni partecipanti e avevano sperimentato almeno un contatto con le strutture dei DSM regionali, ricevendo (o avendo ricevuto) una diagnosi per almeno uno dei 4 disturbi mentali gravi indagati (i.e., schizofrenico, depressivo, bipolare o della personalità). Per ognuna delle 4 patologie indagate, sono state individuate 4 coorti distinte, composte rispettivamente dai pazienti: (1) con disturbo prevalente;

(2) con disturbo incidente; (3) con disturbo all'esordio (età compresa tra i 18 e i 25 anni); e (4) dimessi da un ricovero ospedaliero nel Servizio Psichiatrico di Diagnosi e Cura (SPDC).
Per ogni patologia, attraverso una procedura di record-linkage tra i vari flussi sanitari amministrativi (SISM, Schede di Dimissione Ospedaliera, Specialistica ambulatoriale, Farmaceutica territoriale, File F e Registro delle cause di morte) è stato possibile monitorare i percorsi di cura erogati ai pazienti componenti le 4 coorti.
Per quanto riguarda l'obiettivo primario del progetto, e cioè per la valutazione della qualità dei percorsi di cura, è stato definito e calcolato un set di indicatori clinici (12 di processo, 4 di esito). Tali indicatori sono classificati in base alla (i) unità di osservazione (Paziente, Famigliare, Paziente e famigliare); (ii) tipo di servizio (assistenza territoriale, residenziale, semiresidenziale, ospedaliera e qualsiasi contatto con il DSM) o (iii) contenuto della prestazione (trattamento farmacologico, psicoeducativo, psicoterapico); (iv) coorte alla quale l'indicatore deve essere applicato (prevalente, incidente, incidente con disturbo all'esordio, dei pazienti dimessi da SPDC), e in funzione anche della (v) modalità di calcolo dell'indicatore
Successivamente, sulla sola coorte dei pazienti con disturbo incidente, è stata effettuata la validazione degli indicatori di processo, valutando quanto un appropriato percorso di cura possa aiutare a prevenire un riacutizzarsi della sintomatologia. La relazione tra gli indicatori di processo e di esito clinicamente più rilevanti è stata valutata applicando uno studio case-crossover: su un campione costituito da tutti i pazienti con disturbo incidente che, durante l'anno di osservazione successivo alla data di ingresso nello studio, hanno sperimentato almeno una ospedalizzazione in SPDC, si è andati a stimare l'effetto dell'esposizione corrente sull'outcome considerato. Come esposizione è stata considerata l'erogazione di un percorso di cura appropriato (appropriatezza valutata secondo le due dimensioni di: continuità dell'assistenza territoriale e aderenza al trattamento farmacologico), mentre l'outcome era l'ospedalizzazione in SPDC. Per ogni caso si è andati a valutare l'esposizione nella finestra temporale corrente (immediatamente precedente l'evento) con quella delle 3 finestre temporali precedenti di controllo. Ogni finestra temporale aveva ampiezza pari a 45 giorni. Un modello di regressione logistica condizionata, con appaiamento 1 a 3, è stato implementato per stimare l'odds ratio (OR), e relativi intervalli di confidenza al 95% (95% IC), dell'outcome per l'esposizione ad un adeguato livello di assistenza territoriale e aderenza a trattamento farmacologico. Le stime sono state aggiustate per utilizzo di altri trattamenti farmacologici durante le finestre di riferimento e di controllo. Diverse analisi di sensibilità sono state svolte, facendo variare l'ampiezza temporale delle finestre di riferimento e di controllo, oltre che le soglie per definire una aderenza al trattamento farmacologico adeguata.
La validazione degli indicatori di processo è stata effettuata, momentaneamente, solamente per la regione Lombardia, mentre è in corso di svolgimento nelle altre Regioni.

Risultati

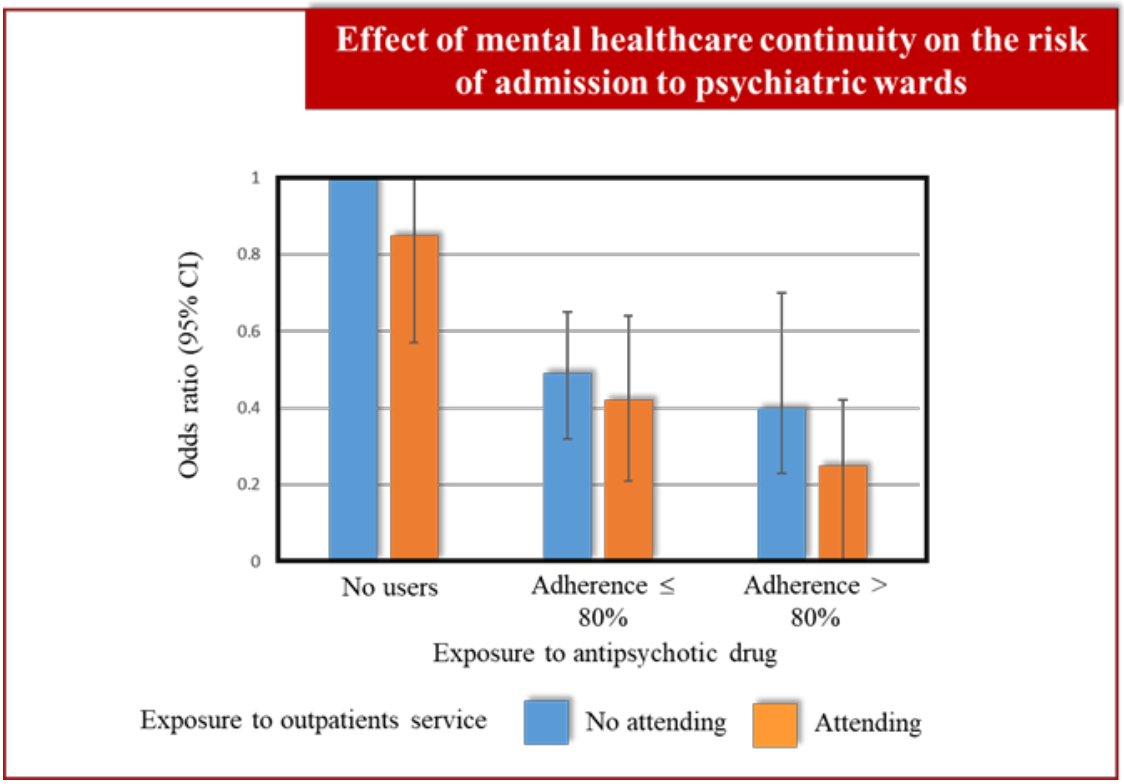
Complessivamente, il progetto QUADIM ha permesso di valutare i percorsi di cura erogati a circa 248,000 pazienti prevalenti con disturbo mentale grave (77,000 schizofrenico, 100,000 depressivo, 34,000 bipolare e circa 37,000 grave della personalità).
Per brevità verranno ora riportati i risultati riguardanti solamente i pazienti con disturbo schizofrenico.
Nel biennio 2016-2017, i pazienti con disturbo schizofrenico nelle 4 regioni considerate sono circa 77,000 per quanto riguarda la coorte prevalente, mentre la coorte dei pazienti con disturbo incidente, all'esordio e dimessi da SPDC hanno rispettivamente numerosità pari a 8,000, 1,000 e 12,000.
I pazienti con disturbo prevalente e quelli dimessi da SPDC godono di una maggiore assistenza territoriale rispetto ai pazienti giovani appena presi in carico. Nella fase di accesso ai servizi dei DSM molto raramente vengono effettuate delle valutazioni standardizzate dei pazienti, mentre l'accesso agli interventi psicosociali (i.e., sedute di psicoterapia, psicoeducazione, interventi rivolti ai familiari) risulta soddisfacente, anche se abbastanza ridotto. La continuità del percorso di cura risulta assicurata maggiormente nella coorte dei casi prevalenti. A circa la metà dei pazienti dimessi da SPDC viene erogato un trattamento in CSM o una visita

domiciliare entro le due settimane successive alla dimissione indice. L'aderenza al trattamento farmacologico con antipsicotici (indicato per i pazienti con disturbo schizofrenico) risulta essere ridotta: due pazienti su tre discontinuano il trattamento entro l'anno. Per i controlli clinici per verificare la sicurezza del trattamento farmacologico, mediamente soltanto 3 pazienti su 10 ricevono un adeguato numero di controlli entro l'anno. Considerando gli indicatori di esito, il tasso standardizzato di mortalità (SMR) è in media due-tre volte maggiore rispetto alla popolazione generale, arrivando ad assumere valori compresi tra 4 e 6 nel caso della coorte dei pazienti dimessi da SPDC. Il tasso di ospedalizzazioni in SPDC risulta invece essere pari a 180, 234 e 847 ospedalizzazioni in SPDC per 100,000 anni-persona (AP) accumulati, rispettivamente per la coorte dei pazienti prevalenti, incidenti e dimessi da SPDC. I risultati presentano una discreta eterogeneità tra le diverse Regioni (e.g., in Sicilia per quasi la totalità degli indicatori considerati, si osservano valori più bassi rispetto alle altre Regioni).
Per la conduzione dello studio *case-crossover* sono stati arruolati circa 890 pazienti lombardi con disturbo schizofrenico incidente che, durante il primo anno di osservazione, hanno sperimentato una ospedalizzazione in SPDC. Sia l'essere assistiti continuativamente a livello territoriale, che essere molto aderenti al trattamento farmacologico con antipsicotici. Rispetto ai pazienti non assistiti a livello territoriale e non in trattamento farmacologico, si osserva una riduzione del rischio pari al 70% per quei pazienti con alta aderenza e con assistenza territoriale continuativa, riduzione del rischio che risulta essere del 60% per i pazienti con alta aderenza agli antipsicotici e poco seguiti a livello territoriale.
I risultati delle analisi di sensibilità, effettuate sullo studio di validazione degli indicatori di processo, hanno confermato i risultati ottenuti nelle analisi principali.

Conclusioni

Il progetto QUADIM ha dimostrato come oggi sia possibile utilizzare i real-world data, derivati dagli archivi sanitari amministrativi, per ricavarne indicatori al fine di valutare la qualità dei percorsi di cura anche nell'ambito della salute mentale, andando a identificare criticità e priorità di intervento.

***Figura 1**. Odds ratio (OR), e relativi intervalli di confidenza al 95% (95% IC), dell'outcome (ospedalizzazione in SPDC) per l'esposizione ad un adeguato livello di assistenza territoriale e aderenza a trattamento farmacologico, negli 890 pazienti lombardi con disturbo schizofrenico incidente nel biennio 2015-2016 che hanno sperimentato almeno una ospedalizzazione in SPDC (disegno case-crossover).*



Bibliografia

[1] Lora A, Conti V, Leoni O, Rivolta AL. Adequacy of treatment for patients with schizophrenia spectrum disorders and affective disorders in Lombardy, Italy. Psychiatric Services, 2011;62(9):1079-84

[2] Lora A, Monzani E, Ibrahim B, et al. Routine quality care assessment of schizophrenic disorders using information systems International Journal for Quality in Health Care, 2016; 1–6

[3] Lora A, Cosentino U, Gandini A, Zocchetti C. Which community care for patients with schizophrenic disorders? Packages of Care provided by Departments of Mental Health in Lombardy (Italy). Epidemiologia e Psichiatria Sociale, 2007;16,4:330-238

[4] Conferenza Unificata Stato-Regioni: "Definizione dei percorsi di cura da attivare nei Dipartimenti di Salute Mentale per i disturbi schizofrenici, i disturbi dell'umore e i disturbi gravi di personalità" (13 novembre 2014)

VISUAL FUNCTION AND BRAIN LESIONS IN PRETERM NEWBORNS: VALIDATION OF THE *NEONATAL ASSESSMENT VISUAL EUROPEAN GRID* (NAVEG)

Gnesi Marco[1], Malerba Laura[2], Rossi Andrea[3], Strobio Caterina[2], Fazzi Elisa[2,3], Montomoli Cristina[1]

[1] *Section of Biostatistics and Clinical Epidemiology, Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy*

[2] *Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy*

[3] *Child and Adolescent Neurology and Psychiatry Unit, Children Hospital, ASST Spedali Civili, Brescia, Italy*

Introduction

Visual function is crucial for children's early development and many authors consider it as a window on the brain [1]. Visual function is an expression of the maturation of the Central Nervous System (CNS), so that its monitoring may inform about CNS functions and their adaptive evolution following a lesion [2]. The *Neonatal Assessment Visual European Grid*, or NAVEG, was developed for the assessment of visual function in newborns. The tool was refined in a previous pilot study [3]. The grid contains the following 17 items:

- *Pupillary Light Reflex* (coded *Present/Slow/Absent*);
- *Fixation – Bull's Eye target* (coded *Present/Difficult/Absent*);
- *Fixation – Face Figure target* (coded as above);
- *Fixation – Human Face* (coded as above);
- *Horizontal Smooth Pursuit, left* (coded as above);
- *Horizontal Smooth Pursuit, right* (coded as above);
- *Vertical Smooth Pursuit, down* (coded as above);
- *Vertical Smooth Pursuit, up* (coded as above);
- *Saccadic Movements, from right* (coded as above);
- *Saccadic Movements, from left* (coded as above);
- *Contrast Sensitivity* (4 levels);
- *Visual Acuity* (7 levels);
- *Visual Field, right* (coded *Typical/Limited*);
- *Visual Field, left* (coded as above);
- *Visual Field, up* (coded as above);
- *Visual Field, down* (coded as above);
- *Fundus Oculi* (optional item, coded *Normal/Abnormal*).

Aim

The aim of the present study was to assess construct validity of NAVEG and concurrent validity brain lesions (US scanning) and Amiel Tison's Neurological Examination at Term (ENTAT).

Methods

Preterm newborns (< 34 weeks Gestational Age), were consecutively enrolled at the NICU of Spedali Civili Hospital (Brescia, Italy). NAVEG was administered at a Corrected Age of 38-42 weeks; this choice was aimed at focusing on impairment in visual function due to brain lesions rather than on the (typical) evolution of visual function due to CNS development. Newborns also underwent US scanning for detecting brain

lesions (non-transient hyperechogenicity, periventricular lesions, intraventricular haemorrhage, parenchymal lesions) and ENTAT for functional impairment (impairment: abnormal in at least two AT items). Anamnestic variables were described according to their nature as means and SDs, medians and IQRs, or frequency distributions. Items were described with their absolute and relative frequency distributions, and then their association with US scans and ENTAT was tested by means of Pearson's Chi-squared test or Fisher's exact test.
Construct validity was inspected by means of Exploratory Factor Analysis (EFA). Because uniquenesses were above 0.30 for almost all items, the Iterated Principal Factors method was used for factors extraction. The number of factors to be retained was determined according to Horn's Parallel Analysis, in which the eigenvalues obtained from the actual correlation matrix are compared to those generated by uncorrelated normal variables in a Monte Carlo simulation process, or by Cattel's Scree plot criterion and Kaiser's criterion [4]. For the sake of interpretation, factors were then rotated using the Quartimax method, which simplifies the structure of items (i.e. it reduces the number of factors needed to "explain" a single item).
Multiple Correspondence Analysis (MCA) was used to inspect the relationship between NAVEG's items and US scanning ('normal' vs. 'pathological lesions') or ENTAT ('normal' or 'impaired'). Both the aforementioned explanatory variables were included as supplementary variables, i.e. passive variables that are represented in the axes after the dimensions have been determined on the other (active) variables. The dimensions were determined by using Burt's method. The variables were represented in principal coordinates.
Finally, after the factorial structure of NAVEG was identified, Cronbach's alpha was computed as an index of internal consistency, in spite of criticisms raised against its use [5].

Results

A total of 111 preterm newborns were enrolled, of which 54 (49.5%) females. 34 (31.2%) had evidence of brain lesions on US scanning, while AT's examination disclosed an impairment in 61 (55.0%). Notably, US scanning and AT's examination showed a low inter-rater agreement. Mean gestational age at birth was 31.5±1.9 weeks in newborns with no evidence of brain lesions, and 30.4±2.0 weeks in those with brain lesions.
NAVEG was administered at a mean Corrected Age of 39.9±1.2 weeks (range 38-42 weeks), with no significant difference between those with and without brain lesions (MW=0.220, p=.83). The occurrence of missing observation was negligible for all items except *Fundus Oculi*, an optional item, which was excluded from further analyses; for some other items, categories of moderate or strong impairment were underrepresented (*Pupillary Light Reflex*, *Fixation Bull's – Eye* and *Face Figure*, *Smooth Pursuit – left* and *up*, *Saccadic Movements – from right* and *from left*, *Contrast Sensitivity*, *Visual Acuity*, *Visual Field – left*). Almost all items were associated to US scanning, with the exception of *Contrast Sensitivity*, *Visual Acuity*, and *Visual Field – up*. Similarly, almost all items are associated to ENTAT, except for *Fixation – Face Figure*. All the items (excluding *Fundus Oculi*) were included in an EFA; two factors were extracted. After rotation, the first factor explained 40.8% of observed variance and the second 8.9%; the first factor saturated almost all items (see Figure), except for *Contrast Sensitivity* and *Visual Field – down* (equally saturated by the first two factors) and for *Visual Acuity* (saturated only by the second factor). Moreover, *Fixation – Bull's Eye* was saturated only by the first factor, but its saturation was lower when compared to other items. Uniquenesses were found to be from moderate to high.
In MCA, the first dimension explained 74.9% of total inertia (i.e. observed variability). Two archetypical profiles, separated on the first dimension, were disclosed: one of impairment and the other of optimality on NAVEG items. Both US scanning and ENTAT, included as supplementary variables, were related to those profiles. In addition, the MCA solution advised about the chance of dichotomising all the items as 'impaired' vs. 'optimal', even when the original ordinal scores had three or more levels.

Thus, all the items were dichotomised, their categories being assigned to impairment or optimality depending on how they located in the previous MCA, and analyses were repeated. Dichotomisation helped in reducing the number of items with underrepresented categories, which were limited to *Fixation Bull's – Eye* and *Face Figure*, and *Visual Field – left*. Tests of association of the dichotomised items to US scanning confirmed what was found with the original items.
EFA (on the matrix of tetrachoric correlations of the dichotomised items) still identified two factors, respectively explaining 64.3% and 8.0% of observed variance (see Figure). In comparison with the previous solution, in this case saturations were higher and uniquenesses lower. Only *Visual Acuity* still had a stronger saturation on the second factor, but was now saturated also by the first factor. MCA results were consistent with the archetypical profiles of optimality and impairment, the first dimension explaining 93.5% of inertia.
Cronbach's alpha was computed for the dichotomised items (excluding *Visual Acuity*), showing a satisfactory internal consistency of items (α=0.91). Item-test correlations were from moderate to high.
Finally, a "NAVEG score" was computed by summing the dichotomised items (1 if the item was impaired, 0 otherwise). Such score was significantly higher among newborns with brain lesions compared to those without brain lesion (MW=3.70, p=.0002) and among newborns with impairment on ENTAT compared to those without impairment (MW=5.62, p<.0001).

Figure. *Results of Exploratory Factor Analysis on the original items and on the dichotomised items. Factors rotation method: Quartimax. For each factor, explained variance is reported in brackets. Saturations $|\lambda|<0.30$ are omitted.*

	EFA on original items		**EFA on dichotomised items**	
	Factor 1 (40.8%)	Factor 2 (8.9%)	Factor 1 (64.3%)	Factor 2 (8.0%)
	λ	λ	λ	λ
Pupillary Light Reflex	0.602	-	0.741	-
Fixation – Bull's Eye target	0.414	-	0.649	-
Fixation – Face Figure target	0.621	-	0.848	-
Fixation – Human Face	0.589	0.305	0.761	-
Horizontal Smooth Pursuit, left	0.593	-	0.748	-
Horizontal Smooth Pursuit, right	0.711	-	0.904	-
Vertical Smooth Pursuit, down	0.654	0.539	0.864	-
Vertical Smooth Pursuit, up	0.777	-	0.858	-
Saccadic Movements, from right	0.769	-	0.862	-
Saccadic Movements, from left	0.664	-	0.832	-
Contrast Sensitivity	0.414	0.398	0.651	0.420
Visual Acuity	0.301	0.581	0.509	0.670
Visual Field, right	0.789	-0.319	0.868	-0.461
Visual Field, left	0.788	-	0.931	-0.307
Visual Field, up	0.767	-	0.935	-
Visual Field, down	0.4989	0.461	0.747	-

Conclusions

Results of the present study are consistent with those of the pilot study [3] and support construct validity of NAVEG. Moreover, they provide evidence regarding the relation between visual function, brain lesions and neurological examination, thus backing up concurrent validity. It is also shown how a dichotomous scoring for all the items could "perform" better. In the future, further research will be needed to demonstrate robust criteria for the computation of a "NAVEG score", and to assess how visual function measured through NAVEG evolves with CNS development.

Bibliografia

[1] Hubel D. H., Wiesel T. N. Brain mechanisms of vision. Scientific American, 1979; 241(3):150–162

[2] Fazzi E., Signorini S.G., Bova S.M., et al. Spectrum of visual disorders in children with cerebral visual impairment. Journal of Child Neurology, 2007 Mar; 22(3):294-301

[3] Rossi A., Gnesi M., Montomoli C., et al. Neonatal Assessment Visual European Grid (NAVEG): unveiling neurological risk. Infant Behavior and Development, 2017; 49:21–30

[4] Ledesma R.D., Valero-Mora P. Determining the Number of Factors to Retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis. Practical Assessment, Research & Evaluation, 2007 Feb; 12(2)

[5] Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's Alpha. Psychometrika, 2009 Mar; 74(1):107-120

ASSOCIAZIONE TRA ESPOSIZIONE AD ANTIDEPRESSIVI / DEPRESSIONE IN GRAVIDANZA ED OUTCOMES NEONATALI: "UMBRELLA REVIEW" DI REVISIONI SISTEMATICHE E META-ANALISI

Biffi Annalisa[1,2], Cantarutti Anna[1,2], Corrao Giovanni[1,2]

1 Centro di Ricerca Interuniversitario Healthcare Research & Pharmacoepidemiology

2 Dipartimento di Statistica e Metodi Quantitativi, Sezione di Biostatistica, Epidemiologia e Sanità Pubblica, Università di Milano-Bicocca

Introduzione

La sindrome depressiva rappresenta un grave problema di salute pubblica. Si stima che nel 2020 questa sarà la patologia più diffusa tra le malattie mentali e in generale la seconda malattia più diffusa dopo le patologie cardiovascolari [1]. Oggigiorno, i valori percentuali più elevati si registrano nelle donne, soprattutto nel periodo durante e/o post- gravidanza [2]. Tra queste, circa il 18% registra una diagnosi di depressione [3] e il 4-10% fa uso di almeno un antidepressivo (AD) durante la gravidanza [4]. In generale, gli AD si classificano in tre categorie: SSRI (inibitori selettivi della ricaptazione della serotonina), TCA (AD triciclici), e AA (AD atipici).
L'insorgenza della sindrome depressiva in/post gravidanza sembra essere collegata a diversi fattori di rischio, quali i) precedenti malattie mentali, ii) mancanza di supporto sociale e finanziario [5], e iii) diverse complicanze insorte durante la gravidanza [6]. Per questo, una precoce diagnosi di depressione è fondamentale per elaborare una adeguata terapia utile ad evitare il peggioramento dei sintomi depressivi [7] e così prevenire i possibili eventi avversi neonatali [8], come il normale sviluppo psicomotorio del neonato [9]. Sono state inoltre avanzate terapie alternative, quali la terapia cognitiva comportamentale [10] e/o cambiamenti dello stile di vita [11] così da migliorare lo stato di salute mentale delle donne in gravidanza.

Obiettivi

Diverse revisioni sistematiche e meta-analisi sono state pubblicate circa la relazione tra depressione/uso di AD in gravidanza e l'insorgenza di eventi avversi neonatali. Per fornire una sintesi della letteratura disponibile, è stata realizzata una "umbrella review" finalizzata a raccogliere le revisioni sistematiche e meta-analisi di studi osservazionali.

Metodi

Ricerca in letteratura. È stata effettuata una ricerca sistematica su MEDLINE fino a settembre 2018. Sono stati selezionati tutti gli studi che analizzavano l'associazione tra esposizione a sindrome depressiva/AD in gravidanza ed esiti neonatali. Se una specifica relazione è stata approfondita da più di una meta-analisi, è stata inclusa quella più recente e con il maggior numero di studi. Quando possibile, è stata effettuata una integrazione degli studi osservazionali riportati dalle diverse meta-analisi che si focalizzavano su un determinato topic, quale nascita pretermine, basso peso alla nascita, disturbi del neurosviluppo, sindrome da scarso adattamento neonatale, problemi respiratori, tremori, malformazioni, aborto spontaneo, ipertensione polmonare persistente, alterazioni di diversi sistemi fisiologici (come nervoso, urogenitale, muscolo-scheletrico). Inoltre, sono stati valutati l'essere piccolo per l'età gestazionale, i rapporti con l'unità di terapia intensiva neonatale, e l'iniziazione all'allattamento.
La sindrome depressiva è stata definita sulla base del Sistema di Classificazione Internazionale, del Manuale Diagnostico e Statistico dei Disturbi Mentali, o su di uno strumento di screening o scala specifica.
L'esposizione di interesse poteva riguardare (i) la sindrome depressiva in gravidanza, in questo caso tutte le utilizzatrici di AD con diagnosi di depressione sono state escluse dall'analisi, e (ii) l'utilizzo di farmaci AD

durante la gravidanza, in questo caso, quando specificato, sono state riportate le stime specifiche per tipo di farmaco assunto e trimestre di esposizione.
Analisi statistica. Per ogni meta-analisi, è stata riportata la stima pooled (rischio relativo o odds ratio) ed il corrispondente intervallo di confidenza 95% (IC) ottenuti con il modello a effetti fissi e casuali [12]. È stato stimato un intervallo di previsione al 95% che considera l'eterogeneità tra gli studi e l'incertezza generata se un nuovo studio andasse ad investigare la stessa associazione [13,14]. Inoltre, l'eterogeneità tra gli studi è stata misurata con l'indice di inconsistenza (0-100%) [15]. Il publication bias è stato valutato con il test di Egger ($p < 0.10$) [16] ed è stato applicato un test per determinare l'eccesso di significatività [17] che compara i risultati osservati (O) rispetto agli attesi (A) con un livello alfa pari allo 0.05. La differenza tra O ed A è stata valutata con un test Chi-Quadro ($p<0.10$ e $O>A$).
Le associazioni meta-analitiche sono state riclassificate a seconda del livello di credibilità in: convincente (I), altamente suggestiva (II), suggestiva (III), debole evidenza, associazione non significativa. Ciò è stato possibile considerando determinati criteri, sulla base di certi valori soglia, quali: il numero di casi, il valore del p-value del modello ad effetti casuali, l'eterogeneità tra gli studi, la significatività dell'intervallo di previsione, la presenza di publication bias, l'eccesso di significatività statistica.
Tutte le analisi statistiche sono state condotte con il software STATA (versione 13.1).

Risultati

Sono stati identificati 3,183 articoli da cui sono state selezionate 26 revisioni sistematiche o meta-analisi, di cui 19 e 7 per esposizione ad AD e depressione in gravidanza e rispettivamente 43 e 6 associazione meta-analitiche.
Esposizione ad AD. Nessuna delle relazioni analizzate è stata classificata nel I livello di credibilità, mentre quattro associazioni meta-analitiche sono state inserite nella II categoria (o evidenza altamente suggestiva), quali: assunzione di AD [18] o SSRI [19] vs nascita pretermine, assunzione di SSRI vs problemi respiratori [19], e SSRI assunti durante il primo trimestre vs malformazioni cardiovascolari del neonato [20].
I vincoli necessari affinchè lo studio sia categorizzato nel II livello di credibilità comprendono l'alto numero di outcome neonatali individuati (> 1000), la significatività della stima *pooled* (p-value < 10^{-6}) e l'IC dello studio più grande non inclusivo del valore nullo.
Per quanto riguarda queste associazioni meta-analitiche, non sempre la significatività della stima viene confermata dall'intervallo di previsione, inoltre per la maggior parte di queste relazioni è stato riscontrato un elevato livello di eterogenità tra gli studi, assenza di *publication bias* ed eccesso di significatività statistica, come confermato dal p-value > 0.10 dei rispettivi test.
Al contrario, non si è riscontrata nessuna associazione tra: i) utilizzo di AD o SSRI vs malformazioni neonatali /difetti del setto ventricolare, ii) venlafaxine/citalopram/fluoxetine/sertraline vs malformazioni neonatali, e iii) fluoxetine/sertraline vs difetti di sistemi fisiologici.
Esposizione a sindrome depressiva. Nessuna delle associazioni è stata inserita nella I o II categoria, mentre solamente la relazione tra depressione in gravidanza vs nascita pretermine [21-24] è stata inclusa nella III classe (o evidenza suggestiva). L'associazione meta-analitica è classificata in questa categoria se caratterizzata da un elevato numero di casi (> 1000) e con significatività statistica della stima *pooled* ottenuta dal modello ad effetti casuali pari a p-value < 10^{-3}.
L'intervallo di previsione di questa relazione conferma la significatività delle stime, presenta alta eterogeneità tra gli studi ed assenza *di publication bias*. Non è stato purtroppo possibile determinare il test circa l'eccesso di significatività a causa di alcune informazioni mancanti.
Non si osserva nessuna associazione tra l'esposizione a sindrome depressiva durante la gravidanza e (i) problemi respiratori riscontrati nel neonato, e (ii) la valutazione del coinvolgimento dell'unità di terapia intensiva neonatale.

Studi	Esposizione	Outcome neonatali	Casi	Livello di significatività (modello ad effetti casuali)	Stima dello studio più grande (95% IC)	95% intervallo di previsione	Eterogeneità	Publication bias	Eccesso di significatività statistica		Stima pooled del modello ad effetti casuali (95% IC)
									O/A	p-value	
ESPOSIZIONE AD ANTIDEPRESSIVI											
	ASSOCIAZIONI SUPPORTATE DA EVIDENZA ALTAMENTE SUGGESTIVA (CLASSE II)										
Huang, 2014	AD	Nascita pretermine	> 1,000	<0.001	1.45 (1.31 to 1.63)	1.23 to 2.30	Non alta	<0.01	18/10.7	<0.01	1.68 (1.52 to 1.86)
Eke, 2016	SSRI	Problemi respiratori	> 1,000	<0.001	1.13 (1.07 to 1.19)	0.83 to 2.12	Molto alta	0.01	5/3.6	0.17	1.33 (1.14 to 1.55)
Eke, 2016	SSRI	Nascita pretermine	> 1,000	<0.001	1.32 (1.26 to 1.38)	0.99 to 1.90	Alta	0.21	6/4.6	0.31	1.37 (1.20 to 1.56)
Zhang, 2017	SSRI, I trimestre	Malformazioni cardiovascolari	> 1,000	<0.001	1.15 (1.05 to 1.26)	0.92 to 1.72	Alta	0.60	6/4.2	0.34	1.25 (1.13 to 1.39)
ESPOSIZIONE A SINDROME DEPRESSIVA											
	ASSOCIAZIONI SUPPORTATE DA EVIDENZA SUGGESTIVA (CLASSE III)										
Grigoriadis, 2013; Jarde, 2016; Accortt, 2015; Staneva, 2015	Depressione	Nascita pretermine	> 1,000	<0.001	1.03 (0.99 to 1.06)	0.46 to 4.21	Molto alta	1.00			1.39 (1.16 to 1.66)

Conclusioni

Questa *umbrella review* analizza diversi esiti neonatali che appaiono essere strettamente correlati con l'uso di AD o l'esposizione a sindrome depressiva in gravidanza. La mancanza di una forte evidenza tra le associazioni riportate potrebbe essere spiegata dalla selezione delle donne gravide, caratterizzate da una diversa severità della sindrome depressiva o tipo di AD assunto, differente stile di vita, o variabilità farmacogenetica. Inoltre, si potrebbe supporre che le donne in gravidanza e depresse potrebbero aver avuto maggior contatti col servizio sanitario con il conseguente accertamento e monitoraggio della patologia stessa [25], oppure queste potrebbero decidere di non utilizzare AD a scopo preventivo o ancora continuare il trattamento per evitare il rischio di ricadute, ad esempio a seguito di una precedente depressione [26]. Inoltre, la presenza di fattori di rischio, quali l'età delle donne gravide, il fumo o l'uso di cotrattamenti [21] potrebbero aver influenzato i risultati.
Diversi limiti sono stati osservati nella conduzione di questo studio, quali la diversa scala applicata per definire la sindrome depressiva o il gruppo di controllo considerato (come le donne non-depresse oppure depresse e non esposte ad AD in gravidanza). In più, la discontinuità degli AD potrebbe aver incrementato il rischio di insorgenza di eventi avversi e di recidive legate alla sindrome depressiva [27]. Infine, non è stato possibile analizzare l'effetto della depressione in uno specifico trimestre, anche se la gravità della sindrome depressiva sembra variare durante la gravidanza [28].
In generale, lo sviluppo della depressione prenatale sembra essere superiore rispetto agli anni passati [29] così un precoce intervento, come un programma di screening [30], potrebbe essere necessario per rilevare i sintomi depressivi e prevenire l'insorgenza di eventi avversi neonatali.

Bibliografia

[1] Murray J.L., Lopez A.D., World Health Organization, et al. The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020: summary. Published by the Harvard School of Public Health on behalf of the World Health Organization and the World Bank.

[2] National Collaborating Centre for Mental Health. Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance: Updated edition. British Psychological Society, 2014.

[3] Gavin N.I., Gaynes B.N., Lohr K.N., et al. Perinatal depression: A systematic review of prevalence and incidence. Obstet Gynecol., 2005 Nov;106(5):1071–83.

[4] Pedersen L.H. The risks associated with prenatal antidepressant exposure: time for a precision medicine approach. Expert Opin Drug Saf., 2017 Aug;16(8):915-21.

[5] Bogaerts A., Devlieger R., Nuyts E., et al. Effects of lifestyle intervention in obese pregnant women on gestational weight gain and mental health: A randomized controlled trial. Int J Obes (Lond)., 2013 Jun;37(6):814-21.

[6] Josefsson A., Angelsiöö L., Berg G., et al. Obstetric, somatic, and demographic risk factors for postpartum depressive symptoms. Obstet Gynecol., 2002 Feb;99(2):223-28.

[7] Howard L.M., Molyneaux E., Dennis C.-L., et al. Non-psychotic mental disorders in the perinatal period. Lancet., 2014 Nov;384(9956):1775-88.

[8] Howard M.M., Mehta N.D., Powrie R. Peripartum depression: Early recognition improves outcomes. Cleve Clin J Med., 2017 May;84(5):388-96.

[9] Stein A., Pearson R.M., Goodman S.H., et al. Effects of perinatal mental disorders on the fetus and child. Lancet., 2014 Nov;384(9956):1800-19.

[10] Stuart S., Koleva H. Psychological treatments for perinatal depression. Best Pract Res Clin Obstet Gynaecol., 2014 Jan; 28(1):61–70.

[11] Omidvar S., Faramarzi M., Hajian-Tilak K., et al. Associations of psychosocial factors with pregnancy healthy life styles. PloS one., 2018 May;13(1):e0191723.

[12] DerSimonian R., Laird N. Meta-analysis in clinical trials. Control Clin Trials., 1986 Sep;7(3):177–88.

[13] Higgins J.P., Thompson S.G., Spiegelhalter D.J. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc., 2009 Jan;172(1):137–59.

[14] Higgins J.P. Commentary: heterogeneity in meta-analysis should be expected and appropriately quantified. Int J Epidemiol., 2008 Oct;37(5):1158–60.

[15] Higgins J.P., Thompson S.G. Quantifying heterogeneity in a meta-analysis. Stat Med., 2002 Jun;21(11):1539–58.

[16] Egger M., Davey Smith G., Schneider M., et al. Bias in meta-analysis detected by a simple, graphical test. BMJ., 1997 Sep;315(7109):629–34.

[17] Ioannidis J.P., Trikalinos T.A. An exploratory test for an excess of significant findings. Clin Trials., 2007;4(3):245–53.

[18] Huang H., Coleman S., Bridge J.A., et al. A meta-analysis of the relationship between antidepressant use in pregnancy and the risk of preterm birth and low birth weight. Gen Hosp Psychiatry., 2014 Jan-Feb;36(1):13-8.

[19] Eke A. C., Saccone G., Berghella V. Selective serotonin reuptake inhibitor (SSRI) use during pregnancy and risk of preterm birth: a systematic review and meta-analysis. BJOG., 2016 Nov;123(12):1900-07.

[20] Zhang T.N., Gao S.Y., Shen Z.Q., et al. Use of selective serotonin-reuptake inhibitors in the first trimester and risk of cardiovascular-related malformations: a meta-analysis of cohort studies. Sci Rep., 2017 Feb;7:43085.

[21] Grigoriadis S., VonderPorten E.H., Mamisahvili L., et al. The impact of maternal depression during pregnancy on perinatal outcomes: a systematic review and meta-analysis. J Clin Psichiatry.,2013 Apr;74(4):e321-41.

[22] Jarde A., Morais M., Kingston D., et al. Neonatal outcomes in women with untreated antenatal depression compared with women without depression: a systematic review and meta-analysis. JAMA Psychiatry., 2016 Aug;73(8):826-37.

[23] Accortt E.E., Cheadle A.C., Dunkel Schetter C. Prenatal depression and adverse birth outcomes: an updated systematic review. Matern Child Health J., 2015 Jun;19(6):1306-37.

[24] Staneva A., Bogossian F., Pritchard M., et al. The effects of maternal depression, anxiety, and perceived stress during pregnancy on preterm birth: A systematic review. Women Birth., 2015 Sep;28(3):179-93.

[25] Ko J.Y., Farr S.L., Dietz P.M., et al. Depression and treatment among US pregnant and nonpregnant women of reproductive age, 2005–2009. J Womens Health (Larchmt)., 2012 Aug;21(8):830-6.

[26] Molyneaux E., Telesia L.A., Henshaw C., et al. Antidepressants for preventing postnatal depression. Cochrane Database Syst Rev., 2018 Apr;4:CD004363.

[27] Cohen L.S., Altshuler L.L., Harlow B.L., et al. Relapse of major depression during pregnancy in women who maintain or discontinue antidepressant treatment. JAMA., 2006 Feb;295(5):499-507.

[28] Banti S., Mauri M., Oppo A., et al. From the third month of pregnancy to 1 year postpartum. Prevalence, incidence, recurrence, and new onset of depression. Results from the Perinatal Depression–Research & Screening Unit study. Compr Psychiatry., 2011 Jul-Aug;52(4):343–51.

[29] Hidaka B.H. Depression as a disease of modernity: explanations for increasing prevalence. J Affect Disord., 2012 Nov;140(3):205-14.

[30] Milgrom J., Gemmill, A.W. Screening for perinatal depression. Best Pract Res Clin Obstet Gynaecol., 2014 Jan;28(1):13–23.

NERVECHECK: A CHEAP AND ACCURATE DIAGNOSTIC TOOL FOR DIABETIC PERIPHERAL NEUROPATHY. A SINGLE-CENTER REAL-LIFE EXPERIENCE.

Pia Clara Pafundi[1], Raffaele Galiero[1,2], Paul Valensi [2]

[1] *University of Campania "Luigi Vanvitelli", Department of Advanced Medical and Surgical Sciences, Piazza Luigi Miraglia 2, 80138, Naples (NA), Italy*

[2] *Department of Endocrinology-Diabetology-Nutrition, AP-HP, Jean Verdier Hospital, Paris Nord University, Sorbonne Paris Cité, Bondy, France*

Introduction

Several studies report that peripheral neuropathies developing in patients with diabetes are known to be heterogenous by symptoms, pattern of nerve involvement, course, risk factors and pathologic alterations [1]. The heterogeneity of symptoms and types of disease sometimes renders difficult an early diagnosis of the disease [2]. Hence, it is important to develop new tools, devices, besides the gold-standard tests, able to do this, alongside with low costs for the patients and a good diagnostic efficacy. Among these, in the last years, NerveCheck Master has been developed, as an innovative Quantitative Sensory Testing medical device that could allow early diagnosis of sensory nerve fibers alterations with a focus on Diabetic Neuropathy [3].

Aims

The aim of our study was to assess NerveCheck tests accuracy in a population of type 1 diabetes mellitus (DM) and to compare with previous data in the literature.

Methods

This is a real-life single-center study. We enrolled 76 patients with type 1 diabetes mellitus in a period of six months, from October 1st 2018 to March 31th 2019. All patients underwent the complete Nerve Check Master investigation, consisting of four tests: vibration, cold, warm and pain, and were assessed using the Michigan Neuropathy Screening Instrument (MNSI).

Results

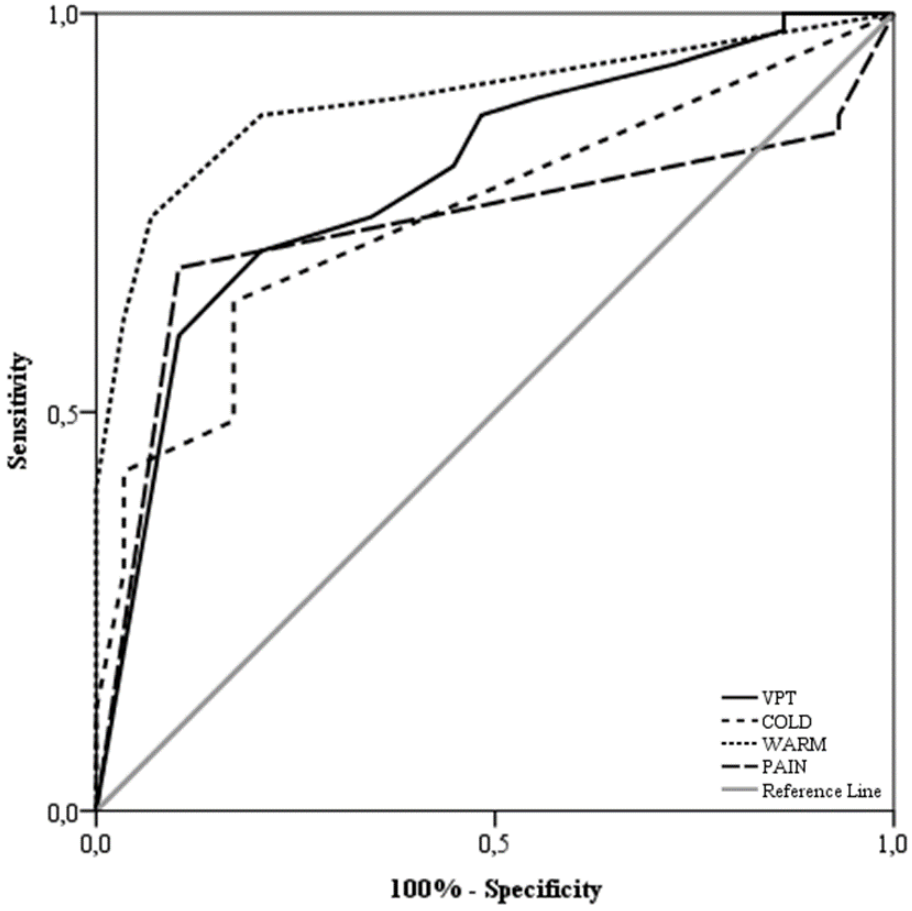
The patients had a median age of 35 years (IQR 23.3-49.8) and duration of diabetes was 13,5 (9-23,8). 20 patients were diagnosed with and 56 without diabetic peripheral neuropathy (DPN) based on MNSI. Sex, diabetes duration, HbA1c and cholesterol levels did not differ between the groups with and without DPN, whereas age ($p<0.0001$), BMI ($p=0.023$) and systolic blood pressure ($p=0.027$) were significantly higher in those with DPN. With the Nerve-Check tests, the group with DPN had a significantly lower cold perception threshold (CPT) ($p=0.003$) and warm perception threshold (WPT) ($p=0.011$) compared to those without DPN, whilst only a trend was observed for vibration perception threshold (VPT) ($p=0.052$). We hence assessed, in our population, the diagnostic performance of each NerveCheck test, which allowed us to define NerveCheck as a valid tool for diagnosis of DPN through a ROC curve analysis (Figure 1). The diagnostic performance of NerveCheck in detecting vibration perception threshold (VPT) showed an AUC of 0.796 (S.E.: 0.053; 95% C.I. 0.693–0.899; $p<0.0001$), with a high specificity (79.3%) and a moderate sensitivity (69.7). Cold perception threshold (CPT) AUC was instead 0.748 (S.E.: 0.056; 95% C.I. 0.639–0.858; $p<0.0001$), showing a perfect specificity (100%) and a high sensitivity (87.2%). Warm perception threshold (WPT) AUC was 0.893 (S.E.: 0.037; 95% C.I. 0.819–0.966; $p<0.0001$) with a moderate sensitivity

(70%) and a high specificity (93.1%) and, finally, Heat Pain Threshold (HPT) was 0.734 (S.E. 0.060; 95% C.I. 0.617–0.851; p=0.001) with a moderate sensitivity 68% and a high specificity (89.7%).

Conclusions

Our findings are consistent with the results of the unique published study on the usage of NerveCheck device in diabetic patients [4], which showed a high specificity and sensitivity for VPT (81% and 84%, respectively), a moderate specificity and high sensitivity for CPT (67% and 89%, respectively) and a moderate specificity and high sensitivity for WPT (66% and 75%, respectively) when taking SNCV, sural nerve conduction velocity as the reference.

Figure 1. ROC curve. *ROC curve analysis was used to evaluate the diagnostic performance of Nervecheck in detecting vibration perception threshold (VPT) (black line). cold perception threshold (CPT) (dashed line). warm perception threshold (WPT) (dotted line) and Heat Pain Threshold (HPT) (broken line) as reference methods for Quantitative sensory testing QST. The grey line is the reference line of the ROC curve. The AUC for VPT was 0.796 (S.E.: 0.053. 95% C.I. 0.693–0.899. p<0.0001). for CPT 0.748 (S.E.: 0.056. 95% C.I. 0.639–0.858. p<0.0001). WPT 0.893 (S.E.: 0.037. 95% C.I. 0.819–0.966. p<0.0001) and for HPT 0.734 (S.E. 0.060; 95% C.I. 0.617–0.851. p=0.001*



References

[1] Dyck PJ, Kratz KM, Karnes JL, et al. The prevalence by staged severity of various types of diabetic neuropathy, retinopathy, and nephropathy in a population-based cohort: the Rochester Diabetic Neuropathy Study. *Neurology* 1993;43:817–824.

[2] Tesfaye S, Boulton AJ, Dyck PJ, et al.; Toronto Diabetic Neuropathy Expert Group. Diabetic neuropathies: update on definitions, diagnostic criteria, estimation of severity, and treatments. *Diabetes Care* 2010;33:2285-93.

[3] Ponirakis G, Odriozola MN, Odriozola S, et al. NerveCheck for the Detection of Sensory Loss and Neuropathic Pain in Diabetes. *Diabetes Technol Ther* 2016;18:800-805.

[4] Ponirakis G, Odriozola MN, Odriozola S, et al. NerveCheck: An inexpensive quantitative sensory testing device for patients with diabetic neuropathy. *Diabetes Res Clin Pract* 2016;113:101-7.

AUTISM SPECTRUM DISORDER AND YOUNG ADULTS SOCIAL COGNITION NETWORKS COMPARISON

Vagnetti Roberto[1], Pino Chiara[1], Mazza M.[1], Masedu Francesco[1], Valenti Marco[1,2]

1 University of L'Aquila, Department of Applied Clinical Sciences and Biotechnology

2 Regional Centre for Autism, Abruzzo Region Health System, L' Aquila, Italy

Introduction

Autism Spectrum Disorder (ASD) is a heterogeneous neurodevelopmental disorder distinguished by a varied severity of symptoms among affected individuals [1;2]. According to leading research in this field [3;4;5;6], the central core of autism is the Social Cognition (SC) abilities impairment. SC is a complex mental process that refers to the capacity of people to store, process, and apply information about other people and social contexts [4]; this ability is essential in our life as it allows us to predict other people's behaviours and to modify one's own behaviour in response [6;7].
A fundamental component of SC is the Theory of Mind (ToM): the ability to attribute mental (cognitive ToM) and emotional (affective ToM) states to oneself and others and to use these attributions to make sense of and predict behaviour [3;5]. According to recent literature [5;6;8], individuals with ASD are characterised by a delay in the development of SC capacities rather than a total lack of these complex domains/constructs. Indeed, in Typically Developing (TD) children, SC capacities emerge in a specific sequence [9;6;8]. At the same time, individuals with ASD are characterised by the same sequences regarding the development of these competencies; however, they develop later than in TD children.
SC can be understood as a complex network diagram that includes distinct components such us emotion processing, biological motion perception, empathy, ToM, self-processing, affiliation, and social identity. All of these components are interrelated, and A method to evaluate the SC association patterns is the graph theory. Indeed, the graph theory allows the exploration of associations among interacting elements in a complex network such as the SC domain [10]. This type of method can be understood as an ecological approach in which the SC components are connected together to resolve a social situation and to construct an adaptive response; in fact, the interrelation among SC components is crucial to developing social behaviours and maintaining stable social relationships [9;8]. In this regard, we suggest that adults with autism, due to the delay in the development of SC components, could have a dysfunctional or poor social network in which the SC components fail to work effectively to ensure good social functioning.

Objectives

This study aims at characterizing ASD and TD SC competencies in terms of networks performances and topology. We applied the graph theory to behavioural data to verify how the SC components interact, and to establish which SC competencies are important within interacting social networks; these were applied to both ASD and TD adults, after which analysis took place.

Methods

Participants. Our study included 65 male ASD participants selected by the Reference Regional Centre for Autism, Abruzzo's Region Health System, L'Aquila, Italy (mean±standard deviation chronological age =21.43±2.06) and 61 male TD participants recruited from the University of L'Aquila, Italy (mean±standard deviation chronological age =21.52±1.97). No differences between the groups (ASD and TD) emerged for chronological age (p= 0.80). The ASD diagnoses were provided by experienced clinicians according to the

new criteria of the DSM-5 (APA, 2013). ASD diagnoses were confirmed using the Autism Diagnostic Observation Schedule, Second Edition [11]. In respect of their chronological age, individuals with ASD and TD were tested with the Wechsler Adult Intelligence Scale [12].
It was crucial that participants in the TD group had not been diagnosed with any neurological or psychological disorders. All the participants were tested individually in a quiet room according to the principles established by the Declaration of Helsinki. The study was approved by the Ethical Committee of the NHS Local Health Unit (Azienda Sanitaria Locale 1) that approved the experimental protocol prior to the recruitment of participants, according to the principles established by the Declaration of Helsinki. Informed consent was obtained from all the participants before the study.

Social Cognition Measures

Basic Empathy Scale [13; 14]. The BES is composed of two sub-scales: the Affective Empathy Sub-scale (AES) and the Cognitive Empathy Sub-scale (CES). The AES is composed of 11 items that measure an individual's ability to share another person's emotions. The CES comprises nine items and measures a person's understanding of another person's emotions.
Eyes Task [15]. The Eyes Task is a revised version of the "Reading the Mind in the Eyes Test"; this test was considered by Baron-Cohen and collaborators to be a 'pure' ToM test. The participants are given 36 photographs depicting the ocular area in an equal number of different actors and actresses. At the corner of every photo, four emotional descriptors (e.g., dispirited, bored, playful, or comforting) are printed, only one of which (the target word) correctly identifies the depicted person's mental state, while the others are included as foils. Maximum total score is 36.
Empathy Quotient [16]. The EQ is a self-reported measure evaluating different aspects of empathy through cognitive, social skill, and emotional sub-scales. The cognitive dimension of empathy is evaluated by two sub-scales of the EQ: cognitive empathy (CEQ) and social skills (SSQ), which measure, respectively, the capacity to understand the perspective of the other person, and a number of regulatory mechanisms that keep track of the origins of one's own and others' feelings. The emotional dimension is evaluated by the emotional sub-scale (EEQ).
Advanced-Theory of Mind Task [17]. The A-ToM is an Italian adaptation of a cognitive task that Blair and Cipolotti used and was first proposed by Happè. The Italian task consists of an abridged version of 13 vignettes, each accompanied by two questions: the comprehension question "Was it true, what X said?", and the justification question "Why did X say that?". The 12 story-types include Lie, White Lie, Joke, Pretense, Misunderstanding, Double Bluff, Contrary Emotions, Figure of Speech, Appearance/Reality, Forgetting, Irony, and Persuasion. The subject obtains a score ranging from 0 to 1 for each question. Maximum score is 13.
Data and network analysis. Preliminarily a multivariate comparisons of the score vector profiles of TD and ASD subjects have been performed using a Hotelling test.
We used the graph analysis to define the relationship between social cognition domains. Edges can be weighted and an edge with an higher weight is more strongly connected with a node than an edge with a lower weight. In our study, SC measures constitutes nodes of the network with correlations between them as weighted and undirected edges. Some of canonical centrality indices are represented by strength, betweenness and closeness [18]. Strength quantifies how much information a node hold, betweenness quantifies how a node act as a 'bridge' between nodes and it is useful to understand communication dynamics and closeness quantifies how a node influence a network most quickly. Two graphs were individually constructed for the ASD and TD groups, where nodes represent psychological domains (in our study, the nodes represent SC components/abilities) obtained from the assessed tests. We transformed all the scores of SC measures into z-scores to allow better comparison of the data. Then networks were estimated using the Gaussian graphical model [19] where edges can directly be interpreted as pairwise

correlation coefficient using covariance matrix as input. Significant correlations were used to select relevant edges.
Afterwards we evaluated some canonical centrality indices, such as strength, betweenness and closeness of each node (SC components) listed below.
Strength: weighted measure of the degree nodes connection

$$k_i = \sum_{j \in N} w_{ij}$$

where k represents the strength and w represents the weight between the nodes I and j, we decided to set was the correlation coefficient.
Betweenness: how many times a node serves as a bridge between two other nodes:

$$b_i = \sum_{j<k} \frac{p_{jk}(i)}{p_{jk}}$$

p_{jk} represent the number of shortest paths between nodes j and k and $p_{jk}(i)$ is the number of shortest paths between j and k that pass-through i. The higher betweenness is the greater number of shortest paths pass through this node.
Closeness: average length of the shortest path between a node and any other node:

$$L_i^{-1} = \frac{n-1}{\sum_{j \in N j \neq i} d_{ij}}$$

d represent the shortest path length between node j and i, n represents the number of nodes. It can be considered a measure of how long an information from a node will it take to reach other nodes.
In order to evaluate differences between TD and ASD networks, we bootstrapped each network 1000 times and for each bootstrap we obtained strength, betweenness and closeness for each node. The analysis was performed using the Bootnet package [20].

Results

Nodes represent different SC components. Lines between nodes represent correlations between measures. The width of the lines indicates how strength a correlation is, while red and blue lines represent positive (blue) or negative (red) correlations. Closeness is represented by node's distance (Fig. 1).
The graph analysis on centrality indices (strength, betweenness and closeness) of each node (SC components) and between groups (ASD and TD) showed significant differences between ASD and TD graph's nodes in strength and betweenness proprieties. No significant differences were found for closeness property (see Table 2). Specifically, TD group showed higher betweenness on SSQ of EQ node ($z= 4.43$, $p< 0.01$), higher betweenness on A-ToM node ($z= 6.12$, $p< 0.01$), compared to ASD group. Moreover, the TD group had higher strength on ET node ($z= 6.84$, $p< 0.01$), SSQ of EQ ($z= 6.36$, $p< 0.01$) and on A-ToM node ($z= 9.53$, $p< 0.01$), compared to ASD group. On the contrary, on CES of BES node, the ASD group showed higher strength ($z= -3.30$, $p= 0.02$), compared to TD group.

It is interesting to note that the betweenness on the AES of BES node showed a substantial trend to be statistically significant with ASD group that showed higher betweenness (z= -3.04, p= 0.05), compared to TD group. The graphs for the ASD and TD groups are reported in Figure 1 and the values of betweenness, strength, and closeness properties are reported in Table 1.

Fig.1. *ASD and TD groups networks*

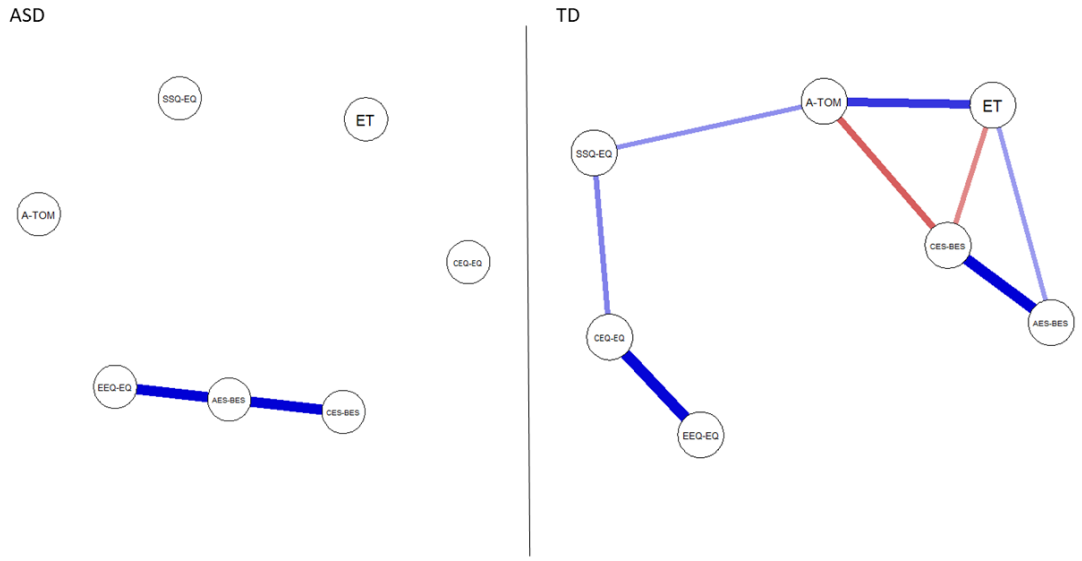


Table 1. *Demographic data for ASD and TD groups and clinical information concerning the ASD group*

	TD group (n = 61) *Mean (SD)*	**ASD group** (n = 65) *Mean (SD)*	p
Chronological Age	21.52 (1.97)	21.43 (2.06)	0.80
ADOS-social communication and social interaction	-	8.18 (2.58)	
ADOS- Repetitive and Stereotyped Behaviours	-	1.47 (1.0)	
ADOS Total scores	-	9.71 (3.35)	
QIV	98.00(23.27)	103.40(19.72)	0.16
QIP	95.00(13.50)	95.60(12.34)	0.79
QIT	96.40(15.23)	97.40(13.03)	0.69

Conclusions

The aim of this study was to explore the domain of the SC network in young adults with autism, compared with the TD adults as a control group, using the graph theory. We have built two separate networks of SC components, one for ASD and one for TD, in order to map the interacting associations among the components that characterise this construct. Indeed, the SC construct can be defined as a complex network in which interacting components, such as ToM, emotion processing, empathy, self-processing, and social identity, influence each other. Individuals with ASD are characterised by a delay in the development of SC competencies. Firstly, we showed that our individuals with ASD have an impairment in all SC measures (Eyes-task, BES, EQ, and A-ToM task) used in this study, compared to TD individuals. This result is in line with literature and it confirmed that ASD individuals have considerable difficulties in both affective and cognitive empathy.

Furthermore, our graph analysis results showed that the SC network is significantly different in TD and ASD groups. In particular, the first thing was that in the TD group all the SC nodes were connected, while in the ASD group the dimensions are highly disconnected (see Figure 1). These results suggest that the SC domains of ASD network are characterized by poor communication between them, unlike the SC network of TD group. Indeed, the SC network of the TD group showed increased betweenness among nodes, specifically the A-ToM (the complex cognitive capacity) and the SSQ of EQ (social skills abilities) nodes, compared to the ASD group. In TD adults, therefore, the information crosses several SC domains, thus realising a more complex and integrated construction response to social situations. Indeed, the TD individuals showed both higher betweenness and strength on A-ToM, ET and SSQ of EQ (cognitive empathy components), compared to ASD group. Specifically, an increased ability to understand the mental states of other people (A-ToM and ET) and the capacity to understand the perspective of another person, in addition to a number of regulatory mechanisms that keep track of the origins of one's own and others' feelings (SSQ of EQ) that are connected to other SC competencies, such as the affective dimension of empathy, as measured by EEQ of EQ and AES of BES (see Figure 1). This suggests that affective and cognitive information in SC networks are integrated through complex cognitive mechanism (A-ToM and SSQ of EQ) in TD individuals.

To contrast, the SC network of the ASD group showed that the efficiency connection among all SC components/domains mediated by complex cognitive domain (A-ToM and SSQ of EQ) elaboration is completely absent. Indeed, as shown in our results, the ASD network is characterised by increased betweenness, i.e. major functional communication between nodes, in the affective component of empathy (AES of BES), when compared to the TD network. This means that ASD individuals, unlike the TD group, use increased levels of their affective empathy ability (even if their showed lower scores in these affective abilities), and in a more isolated way, to construct a social representation; this result underline as the specific difficulties in ToM abilities can be considered the "central core" of ASD individuals [21]. Moreover, the ASD group showed higher strength, i.e. higher activation between nodes, on CES of BES (cognitive empathy) compared to TD group. Nevertheless, the lack of connections between cognitive and affective empathy and the general disconnection observed in the ASD network explains the poor functional communication among nodes (SC abilities) that characterise their network. Overall, our results support the idea that young adults with TD have a functional social network, as demonstrated by the fact that all SC competencies communicate with each other, all information is exchanged between different nodes, and the existence of a strong betweenness of the cognitive abilities that mediate the proper functioning of the other domains. On the contrary, the lack of connection and communication among SC domains is characteristic of the ASD network.

REFERENCES

[1] Lenroot R. K., Yeung P. K., Heterogeneity within autism spectrum disorders: what have we learned from neuroimaging studies? Frontiers in human neuroscience, 2013 Oct; 7: 733.

[2] Peretti S., Mariano M., Mazzocchetti C., et al., Diet: the keystone of autism spectrum disorder? Nutritional neuroscience, 2018 Apr; 19: 1-15.

[3] Baron-Cohen S., Leslie A. M., Frith U., (1985) Does the autistic child have a "theory of mind"? Cognition, 1958 Oct; 21(1): 37-46.

[4] Happé F., Cook J. L., Bird G., (2017). The structure of social cognition: In (ter) dependence of sociocognitive processes. Annual review of psychology, 2017 Jan; 68: 243-67.

[5] Mazza M., Mariano M., Peretti S., et al., The role of theory of mind on social information processing in children with autism spectrum disorders: A mediation analysis. Journal of autism and developmental disorders, 2017 May; 47(5): 1369-79.

[6] Pino M. C., Mazza M., Mariano M., et al., Simple mindreading abilities predict complex theory of mind: Developmental delay in Autism Spectrum Disorders. Journal of autism and developmental disorders, 2017 Sep; 47(9): 2743-56.

[7] Vetter N. C., Leipold K., Kliegel M., et al., Ongoing development of social cognition in adolescence. Child Neuropsychology, 2012 Aug; 19(6): 615-29.

[8] Pino M. C., Mariano M., Peretti S., et al., When do children with autism develop adequate social behaviour? Cross-sectional analysis of developmental trajectories. European Journal of Developmental Psychology, 2018 Oct; 1-17.

[9] Happé F., Frith U., Annual research review: Towards a developmental neuroscience of atypical social cognition. Journal of Child Psychology and Psychiatry, 2013 Nov; 55(6): 553-77.

[10] Ibrahim G. M., Morgan B. R., Vogan V. M., et al., Mapping the network of neuropsychological impairment in children with autism spectrum disorder: A graph theoretical analysis. Journal of autism and developmental disorders, 2016 Sep; 46(12): 3770-77.

[11] Lord C., Rutter M., DiLavore P. C., et al., Autism Diagnostic Observation Schedule (ADOS-2): Manual, 2ndEdn. Los Angeles: CA: Western Psychological Services, 2012.

[12] Wechsler D., Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV). San Antonio, TX: The Psychological Corporation, 2008.

[13] Albiero P., Matricardi G., Speltri D., et al., The assessment of empathy in adolescence: A contribution to the Italian validation of the "Basic Empathy Scale". Journal of adolescence, 2009 Apr; 32(2): 393-408.

[14] Jolliffe D., Farrington D. P., Development and validation of the Basic Empathy Scale. Journal of adolescence, 2006 Aug; 29(4): 589-611.

[15] Baron-Cohen S., Wheelwright S., Hill J., et al., The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. Journal of child psychology and psychiatry, 2003 Oct; 42(2): 241-51.

[16] Baron-Cohen S., Wheelwright S., The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. Journal of autism and developmental disorders, 2004 Apr;34(2): 163-75.

[17] Happé F. G., An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. Journal of autism and Developmental disorders, 1994 Apr; 24(2): 129-54.

[18] Rubinov M., Sporns O., Complex network measures of brain connectivity: uses and interpretations. Neuroimage, 2010 Sep;52(3): 1059-69.

[19] Lauritzen S. L., Graphical Models. Oxford, United, Kingdom: Clarendon Press, 1996.

[20] Epskamp S., Borsboom D., Fried E. I., Estimating psychological networks and their accuracy: A tutorial paper. Behavior Research Methods, 2018 Feb; 50(1): 195-212.

[21] Warrier V., Baron-Cohen S., Genetic contribution to 'theory of mind' in adolescence. Scientific reports, 2018 Feb; 8(1): 3465.

AN ANALYSIS OF THE FOOD AND DRUG ADMINISTRATION'S ADVERSE EVENT REPORTING SYSTEM FOR DETECTING SIGNALS OF SUICIDAL BEHAVIOR IN SCHIZOPHRENIC PATIENTS TREATED WITH SECOND GENERATION ANTIPSYCHOTICS

Azzolina Danila[1,2], Lorenzoni Giulia[1], Bottigliengo Daniele[1], Gregori Dario[1]

[1] *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, Vascular Sciences, and Public Health, University of Padova, Italy*

[2] *Department of Translational Medicine, Università del Piemonte Orientale, Italy*

Introduction

Life expectation of people affected by schizophrenia is 10 years shorter than general population and one of the major cause of death is suicide [1,2]. Some second-generation antipsychotics (SGAs) have been hypothesized to be protective against suicide, but results are inconsistent [3].

Objectives

The present study aimed at conducting an analysis, based on the Food and Drug Administration (FDA) Adverse Events Reporting System (FAERS) database, about the association between suicide and SGAs, in schizophrenic patients.

Methods

The FAERS database represents a data collection system of adverse drug events. Events are notified, on a voluntarily basis, from all over the world. The structure of the data collection system is based on 7 datasets. They could be linked to each other to form a single relational dataset. The DEMO one is the most important because it is the one to which all other datasets could be linked to. It reports patients' demographics and other administrative data. The following datasets concern the drug therapy: DRUG (drug/s suspected to be involved in the adverse event), INDI (diagnosis/indication for which the drug has been prescribed, coded using the Preferred Terms (PT) of the Medical Dictionary for Regulatory Activities (MedDRA) (https://www.meddra.org/) [4]), and THER (therapy start and end date, for each drug). The REAC dataset reports the type of reaction/adverse event coded using the MedDRA PT, and the OUTC dataset refers to patient outcome. The RPSR reports information about the source notifying the adverse event.
Records reporting cases of completed suicide, tentative suicide, and suicidal ideation in schizophrenic patients treated with SGAs (clozapine, risperidone, aripiprazole, quetiapine, and olanzapine) were extracted from the FAERS database. The adverse events of suicidal ideation, completed suicide, and suicide attempt were considered as the main outcomes and they were analyzed both individually and grouped all together. A disproportionality analysis was conducted to detect quantitative signals identifying drug-event association, using the reporting odds ratio (ROR) algorithm [5]. Both crude and adjusted (according to age and gender) ROR were estimated, using a multivariable model based on a Generalized Linear Model (logit link). Clozapine was chosen as the reference drug since it is the only SGA considered to be protective against suicide.

Results

The analysis of the FAERS database, for the period considered (from the 4th quarter of 2012 to the1st quarter of 2019), identified 70.086 records with the diagnosis/indication of schizophrenia. The outcomes of interest were reported in 740 out of 70.086 records. The most prevalent outcome of interest was suicidal ideation (reported in about 40% of records), followed by suicide attempts and completed suicide. Generally, the male

gender and the adult age (18-65 years) were the most prevalent, even if gender and age distribution varied significantly by drugs (p-value <0.001). Considering the three outcomes pooled together, all SGAs were found to have a significant higher risk for the outcome of interest compared to clozapine. Similar results were obtained by modelling two out of the three outcomes of interest individually (suicide ideation and suicide attempt). Conversely, looking only at the cases of completed suicide, risperidone was found to have a protective role compared to clozapine at both unadjusted (ROR 0.55, 95% C.I. 0.32-0.97) and adjusted analysis (ROR 0.56, 95%C.I. 0.32-0.98). Also the other SGAs analyzed were found to have a protective role against completed suicide compared to clozapine, but the effect was not significant.

Conclusions

Results of the present study confirmed the protective role of clozapine on suicide (considering the outcomes of interest pooled together). However, when the outcomes were considered individually, the results obtained were not consistent to each other. Further studies are needed to better clarify the role of each SGA in affecting each outcome of interest.

References

[1] McGrath J, Saha S, Chant D, et al., Schizophrenia: A Concise Overview of Incidence, Prevalence, and Mortality. Epidemiol Rev., 2008; 30:67–76

[2] Pompili M, Amador XF, Girardi P, et al., Suicide risk in schizophrenia: learning from the past to change the future. Ann Gen Psychiatry., 2007; 6:10

[3] Pompili M, Baldessarini RJ, Forte A, et al., Do Atypical Antipsychotics Have Antisuicidal Effects? A Hypothesis-Generating Overview. Int J Mol Sci., 2016;17

[4] Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). Drug Saf., 1999; 20:109–17

[5] Rothman KJ, Lanes S, Sacks ST. The reporting odds ratio and its advantages over the proportional reporting ratio. Pharmacoepidemiol Drug Saf., 2004; 13:519–23

APPLICATION AND VALIDATION OF A BIOLOGICAL AGING PREDICTOR BASED ON MACHINE LEARNING AS A MARKER OF PUBLIC-HEALTH IN THE ITALIAN POPULATION

Gialluisi Alessandro[1*], Di Castelnuovo Augusto[2], Costanzo Simona[1], De Curtis Amalia[1], Magnacca Sara[2], Bonaccio Marialaura[1], Cerletti Chiara[1], Donati Maria Benedetta[1], De Gaetano Giovanni[1], Iacoviello Licia[1,3] on behalf of the Moli-sani Study Investigators[4]

[1] *Department of Epidemiology and Prevention, IRCCS NEUROMED, Pozzilli, Italy*

[2] *Mediterranea Cardiocentro, Napoli, Italy*

[3] *Research Center in Epidemiology and Preventive Medicine (EPIMED), Department of Medicine and Surgery, University of Insubria, Varese, Italy*

[4] *A complete list of the Moli-sani Project Investigators is reported online at https://www.moli-sani.org/?page_id=181*

Introduction

Recently, deep learning algorithms have been developed for the estimation of Biological Age (BA), the hypothetical underlying age of an organism, which can differ from its chronological age (CA) and can be computed based on circulating biomarkers [1-3]. However, these algorithms are population-specific, and warrant further validation as public health screening tools [1].

Aims

To develop deep learning BA estimates in the Italian population and test their predictivity of clinical events, and potential environmental influences.

Methods

We trained Deep Neural Networks (DNN) based on 38 circulating biomarkers (including haemochrome and biochemical measures), to compute BA in an Italian general population cohort, the Moli-sani study (N=24,325; age≥35 years; 51.5% women) [4]. Then we computed the difference between BA and CA (Δage), to have a measure of higher (Δage>0) or lower than expected ageing (Δage<0) for each participant. Finally, we tested its predictivity of all-cause mortality and hospitalization risk through a cox regression. Moreover, we investigated associations of Δage with different exposures, including education, obesity and main lifestyle factors (diet, smoking and physical activity), through generalized linear models. All models were adjusted for age (CA) and sex.

Results

The best performing DNN showed accurate BA estimates, compared to CA values (Mean Average Error 6.29 years; Pearson´s r=0.75; R^2=0.56). The resulting Δage significantly predicted the risk of all-cause mortality (HR = 1.08[1.06-1.11] per year increase; $p=1.0x10^{-12}$) and hospitalizations (1.03[1.02-1.04]; $2.2x10^{-14}$). Δage was significantly associated with obesity (β(SE)=0.20(0.02) years per unit increase of BMI), diet (-0.20(0.06) per unit increase in adherence score to Mediterranean Diet [4]), smoking status (β(SE)=1.14(0.24) for smokers vs non-smokers) and educational level completed (-0.83(0.27) and -1.76(0.34) for upper secondary and post-secondary vs primary school).

Conclusion

These findings provide a robust body of evidence supporting the use of blood-based BA as a marker of public health in the general population.

References

[1] Putin E, Mamoshina P, Aliper A, et al. Deep biomarkers of human aging: Application of deep neural networks to biomarker development. Aging (Albany NY), 2016 May; 8(5):1021-33.

[2] Mamoshina P, Kochetov K, Putin E, et al. Population Specific Biomarkers of Human Aging: A Big Data Study Using South Korean, Canadian, and Eastern European Patient Populations. J Gerontol A Biol Sci Med Sci., 2018 Oct; 73(11):1482-1490.

[3] Mamoshina P, Kochetov K, Cortese F, et al. Blood Biochemistry Analysis to Detect Smoking Status and Quantify Accelerated Aging in Smokers. Sci Rep., 2019 Jan; 9(1):142.

[4] Bonaccio M, Di Castelnuovo A, Costanzo S, et al. Mediterranean diet and mortality in the elderly: a prospective cohort study and a meta-analysis. Br J Nutr., 2018 Oct; 120(8):841-854.

ESPOSIZIONE A GASTROPROTETTORI NEL PRIMO ANNO DI VITA (INIBITORI DI POMPA PROTONICA E ANTAGONISTI DEL RECETTORE H2) E IL SUCCESSIVO SVILUPPO DI ASMA NEI BAMBINI.

Comoretto Rosanna Irene[1], Barbiellini Amidei Claudio[1], Zanier Loris[2], Canova Cristina[1]

[1] Department of Cardiac, Thoracic and Vascular Sciences and Public Health, University of Padua, Padua, Italy

[2] Epidemiological Service, Health Directorate, Friuli Venezia-Giulia Region, Udine, Italy

Introduzione

I gastroprotettori, in particolare gli Inibitori di Pompa Protonica (IPP) e gli Antagonisti dei Recettori H2 (H2RA) costituiscono una delle principali soluzioni terapeutiche per la malattia da reflusso gastroesofageo, una condizione comune nell'età pediatrica. L'alterazione del pH gastrico che ne consegue può indurre cambiamenti nel microbioma intestinale, dimostrata associata allo sviluppo di patologie allergiche [1]. Diversi lavori hanno evidenziato un incremento del rischio di sviluppare l'asma nei bambini le cui madri avevano assunto gastroprotettori durante la gravidanza [2,3]. Tuttavia, l'associazione tra l'esposizione a gastroprotettori nel primo anno di vita e lo sviluppo di asma è limitata ad un recente lavoro.

Obiettivi

Quantificare il rischio di asma associato al consumo di gastroprotettori nel primo anno di vita, in una popolazione pediatrica.

Metodi

Dai certificati di assistenza al parto sono stati identificati i bambini nati in Friuli-Venezia Giulia nel periodo 1995-2011. Si è definita l'esposizione come la presenza di ≥1 prescrizione farmaceutica di IPP (ATC: A02BC) o H2RA (ATC: A02BA) nei primi 365 giorni di vita. Alcune categorie di farmaci anti-asmatici sono state considerate indicatrici di asma in presenza ≥2 prescrizioni in 12 mesi. Il follow-up dei soggetti parte dal terzo anno di età fino alla morte, emigrazione, insorgenza di asma o fine del follow-up (31 Dicembre 2012). Il rischio di sviluppare asma è stato stimato attraverso l'Hazard Ratio (HR) utilizzando il modello di regressione di Cox aggiustato per genere, anno di nascita del bambino, età gestazionale, peso alla nascita, età ed istruzione della madre. Analisi di robustezza e stratificazione per *propensity score* sono in fase di elaborazione.

Risultati

Da una coorte di 111.447 soggetti (680.199p/y), si sono identificati 1.768 bambini (1,6%) con almeno una prescrizione per gastroprotettori nel primo anno di vita (IPP: 151 soggetti (0,1%); H2RA: 1.655 soggetti (1,5%)). Tra i bambini esposti a gastroprotettori il 23,8% (51,76 1000p/y) è stato identificato come affetto da asma, mentre il 17,3% (28,17 1000p/y) tra i non trattati, con un significativo aumento del rischio del 60% (HR: 1,60 95%CI 1,45-1,77). Il 23,2% (66,73 1000p/y) degli esposti a IPP ha poi sviluppato l'asma, rispetto al 17,4% dei non esposti (28,43 1000p/y) con un significativo incremento del rischio dell'81% (HR: 1,81 95%CI 1,30-2,52). Per gli H2RA la proporzione con asma trattato era del 24,2% tra gli esposti (52,01 1000p/y) e del 17,3% tra i non esposti (28,20 1000p/y) con un rischio di 1,62 (95%CI 1,46-1,79).

Conclusioni

Il consumo di gastroprotettori nel primo anno di vita sembra essere associato a un aumentato rischio di sviluppare asma trattato in età pediatrica. Sebbene sulla base di questi dati una relazione causale non possa essere dimostrata, questi risultati confermano la necessità di valutare con attenzione il rapporto rischio-beneficio della prescrizione di farmaci gastroprotettori nei primi mesi di vita.

Bibliografia

[1] Mitre E, Susi A, Kropp LE, Schwartz DJ, Gorman GH, Nylund CM. Association Between Use of Acid-Suppressive Medications and Antibiotics During Infancy and Allergic Diseases in Early Childhood. JAMA Pediatr. 2018;172:e180315.

[2] Devine RE, Sheikh A, Nwaru BI. Acid-suppressive medications during pregnancy and risk of asthma and allergy in the offspring: protocol for a systematic review. npj Prim Care Resp Med. 2016;26:16001.

[3] Andersen ABT, Erichsen R, Farkas DK, Mehnert F, Ehrenstein V, Sørensen HT. Prenatal exposure to acid-suppressive drugs and the risk of childhood asthma: a population-based Danish cohort study. Aliment Pharmacol Ther. 2012;35:1190–8.

DISCONTINUITÀ DEI TRATTAMENTI ANTIPSORIASICI DURANTE LA GRAVIDANZA: STUDIO DI COORTE PSO-MOTHER

Francesca Romana Poggi[1], Valeria Belleudi[1], Serena Perna[2], Luigi Naldi[3], Renata Bortolus[4], Alessandro Cesare Rosa[1], Ursula Kirchmayer[1], Marina Davoli[1], Antonio Addis[1].

1 Dipartimento di Epidemiologia, Servizio Sanitario Regionale della Regione Lazio. Roma, Italia;

2 Agenzia Italiana del Farmaco. Roma, Italia;

3 Dipartimento di Dermatologia, Azienda USL 8 Berica - Ospedale San Bortolo, Vicenza, Italia;

4 Ufficio per la Promozione della Ricerca, Dipartimento di Direzione Medica Ospedaliera, AOUI Verona, Italia.

Introduzione

La psoriasi è una condizione cronica con una prevalenza mondiale di circa 2-3%. Il 50% dei pazienti sono donne, per lo più diagnosticate in età fertile. [1] Le informazioni riguardo l'aderenza alle terapie farmacologiche, da parte delle pazienti con psoriasi durante la gravidanza, sono limitate per tutti i tipi di trattamento: topico e sistemico (sia convenzionale che biologico).[2] Per la maggioranza dei trattamenti disponibili vengono riportate importanti avvertenze di potenziale embrio-fetotossicità che potrebbero influenzare la decisione di interrompere o meno i trattamenti.[1,3]

Obiettivo

Descrivere l'andamento prescrittivo delle terapie farmacologiche nelle donne con psoriasi prima e durante la gravidanza.

Metodi

Attraverso il flusso informativo del Certificato di Assistenza al Parto, sono stati identificati tutti i nati tra il 2009 e il 2016 da donne residenti e assistite nel Lazio di età compresa tra 18 e 45 anni. Tra tutte le gravidanze a termine, sono state identificate quelle relative a donne con psoriasi e in trattamento farmacologico nei sei mesi precedenti la data di inizio gravidanza (calcolata sulla base dell'età gestazionale). Il pattern d'utilizzo dei farmaci antipsoriasici (biologici, sistemici o topici) prima e durante la gravidanza è stato analizzato secondo le caratteristiche cliniche e socio-demografiche delle donne al parto. Inoltre, è stato approfondito l'uso di farmaci specificamente controindicati in gravidanza (metotrexato, acitretina).

Risultati

Dei 1.876 parti di donne affette da psoriasi, il 28% (N=525) ha avuto almeno una prescrizione di trattamento antipsoriasico nei sei mesi precedenti l'inizio della gravidanza. L'età mediana delle donne al parto era di 33 anni. Nel corso del trimestre precedente la gravidanza, il 68,8% delle donne presentava almeno una prescrizione di farmaci per la psoriasi. Per ogni tipo di trattamento considerato, si è riscontrato un decremento generale dell'utilizzo durante la gravidanza. In particolare, le percentuali di donne in trattamento nel trimestre precedente l'inizio della gravidanza e nel terzo trimestre di gravidanza sono diminuite: dall'8,4% allo 0% per i biologici, dal 5,9% al 2,5% per i convenzionali sistemici e dal 52% al 9,4% per i trattamenti topici. Il numero di donne con almeno una prescrizione di metotrexato era pari a otto nei tre mesi precedenti l'inizio della gravidanza e si riduceva a uno durante la gravidanza (prescrizione al terzo trimestre). Non si sono osservate prescrizioni di acitretina né prima né durante la gravidanza.

Conclusioni

La gravidanza sembra influenzare in modo significativo il pattern prescrittivo dei diversi trattamenti farmacologici per la psoriasi. Nell'ambito dello studio PSO-MOTHER sono previsti degli approfondimenti per capire l'impatto di tali andamenti sugli esiti di salute della mamma con psoriasi e del bambino.

Bibliografia

[1] Kurizky PS, Ferreira Cde C, Nogueira LS et al. Treatment of psoriasis and psoriatic arthritis during pregnancy and breastfeeding. An Bras Dermatol. 2015;90:367-75.

[2] Vena GA, Cassano N,Biella G et al. Psoriasis in pregnancy: challenges and solutions Psoriasis (Auckl). 2015; 5: 83–95.

[3] Lam J1, Polifka JE, Dohil MA.Safety of dermatologic drugs used in pregnant patients with psoriasis and other inflammatory skin diseases. J Am Acad Dermatol. 2008 Aug;59(2):295-315.

IMPACT OF ADHERENCE TO STANDARDS OF CARE FOR EARLY RHEUMATOID ARTHRITIS ON POPULATION HEALTH. RESULTS FROM THE RECORD LINKAGE OF RHEUMATIC DISEASE STUDY OF THE ITALIAN SOCIETY FOR RHEUMATOLOGY.

Zanetti Anna[1,2], Scirè Carlo Alberto[1,3], Guerrini Giulio[3], Carrara Greta[1], Zambon Antonella[2]

[1]*Epidemiology Unit, Italian Society for Rheumatology, Milan, Italy;*

[2]*Department of Statistics and Quantitative Methods, Division of Biostatistics, Epidemiology and Public Health, University of Milano-Bicocca, Milan, Italy;*

[3]*Rheumatology Unit, Department of Medical Sciences, University of Ferrara, Ferrara, Italy.*

Background

Early diagnosis and treatment of rheumatoid arthritis (RA) have extensively shown a strong potential of improvement of the disease outcome trajectory. [1-2] and better outcomes in terms of survival, which is dependent from early suppression of disease activity and functional impairment. [3]
The model of care including early diagnosis, prompt access to disease modifying anti-rheumatic drugs (DMARD) and early assessment of response in a treat-to-target management strategy is consistently regarded as a standard of care according to current international RA guidelines.[4]
Despite such strength of evidence and widespread consensus, the actual application of such care standards into practice is not fully completed. Being able to measure evidence-based health care quality indicators (HCQI) is advocated to evaluate the performance of health care systems and to assess the actual outcome of healthcare models at population level. Several scientific efforts focused on the development of set of HCQI to be applied at different care levels, using different sources of information.[5-7]

Objective

The general objective of these analyses was to assess the frequency and impact of the implementation of early RA standards of care at population level. The main objective was to assess in early RA the impact of the adherence to HCQI on the risk of disease-associated hospitalizations and the overall risk of hospitalizations.

Methods

This is a retrospective cohort study performed using the data of the RECORD study—promoted by the Italian Society for Rheumatology — aiming to set up a national surveillance system to monitor the health burden of rheumatic diseases in Italy. The data sources for this study were the healthcare administrative databases (AHD) of Lombardy.
The target population included Lombardy residents, aged 18 and older, who were beneficiaries of the NHS between 2004 and 2013.
To identify patients with RA, a validated algorithm developed in the RECORD study was applied.[8] In these analyses we applied the first two steps of the algorithm characterized by 85% of sensitivity and 96% of specificity. For these highly potential RA cases, the RA onset was identified considering the first access to NHS for: drug treatment with DMARD or glucocorticoids (GCs), rheumatologist certification of RA or hospitalization for RA.
To identify incident RA cases and ensure at least two years disease free and a minimum period of follow-up, only new RA cases identified between 2006 and 2012 were included. These patients accumulated person-years of follow-up starting from the cohort entry (i.e., RA onset) until the earliest date among outcome

onset, death, embarkment, or end of follow-up period (1.5 year after the cohort entry). Finally, to avoid potential misclassification of RA patients, those patients without any exposure to RA drugs or monitoring during follow-up were excluded.
The main study outcome was the hospital admission during follow-up, with primary or secondary diagnosis of RA (ICD9-CM: 714). The secondary outcome was the hospital admission for all causes.
The main exposure was defined as an adherence score based on six quality indicators related to the care process of new RA cases, measurable from AHD and derived from a set of measures already developed by the NHS to monitor the clinical pathway of RA patients. [4] The indicators concerned the treatment with DMARDs, GCs and the disease monitoring (C-reactive protein (CRP) or erythro-sedimentation rate (ESR) and consultation within 3 months from the RA onset).
To characterize the clinical profile of patients, health data from AHD related to the two years period before RA onset were considered. Data from HDFs were used to derive the Charlson Comorbidity Index (CCI), dichotomized in two classes (CCI = 0 and CCI >0) and the drug prescription database to calculate a polypharmacy indicator as the number of different drug classes (3 digits of ATC) delivered to the patients (discrete variable).
The adherence score was estimated combining the adherence indicators (with appropriate weights), following the approach suggested by Mehta et al.[9] To ensure robustness to the weights system, a thousand bootstrap-samples were extracted from the original sample and, for each of them, a multivariate Cox proportional hazards regression model was fitted. The final bootstrap estimates of beta coefficients related to the six adherence indicators were changed in sign (in order to attribute an increment of risk to covariates with negative sign), multiplied by ten and rounded to the nearest integer to obtain the weights.
Each patient was classified into one of the following three adherence score categories (low, medium and high), identified in order to ensure a balanced number of events across groups and the interpretability at clinical level.
Finally, the Kaplan-Meier curves and the corresponding hazard ratios were obtained for each adherence score category.

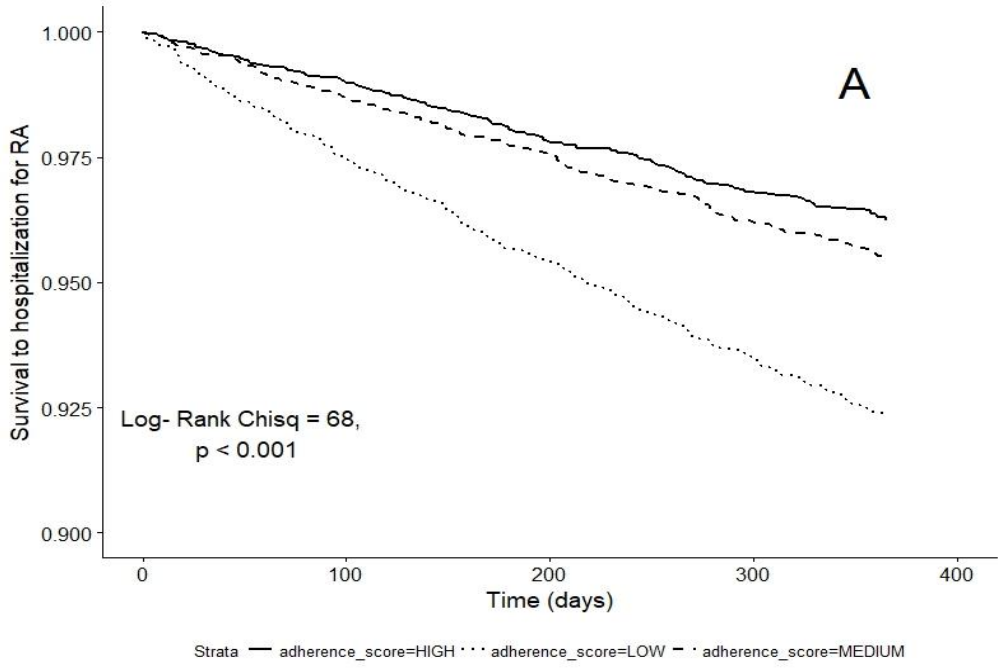
Results

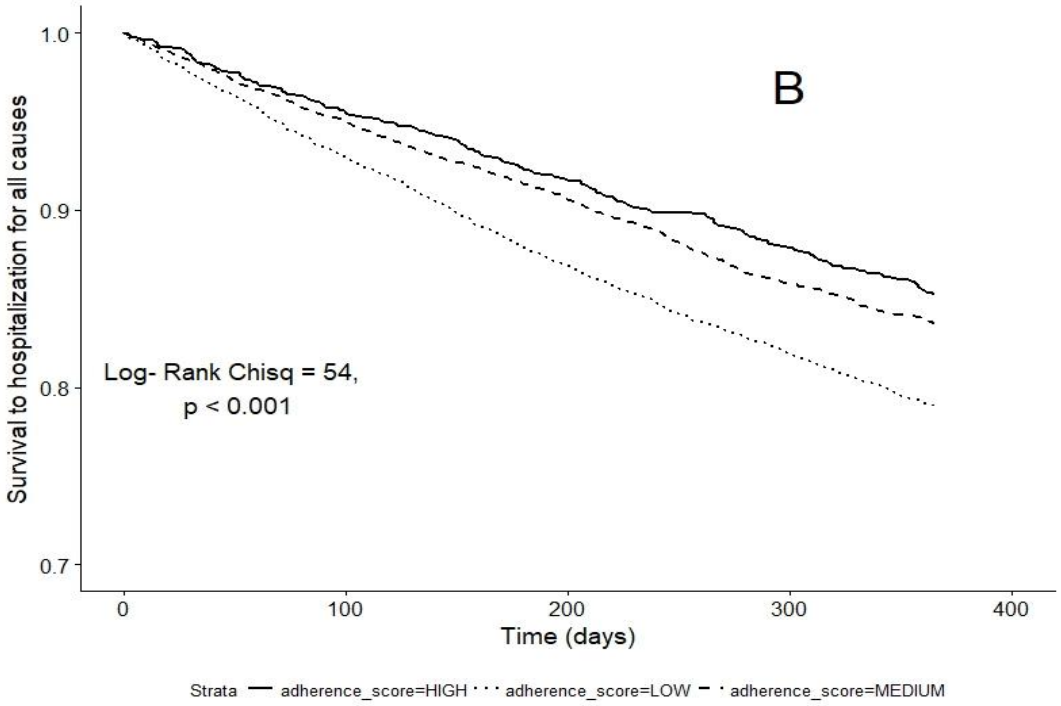
A total of 14,203 patients were included in the final cohort.
The subjects were predominantly women (71%), with a mean (SD) age of 60 years (14.1); the proportion of patients with CCI greater than 0 was 12%; the mean (SD) of drug classes prescribed was 4.1 (2.7).
A total of 704 patients (5%) were hospitalized for RA during follow-up and 2609 (18.4%) were hospitalized.
The adherence score (ranging from -4 to 4) was categorized in three levels: low adherence (adherence score=-4, 36%), intermediate adherence (from -1 to 0, 36%), high adherence (from 1 to 4, 28%). Figure 1A reports the Kaplan-Meier curves for each adherence score category, using RA hospitalization as outcome.
Considering low adherence as reference category, intermediate and high adherence significantly reduced the risk of hospitalization for RA (HR 0.65 [95%CI 0.54, 0.77] and p-value <0.001, and HR 0.58 [95%CI 0.48, 0.70] and p-value <0.001, respectively). Increasing of age and CCI level is associated with a greater risk of hospitalization for RA (HR 1.03 [95%CI 1.02, 1.04] and HR 1.51 [95%CI 1.25, 1.84] respectively, all p-values <0.001). Sex and number of drug classes (polypharmacy indicator) were not associated with the outcome (HR 1.07 [95%CI 0.91, 1.26] and p-value 0.411 and HR 1.03 [95%CI 0.99, 1.06] and p-value 0.065 respectively).
Figure 1B shows curves considering the hospitalization for all causes. Intermediate and high adherence significantly reduced the risk of hospitalization (HR 0.85 [95%CI 0.78, 0.94] and p-value <0.001, and HR 0.76 [95%CI 0.69, 0.84] and p-value <0.001, respectively).

Discussion

Despite the great progresses made in the last decades in the understanding pathogenetic mechanisms of RA, in the development of new targeted therapies and increasingly effective intervention strategies, the translation of this knowledge to health systems is not yet complete.
The main result of our analysis was that the adherence of the individual care pathway to HCQIs is associated with a decrease in the risk of hospitalization for RA, with a reduction of about 40% for patients with the highest adherence. This result reflects at system level what has already been defined in clinical studies on limited and selected groups of patients. Early hospitalization (within the first 2 years) is a measure of failure of the RA care pathway and represents one of the most important mortality risk factors after age and persistent disease activity.[10]
In conclusion, we provided one of the first evidence of the impact of the quality of care of early RA on hospitalization. These results summarize the relevance of promoting the early arthritis clinics model to improve early access to disease modifying treatments in RA, in terms of individual and system outcomes. This healthcare model appears to be a good candidate to improve the health status of patients with RA, ameliorate the treatment experience, and limit the use of health resources, leveraging on the improvement of the quality of care.

Figure 1. *Kaplan-Meier survival curves according to categories of adherence score and A) RA hospitalization, B) hospitalization for all causes.*





References

[1] Lard LR, Visser H, Speyer I, et al. Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies. *Am J Med* 2001;111:446–451.

[2] Scirè CA, Verstappen SMM, Mirjafari H, et al. Reduction of long-term disability in inflammatory polyarthritis by early and persistent suppression of joint inflammation: results from the Norfolk Arthritis Register. *Arthritis Care Res* 2011;63:945–952.

[3] Symmons DP, Jones MA, Scott DL, et al. Longterm mortality outcome in patients with rheumatoid arthritis: early presenters continue to do well. *J Rheumatol* 1998;25:1072–1077.

[4] Smolen JS, Landewé R, Bijlsma J, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017;76:960–977.

[5] Govoni M, Sciré CA, Manara M, et al. Does early arthritis clinic organisation improve outcomes? What evidence is there? A systematic review. *Clin Exp Rheumatol* 2013;31:443–451.

[6] Petersson IF, Strömbeck B, Andersen L, et al. Development of healthcare quality indicators for rheumatoid arthritis in Europe: the eumusc.net project. *Ann Rheum Dis* 2014;73:906–908.

[7] Barber CEH, Marshall DA, Mosher DP, et al. Development of System-level Performance Measures for Evaluation of Models of Care for Inflammatory Arthritis in Canada. *J Rheumatol* 2016;43:530–540.

[8] Carrara G, Scirè CA, Zambon A, et al. A validation study of a new classification algorithm to identify rheumatoid arthritis using administrative health databases: case-control and cohort diagnostic accuracy studies. Results from the RECord linkage On Rheumatic Diseases study of the Italian Society for Rheumatology. *BMJ Open* 2015;5:e006029.

[9] Mehta HB, Mehta V, Girman CJ, et al. Regression coefficient-based scoring system should be used to assign weights to the risk index. *J Clin Epidemiol* 2016;79:22–28.

[10] Lezcano-Valverde JM, Salazar F, León L, et al. Development and validation of a multivariate predictive model for rheumatoid arthritis mortality using a machine learning approach. *Sci Rep* 2017;7.

INCREASING RISK OF EMERGENCY ROOM ADMISSIONS FOR BRONCHIOLITIS IN INFANTS EXPOSED TO AIR POLLUTION

Gallo Elisa[1], Bottigliengo Daniele[1], Bressan Silvia[2], Lanera Corrado[1], Geremia Sara[3], Zagolin L.[4], Marson G.[4], Gregori Dario[1]

[1] *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padova, Padova, Italy*

[2] *Paediatric Emergency Unit, Department of Women's and Children's Health, University Hospital of Padova, Padova, Italy*

[3] *University of Bologna, Bologna, Italy*

[4] *Regional Agency for Environmental Prevention and Protection of the Veneto region (ARPAV)*

Introduction

In the last decade air pollution has been linked to an increased risk of several respiratory diseases in children, especially respiratory tract infection [1–3]. Furthermore, the developing lungs and immune systems of children makes them more vulnerable to infections [4].
Bronchiolitis consists in a lower-respiratory tract viral infection, usually due to syncytial respiratory virus, that leads to increased expiratory wheezing and respiratory distress [5].
The time course of the disease illustrated by Karr and colleagues shows the beginning of symptoms 6-7 days after the infection and a peak, compatible with the period of Emergency Room visits, on days 8-9 [6].
Moreover, the development of bronchiolitis during infancy has been linked to an increased risk of asthma or of lifetime wheezing later in life [7].
The role of air pollution on the development of bronchiolitis is still controversial. Lack of association between hospitalization for bronchiolitis and particulate matter of less than 2.5 μm ($PM_{2.5}$) and less than 10 μm (PM_{10}) in aerodynamic diameter and nitrogen dioxide (NO_2) was found by Karr and colleagues in different time-windows (lag 1, lag 4, lag 1-2, lag 3-5, lag 6-8) [6], whereas a significant association was found in the week before the hospitalization [8] or in a 5 days window [9], but not with other time-windows.

Objectives

The present study aims to (i) evaluate the association between Pediatric Emergency Room (PER) admissions for bronchiolitis and PM_{10}, $PM_{2.5}$ and NO_2, (ii) evaluate how the association changes if only hospitalized children with severe symptoms are considered.

Methods

PER admissions due to bronchiolitis in children aged less than 1 year were retrospectively collected from 2007 to 2018 in Padua, Italy. Daily environmental data were obtained from monitoring stations of the Regional Agency for Environmental Prevention and Protection of the Veneto Region (ARPAV) located within 30 km from the children residence.
A conditional logistic regression based on a time-stratified case-crossover design was used to evaluate the association between PER admissions and exposure to NO_2, $PM_{2.5}$ and PM_{10} [10]. This design accounts for the confounding role that seasonal patterns may play by selecting controls on the same day of week of the same month of the same year of cases. Thus, every subject serves as his/her own control and confounders are matched by design [11].

Models were adjusted for atmospheric pressure, temperature, humidity and national holidays. Odds Ratio for lags from 0 to 15 were obtained. The same analysis was repeated considering only hospitalized children.

Results

Overall, 2333 children presented to PER for bronchiolitis. A positive significant association was found on the first two lags for both $PM_{2.5}$ and PM_{10}. Moreover, an increasing risk, even though not statistically significant, from lag 10 to lag 13 was observed.
Regarding NO_2, a significant association with PER admissions was found from lag 10 to 13.
The analysis restricted to infants with severe symptoms showed a further increase of risk in the three days before the hospitalization for both PM_{10} and $PM_{2.5}$.

Conclusion

In the present study the association between air pollution and PER admissions for bronchiolitis was evaluated. Results show that exposure to PM (lag 0, lag1 that correspond to 7-8 days after the infection) can worsen health of children with mild symptoms of respiratory disease.
Moreover, the association between PER visits for bronchiolitis and NO_2, PM_{10} and $PM_{2.5}$ from lag 10 to lag 13 highlights that exposure of infants to high levels of NO_2 and PM could damage the respiratory tract mucosa, facilitating viral infections.

Bibliography

[1] Nhung N.T.T., Schindler C., Dien T.M., et al., Acute effects of ambient air pollution on lower respiratory infections in Hanoi children: An eight-year time series study. *Environ. Int.* 2018, *110*, 139–148.

[2] Xia X., Zhang A., Liang S., et al., The Association between Air Pollution and Population Health Risk for Respiratory Infection: A Case Study of Shenzhen, China. *Int. J. Environ. Res. Public. Health* 2017, *14*.

[3] Zheng P.-W., Wang J.-B., Zhang Z.-Y., et al., Air pollution and hospital visits for acute upper and lower respiratory infections among children in Ningbo, China: A time-series analysis. *Environ. Sci. Pollut. Res.* 2017, *24*, 18860–18869.

[4] Darrow L.A., Klein M., Flanders W.D., et al., Air Pollution and Acute Respiratory Infections Among Children 0–4 Years of Age: An 18-Year Time-Series Study. *Am. J. Epidemiol.* 2014, *180*, 968–977.

[5] Meissner H.C., Viral Bronchiolitis in Children. *N. Engl. J. Med.* 2016, *374*, 1793–1794.

[6] Karr C., Lumley T., Shepherd K., et al., A Case–Crossover Study of Wintertime Ambient Air Pollution and Infant Bronchiolitis. *Environ. Health Perspect.* 2006, *114*, 277–281.

[7] Lee J.-Y., Leem J.-H., Kim H.-C., et al., Effects of traffic-related air pollution on susceptibility to infantile bronchiolitis and childhood asthma: A cohort study in Korea. *J. Asthma* 2018, *55*, 223–230.

[8] Yitshak-Sade M.; Yudovitch D., Novack V., et al., Air Pollution and Hospitalization for Bronchiolitis among Young Children. *Ann. Am. Thorac. Soc.* 2017, *14*, 1796–1802.

[9] Ségala C., Poizeau D., Mesbah M., et al., Winter air pollution and infant bronchiolitis in Paris. *Environ. Res.* 2008, *106*, 96–100.

[10] Maclure M., The case-crossover design: a method for studying transient effects on the risk of acute events. *Am. J. Epidemiol.* 1991, *133*, 144–153.

[11] Janes H., Sheppard L., Lumley T., Case–Crossover Analyses of Air Pollution Exposure Data: Referent Selection Strategies and Their Implications for Bias. *Epidemiology* 2005, *16*, 717.

A WEIGHTED QUANTILE SUM (WQS) REGRESSION WITH DOUBLE INDEX

Renzetti Stefano[1], McRae Nia[2], Gennings Chris[2], Wright Robert O.[2], Calza Stefano[1]

[1]*Università degli Studi di Brescia*

[2] *Icahn School of Medicine at Mount Sinai*

Introduction

In the last few years the importance of identifying mixtures as a combination of chemicals like nutrients, metals, pesticides etc. was brought to light in public health and epidemiology since the hypothesis that it is not only one agent but it is the combination of multiple components to influence the disease or health state is always more strengthen [1,2,3]. Weighted Quantile Sum (WQS) regression is a recent statistical model that is increasingly used in epidemiological studies. This method allows to build an empirical weighted body burden index that represents the overall mixture exposure and test the association with the outcome of interest. The body burden index reduces the dimensionality and is more robust to multicollinearity [4,5].

Objectives

The original methodology provides the estimate of a single index that allows to measure the association between the mixture and the dependent variable in only one direction (either positive or negative).
In this study we propose an extension of the WQS where two indices, one looking to the positive and the second to the negative direction, will be built in the same model both at the weights computation and at the final model estimation step.

Methods

The general formula for the WQS regression model where c components are included in the mixture is the following:

$$g(\mu) = \beta_0 + \beta_1 \left(\sum_{i=1}^{c} w_i \, q_i \right) + \boldsymbol{z}'\boldsymbol{\varphi}$$

where w_i are the unknown weights associated to each component of the mixture to be estimated through a bootstrap, q_i are the values of the components scored into quantiles (quartiles, deciles, …) and $\boldsymbol{z}'\boldsymbol{\varphi}$ are the vector of covariates and its parameters respectively. The dataset is split in a training set, used for weight derivation, and a validation set where the mixture index effect is estimated.
The weight estimation step is performed on B bootstrap sample of the training set. The following constrains are imposed during the weights estimation step: the weights have to sum to 1 $\sum_{i=1}^{c} w_i = 1$, and have to be between 0 and 1 $0 \leq w_i \leq 1$. Once the weights are estimated for each bootstrap sample, the average of the weights associated to a β_1 with the same sign of the chosen direction of the association between the mixture and the outcome is estimated. The final model will be fitted on the validation dataset using the formula $g(\mu) = \beta_0 + \beta_1 WQS + \boldsymbol{z}'\boldsymbol{\varphi}$.
In this study we propose to included two indices in the same model to allow an estimate of the effect of the mixture both in a positive and a negative direction in the same model at the weights and regression parameters estimate. We called the new method Double Weight Quantile Sum Regression (DWQS). The new general formula will be the following:

$$g(\mu) = \beta_0 + \beta_{1p}\left(\sum_{i=1}^{c} w_{ip}\, q_i\right) + \beta_{1n}\left(\sum_{i=1}^{c} w_{in}\, q_i\right) + \boldsymbol{z'\varphi}$$

where w_{ip} and w_{in} are the weights associated to each component for the positive and negative direction respectively, β_{1p} and β_{1n} are the two parameters that measure the positive and negative effect of the mixture on the outcome. The equality and inequality constraints will be applied to both the sets of weights besides a constraint on each β_1 parameter: $\beta_{1p} \geq 0$ and $\beta_{1n} \leq 0$.
The final model will be of the form $g(\mu) = \beta_0 + \beta_{1p}WQS_p + \beta_{1n}WQS_n + \boldsymbol{z'\varphi}$.
We applied this model on a dataset from a previous study on the Mexican Programming Research in Obesity, Growth, Environment, and Social Stress (PROGRESS) cohort where separate WQS models were used to test for a positive and negative association between a mixture of nutrients and some neurocognitive outcomes [6]. We performed a simulation study to test the improvements in the method when accounting for the double directionality in the association between the mixture and the dependent variable in the same model. A hundred datasets were generated using the variance and covariance matrix of the nutrient variables measured through a food frequency questionnaire administered to the parents when children were 4-6 years old.

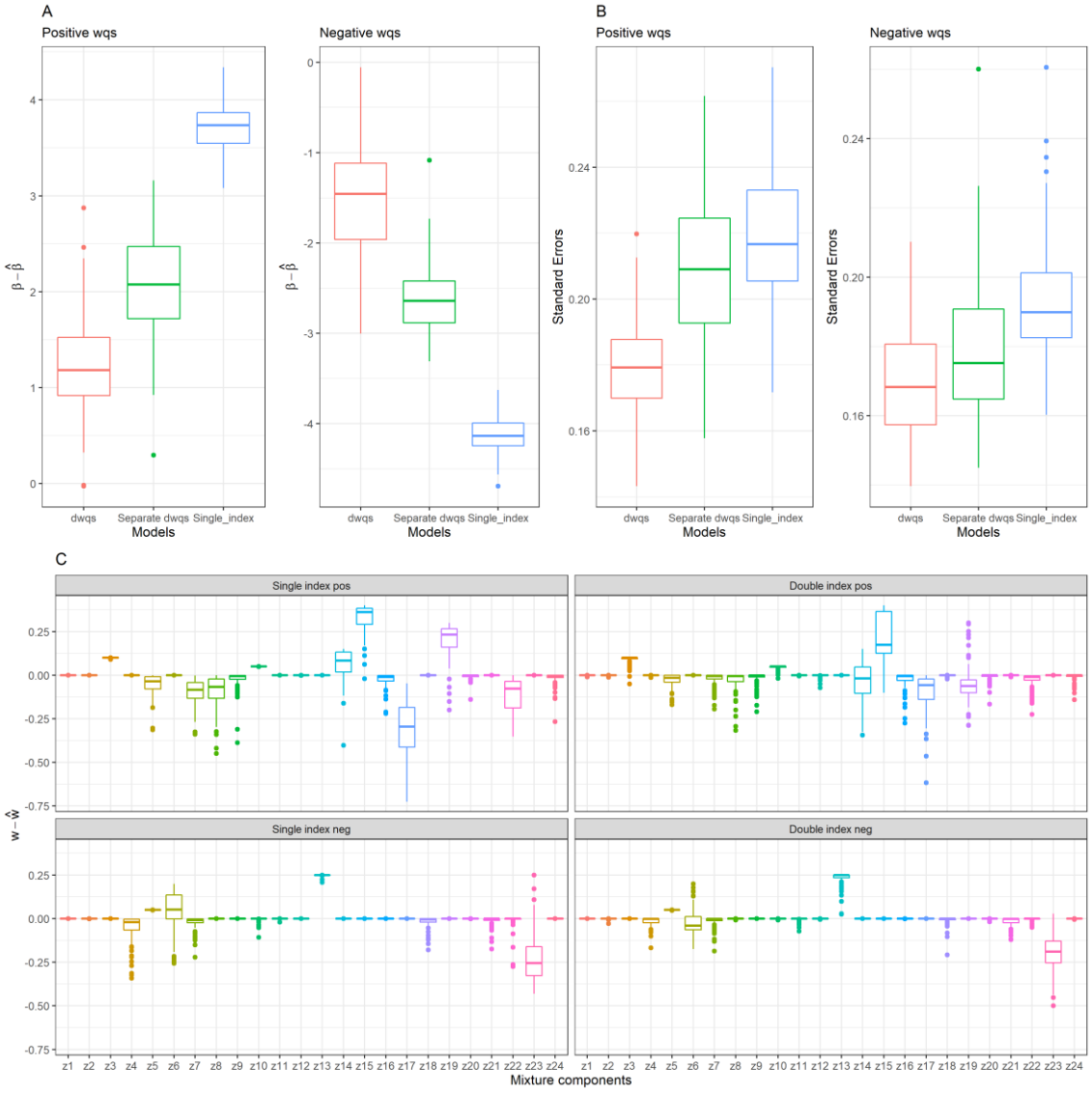
Results

Among the 24 nutrients included in the mixture five variables were selected in the positive and four in the negative association. The following formula was used to generate the dependent variable:

$$y = 2 + 5(0.4z_{15} + 0.3z_{19} + 0.15z_{14} + 0.1z_3 + 0.05z_{10}) - 5(0.5z_{23} + 0.25z_{13} + 0.2z_6 + 0.05z_5)$$

Figure 1 shows three plots that allow to compare the different methods. Figure 1A shows the bias in the beta estimates both for the positive and negative index ($\beta_{1p} - \hat{\beta}_{1p}$ and $\beta_{1n} - \hat{\beta}_{1n}$ respectively) for the DWQS model, for the WQS regression when the double index was taken into account in the beta's estimate and when looking at the two directions in two separate models. The boxplots show a clear improvement of the estimate and lower standard errors (figure 1B) for the DWQS. Moreover, we can also appreciate a better performance in the estimate of the single weights attributed to each nutrient (figure 1C).

Figure 1. *Boxplots of the difference between the real beta and the estimate value associated to the positive and negative index for the DWQS, the WQS model where the two indices were included in the parameter regression estimates and the two WQS models with a positive and negative index (A); the standard errors of the betas for the three models (B); the difference between the real value of each weight and the estimate in the WQS and the DWQS (C).*



Conclusions

The DWQS regression showed an improvement in the estimate of the regression coefficients associated to the WQS indices and the weight estimates related to each nutrient included in the mixture. The standard errors of the regression parameters were also lower compared to the WQS models where we either kept the two indices in the estimate of the final model or we fit two separate models for the two directions. Overall, when there is a potential double directionality in the association between the mixture and the outcome,

DWQS showed a lower bias in the estimates of the parameters and higher precision compared to the standard WQS regression.
This analysis suggests that in the presence of a positive and a negative association between the mixture and the outcome it is recommended to keep in the model the two indices both when the weights and the parameters of the model are estimated.

Bibliography

[1] Carlin D. J., Rider C. V., Woychik R. et al., Unraveling the Health Effects of Environmental Mixtures: An Niehs Priority. Environ Health Perspect, 2013 Jan; 121 (1): A6–8.

[2] Carpenter D., Arcaro K. and Spink D. C., Understanding the Human Health Effects of Chemical Mixtures. Environ Health Perspect, 2002 Feb; 110(Suppl 1): 25–42.

[3] Patel C. J., Analytic Complexity and Challenges in Identifying Mixtures of Exposures Associated with Phenotypes in the Exposome Era. Curr Epidemiol Rep, 2017 Jan; 4(1): 22–30.

[4] Carrico C., Gennings C., Wheeler D. C. et al., Characterization of Weighted Quantile Sum Regression for Highly Correlated Data in a Risk Analysis Setting. J Agric Biol Environ Stat, 2015 Mar; 20(1):100-120.

[5] Czarnota J., Gennings C. and Wheeler D. C., Assessment of Weighted Quantile Sum Regression for Modeling Chemical Mixtures and Cancer Risk. Cancer Inform, 2015 May; 14(Suppl 2):159-71.

[6] Malin A.J., Busgang S. A., Cantoral A. J. et al., Quality of Prenatal and Childhood Diet Predicts Neurodevelopmental Outcomes among Children in Mexico City. Nutrients, 2018 Aug; 10(8).

APPLICATION OF A CART ANALYSIS TO STRATIFY PROGNOSIS OF GRADE II GLIOMAS

Battistella Claudio[1], Ius Tamara[2], Cesselli Daniela[1,3], Pegolo Enrico[3], Skrap Miran[2], Polesel Jerry[4], Isola Miriam[1]

[1] *Department of Medicine, University of Udine, Udine, Italy.*

[2] *Neurosurgery Unit, Department of Neurosciences, University Hospital of Udine, Udine, Italy.*

[3] *Institute of Anatomic Pathology, University Hospital of Udine, Udine, Italy.*

[4] *Unit of Cancer Epidemiology and Biostatistics, CRO Aviano-National Cancer Institute, IRCCS, Aviano, Italy.*

Introduction

Recently it has been shown that extent of resection (EOR) and molecular classification of low-grade glioma (LGG) are endowed with prognostic significance. [1-2] However, a prognostic stratification of patients able to give a specific weight to the different parameters, singularly able to predict patient prognosis, is still missing.
Cox multivariate regression model is a useful tool to identify the predictive variables characterizing patients with a high or low risk of events. Unfortunately, in the clinical context a lot of variables may be continuous, making difficult to assign a score to patients based on hazard ratios in order to predict prognosis. A solution could be turning continuous variables into ordinal variables, but the cut-off choice is arbitrary.
Another problem that could afflict a multivariate Cox model is that it does not take into account the interaction between variables. CART analysis may offer an alternative method to create a predictive score considering the interaction between variables. [3] In this study, we adopted CART analysis to define a prognostic score of grade II glioma patients.

Aim

The aim of this study was to generate a score able to predict the time-to-death risk in a low-grade glioma population applying a CART analysis.

Methods

We considered 241 adults who underwent surgery for a supratentorial low-grade glioma (LGG). CART analysis was used to generate a score predicting the time to death. We considered all clinical, neuroradiological and histopathological variables with clinical relevance. Overall survival was described using the Kaplan-Meier approach. Differences in terms of overall survival probability among the scoring categories were investigated using log-rank test. The performance of the score in predicting time to death was estimated through Harrell's c-index.

Results

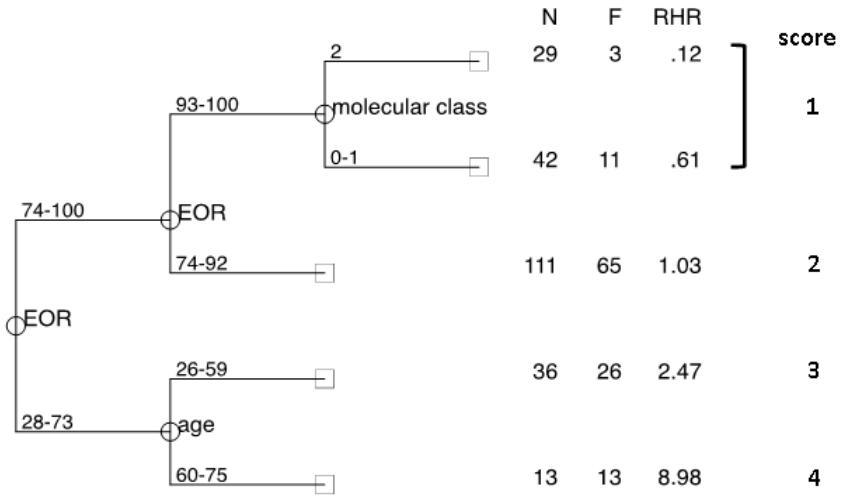
Factors included in the CART analysis were: age, Karnofsky performance status (KPS), pre-operative tumor volume (T2 pre), infiltrative growth index ΔVT2T1, both extent of resection (EOR) and post-operative tumor volume (T2 post), expression of Ki67, tumor molecular class and methylation status of the MGMT promoter (MGMTm). The score was created using the CART analysis for failure time data resulting in 5 groups based on EOR, age and tumor molecular class.

The resulting tree was trimmed according to similarities in groups' relative hazard ratios, leading to a final 4-group classification tree. The 4 groups were different in terms of overall survival with p<0.001. The performance of the score in predicting death was a Harrell's c-index 0.73 (95%IC [0.68;0.78]).

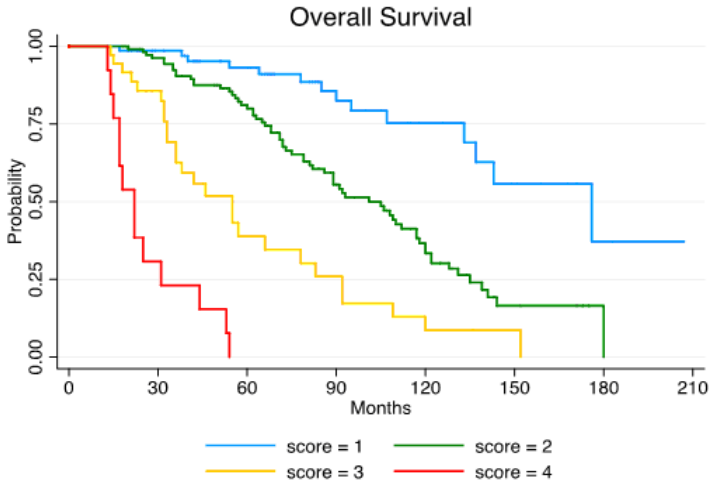
Conclusion

CART analysis for time failure data may be useful to generate a well performing predictive score taking into account the presence of continuous variables in the model.
A risk score, which considered a low number of variables may be useful for clinical application to improve the management of the disease.

Figure 1. *a) Classification and regression tree analysis. Split if p (adjusted <0.05). Variables: age, KI67, molecular class, KPS, T2 pre, ΔVT2T1, EOR, T2 post, MGMTm. b) Kaplan-Meier curves estimation for overall survival by score classification.*



a.



b.

References

[1] Ius T., Isola M., Budai R. et al. Low-grade glioma surgery in eloquent areas: volumetric analysis of extension of resection and its impact on overall survival. A single-institution experience in 190 patients: clinical article. J Neurosurg., 2012; 117(6):1039-1052.

[2] Clark V.E., Cahill D.P. Extent of Resection Versus Molecular classification: what matters when?. Neurosurg Clin N Am. 2019; 30(1):95-101.

[3] Breiman L. Classification and regression trees. New York: Routledge; 1984.

IMPACT OF TIME-DEPENDENT MEDICATION ADHERENCE ON HEART FAILURE PATIENTS USING A JOINT MODELLING FRAMEWORK

Spreafico Marta[1,2], Ieva Francesca[1,2,3]

[1] *MOX - Modeling and Scientific Computing lab, Department of Mathematics, Politecnico di Milano, Italy*

[2] *National Centre for Healthcare Research & Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy*

[3] *CADS – Center for Analysis Decisions and Society, Human Technopole, Milano, Italy*

Introduction

Heart failure (HF) is a chronic cardiac disease which represents a leading and increasing cause of morbidity and mortality worldwide [1]. Different studies proved that pharmacological treatments in HF patients could improve their clinical status, functional capacity and quality of life, prevent hospital admission and reduce mortality [2]. In pharmacoepidemiology, medication adherence is generally defined as the process by which patients take their medications as prescribed [3]. Medication non-adherence is a common issue and it is associated with adverse health conditions and increased economic burden to the healthcare system, especially in case of chronic diseases such as HF [4]. In the past decade there has been substantial growth in adherence research, partly owing to increasing awareness of the problem and partly because of the pervasiveness of non-adherence [5]. This was also possible thanks to the increasingly intensive growth and use of healthcare administrative registries, which include data about many different healthcare services (prescriptions of drug medicaments, medical visits, laboratory and instrumental examinations, hospitalizations, etc) [6].
Adherence was usually computed on a predefined time-period and analysed as a fixed baseline covariate. However, in healthcare analytics associating functional responses (e.g. biomarkers) with an event-time outcome related to a specific event of interest (e.g. treatment failure or death) represents a challenging task that could be tackled using Joint Models for longitudinal and time-to-event data [7]. Therefore, our idea was to exploit healthcare databases in order to reconstruct the dynamics of drugs consumption and adherence through pharmaceutical administrative registries and to investigate its impact on survival in HF patients. In fact, since the costs associated with HF management are high [1], studying HF condition using a time-dependent variable for medication adherence can lead to healthcare improvements, social benefits and economic utilities.

Aims

This work had two different purposes. Firstly, we computed a time-varying variable for drug adherence during the first year of follow-up, which resulted more informative than a fixed dichotomous variable. Secondly, we wanted to analyse the dynamics of drug consumption using a joint modelling framework for longitudinal and time-to-event data, in order to explore the effects of time-dependent adherence to pharmacological therapies on patients' survival outcomes.

Methods

In the Lombardy Region dataset "new incident" patients (i.e. patients at their first HF hospitalization) hospitalized from 2000 to 2012 with a primary diagnosis of HF were identified, as described in [1]. For our work, we focused on a representative sample of patients with the first HF discharge between 2006-2012, excluding patients who died during the index hospitalization. Since we were interested in long-term mortality, we excluded those patients who died within 30 days of follow-up. Moreover, since we wanted to analyse the

effect of 1-year adherence to evidence-based (EB) drugs recommended for HF [2] on patients' survival, we considered only patients with at least one purchase of EB pharmacological treatments in his/her first year of follow-up. In particular, in this work we reported only analyses related to Angiotensin-Converting Enzyme inhibitors (ACE) or Angiotensin Receptor Blockers (ARB) therapy.
For each cohort member, identified by its unique anonymous ID code, data included gender, age and total number of comorbidities at the index hospitalization. Comorbidities were evaluated with the method proposed by [8]. Each patient was followed from the index date until the end of the study or the date of death. Administrative censoring date is December 31st, 2012. Each record in the dataset was related to an event, which could be a hospitalization or a drug purchase. In the first case, the date of discharge from hospital and the length of stay in hospital were given. For drug purchases, identified by their Anatomical Therapeutic Chemical (ATC) codes, the date of purchase and the number of days of treatment covered by the prescription, based on the number of boxes and the Defined Daily Dose (DDD) for that specific medicinal product, were given.
To compute the time-dependent variable of adherence to ACE/ARB therapy, we started from the definition of one of the most used adherence measures, the Proportion of Days Covered (PDC) [3,4]. PDC definition is founded on prescribed drugs on a fixed study period. Our analysis was based on purchased drugs instead of prescribed drugs [9] and the study period varied among time. In particular, for each month and for each patient i we computed adherence using the PDC definition up to time t:

$$PDC_i(t) = \frac{number\ of\ distinct\ coverage\ days\ up\ to\ time\ t}{number\ of\ days\ up\ to\ time\ t}$$

For the computation, we hypothesize that ACE/ARB therapy was assumed by patients during the whole periods of re-hospitalization. Hence, for each patient i we obtained a monthly time-dependent variable which represented the proportion of months covered up to month t.
Finally, for each patient i and for each month t, we defined a dichotomous time-dependent variable for patient's adherence to therapy using an 80% threshold [3]:

$$adherent_i(t) = \begin{cases} 1 & if\ PDC_i(t) \geq 0.80 \\ 0 & if\ PDC_i(t) < 0.80 \end{cases}$$

To analyse the effect of time-dependent adherence on patients' survival we used the joint modelling framework introduced by Rizopoulos in 2012 [7,10,11]. Joint models are used in follow-up studies where interest is in associating an endogenous longitudinal response with an event time outcome (in our case the death of the patient). The key point of endogenous covariates [12] is that their existence and future path are directly related to the event status. In our case, since therapies were usually modified according to the development of the illness, adherence to the therapy was an endogenous covariate related to the behaviour of the individual over time, hence its time path were jointly determined with the responses of interest.
Among the possible joint modelling frameworks for longitudinal and event processes offered in `JMbayes` package [11], we decided to analyse how our time-dependent variable, which indicates the adherence to ACE/ARB therapy up to time t, affect patients' survival using the following submodels:

- *Longitudinal submodel:* we postulated a logistic mixed effect model

$$\log \frac{\Pr[y_i(t) = 1]}{1 - \Pr[y_i(t) = 1]} = \eta_i(t) = \beta_0 + b_{i0} + (\beta_1 + b_{i1})\, t$$

where:

- $y_i(t)$ denotes the value of the longitudinal outcome for the i-th patient at time point t
- $\boldsymbol{\beta}$ is the vector of the unknown fixed effect parameters
- $\boldsymbol{b}_i$ is the vector of random effect for the i-th patient, with $\boldsymbol{b}_i \sim N(0, \boldsymbol{D})$
- $\eta_i(t)$ is the subject-specific linear predictor for the i-th patient at time point t

- *Event submodel:* we assumed that the risk of death depends on the current true value of the linear predictor at time t, adjusting according to age, gender and number of comorbidities at index hospitalization:

$$h_i(t) = h_0(t) \exp\{ \gamma_1 age_i + \gamma_2 gender_i + \gamma_3 comorbidities_i + \alpha \eta_i(t) \}$$

where:

- $h_0(t)$ denotes the baseline hazard function at time t, modelled using a B-spline approach
- $\boldsymbol{\omega}_i = (age_i, gender_i, comorbidities_i)$ is the vector of exogenous baseline covariates with corresponding regression coefficients $\boldsymbol{\gamma}$
- α is the parameter that quantifies the association between the marker process, i.e. the subject specific linear predictor, up to time t and the hazard of the event of interest at the same time point

The estimation of the joint models' parameters proceeded under a Bayesian approach, using Markov Chain Monte Carlo algorithms [13]. See [11] for further details.

Results

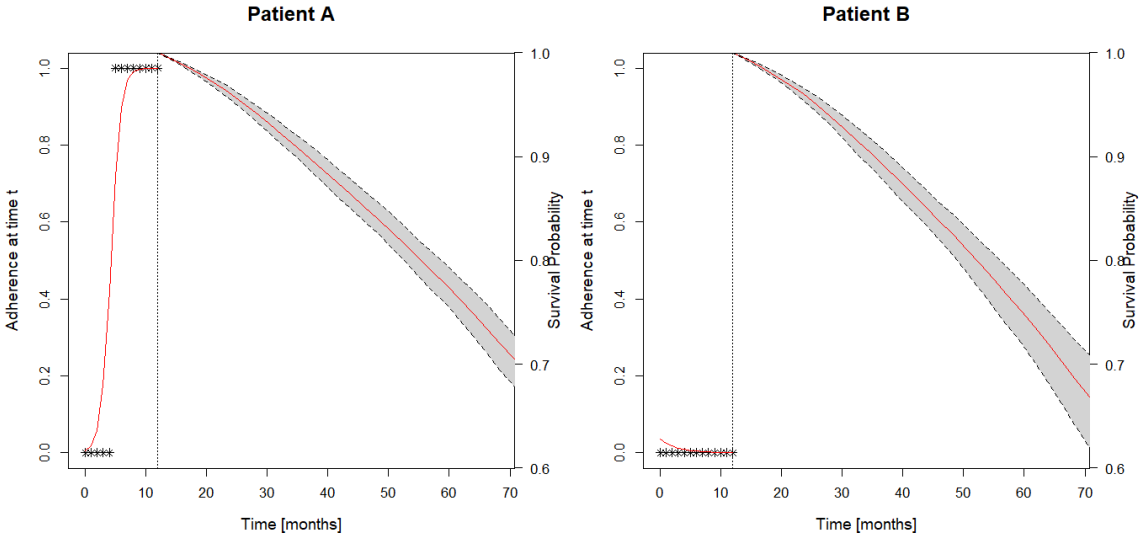
Starting from a representative sample cohort of 4,870 patients with principal diagnostic code of HF during the period 2006-2012, we identified 3,974 (81.6%) patients that purchased at least one box of ACE or ARB in their first year of follow-up and met all the study selection criteria.
Overall, at index hospitalization, mean age was 72.82 years with a substantial proportion of male patients (55.8%). Comorbidity burden was high (median of total number of initial comorbidities 2; 30.5% of patients presenting a total number of initial comorbidities ≥ 3). During a median follow-up of 48.85 (IQR 30.72-66.94) months, 1,012 (25.6%) patients died.
From the longitudinal process, we observed that patients with an overall adherence trend during time had positive and increasing values of their linear predictors over time, $\eta_i(t)$, whereas patients with non-adherence trends presented negative and decreasing values of the linear predictors.
From the joint modelling analysis, it resulted that all the covariates were associated with the risk for the composite event. In particular, being younger or a female corresponded to a higher survival probability, whereas having a higher number of initial comorbidities corresponded to a lower survival probability, as it might be expected. Moreover, the value of the longitudinal linear predictor $\eta_i(t)$ had a protective role, i.e. the higher the value of the linear predictor the higher the survival. Hence, having an adherence trend during time reflected a protective role on patients' survival. Furthermore, as shown in Figure 1, we also observed that patients with a non-adherent trend during the first year (right panel) had larger confidence intervals over time, so the uncertainty about the prediction of the survival outcome for non-adherent patients increased.

Conclusions

In this work we observed that modelling the drug assumption process as time-varying covariates in a joint modelling setting represents an interpretative and forecasting tool for exploring the effects of

pharmacological treatments on patients' survival. In fact, when the outcome processes are correlated, joint modelling has empirically demonstrated to reduce biases, improve efficiency and prediction and they can be applicable to outcome surrogacy [14]. Moreover, they can provide subject-specific predictions and constitute an excellent tool for personalized medicine [11].
Some improvements may be included into the model in order to simultaneously combine all the EB therapies for HF (ACE/ARB, Beta Blocking agents, Anti Aldosterone agents, Diuretics) using a multivariate joint modelling framework.
Nevertheless, since medication nonadherence is usually associated with adverse health conditions and increased costs [4], joint models for pharmacoepidemiology represent a highly stimulating area that could also have a strong impact on the society, leading to new answers in the field of policy-making and in the setting of healthcare databases.

Figure 1. *Longitudinal trajectories (left side of both panels) and survival probability plots (right side of both panels) for two male patients with 72 years old and two initial comorbidities. Patient A (left panel) present an adherence trend with* $y_A(12) = 1$ *and* $\eta_A(12) = 9.61$. *Patient B (right panel) present a non-adherence trend with* $y_B(12) = 0$ *and* $\eta_B(12) = -7.65$. *Patient B has lower survival and larger confidence interval over time.*



Bibliography

[1] Mazzali C., Paganoni A.M., Ieva F. et al., Methodological issues on the use of administrative data in healthcare research: the case of heart failure hospitalizations in Lombardy region, 2000 to 2012. BMC Health Serv Res, 2016; 16(234)

[2] Ponikowski P., Voors A.A., Anker S.D. et al., 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J, 2016; 37(27):2129–2200.

[3] Andrade S.E., Kahler K.H., Frech F. et al., Methods for evaluation of medication adherence and persistence using automated databases. Pharmacoepidemiol Drug Saf 2006; 15:565-574.

[4] Karve S., Cleve M.A., Helm M. et al., Prospective Validation of Eight Different Adherence Measures for Use with Administrative Claims Data among Patients with Schizophrenia. Value In Health 2009;12(6).

[5] Vrijens B., De Geest S., Hughes D.A. et al., A new taxonomy for describing and defining adherence to medications. Br J of Clin Pharmac, 2012 May; 73(5):691-705.

[6] Corrao G., Mancia G., Generating Evidence From Computerized Healthcare Utilization Databases. Hypertens, 2015; 65:490-498.

[7] Rizopoulos D., Joint Models for Longitudinal and Time-to-Event Data: With Applications in R. Chapman and Hall, 2012.

[8] Gagne J.J., Glynn R.J., Avorn J. et al., A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol, 2011; 64:749-59.

[9] Spreafico M., Gasperoni F., Barbati G. et al., Adherence to disease-modifying therapy in Heart Failure patients: a community-based study. 2019. [submitted]

[10] Rizopoulos, D., JM: An R Package for the Joint Modelling of Longitudinal and Time-to-Event Data. J Stat Softw, July 2010; 35(9):1-33.

[11] Rizopoulos, D., The R Package JMbayes for Fitting Joint Models for Longitudinal and Timeto-Event Data Using MCMC. J Stat Softw, Articles, 2016; 72(7):1-46.

[12] Kalbfleisch J.D., Prentice R.L., The statistical analysis of failure time data. John Wiley & Sons, 2011; 360.

[13] Jackman S., Bayesian analysis for the Social Sciences. Wiley, 2009.

[14] Hickey G., Philipson P., Jorgensen A. at al., Joint Models of Longitudinal and Time-to-Event Data with More Than One Event Time Outcome: A Review. Int J Biostat, 2018; 14(1).

TRAJECTORIES IN COGNITIVELY HEALTHY ITALIAN ELDERS

Ferraro Ottavia Eleonora[1], Guaita Antonio[2], Villani Simona[1]

1 Unit of Biostatistics and Clinical Epidemiology, Department of Public Health, Experimental and Forensic Medicine University of Pavia

2 Golgi Cenci Foundation, Corso San Martino 10, 20081 Abbiategrasso

Introduction

Life expectancy is increasing and the number of people aged more than 65 is constantly growing: in 2017 was estimated that in 2050 there will be 2.1 billion people in the world will have almost 60 years and 434 million of these persons will have 80 or more years [1].
Research activity, actually, is focalizing to find all the possible modifiable risk factors, in order to improve the prevention of diseases due to aging, like dementia. These pathologies bring to loss memory and other mental disabilities, and to interfere with daily life [2]. Much less is known about trajectories of cognitive health, physical functioning and disability on elderly people without diagnosis of dementia.

Aims

The primary aim was to identify multi-trajectories trend of cognitive, physical functioning and disability on non-demented elderly people living in North of Italy. The secondary aim was profiling each multi-trajectories group identified, in order to understand which are the main differences that characterized subjects with different trends.

Methods

Data were retrieved from the ongoing prospective population based study "InveCeAb", involving all elderly people born between 1935 and 1939 (1644 eligible at baseline) and living in the Abbiategrasso town located near to Milan in the northern of Italy. The study was settled in 2010 and had 2 follow-up (f-up) in 2012 and 2014 [3]. From the 1321 (80.4%) responders at baseline, 1245 individuals were eligible at 2nd follow-up assessment (89.5%). In 2014, 1061 still eligible (living, who had participated in the baseline and 1st follow-up assessment or only in the baseline assessment) were invited to attend a further follow up: 988 of them accepted.

Outcomes. The cognitive impairment was measured by Mini Mental Score (MMSE), disability by six of Basic Activity of Daily Life score (B-ADL) and physical functioning through walking test.

Statistical Analysis. Quantitative variables were summarized by mean and standard deviation or median and IQR, qualitative ones by percentage.
The trajectories were implemented using the group-based trajectory models (GBTM) [4]. This methodology is based on finite mixture model and identifies different groups of people with the same trajectories on multiple outcomes.
Profiles, on groups identified, were explored with a multinomial logistic regression. The characteristic investigated are gender, age at baseline, ApoE-ε4, education level, obesity. The RRR with 95% Confidence Interval (95%CI) were reported. A p-value less than 0.05 was considered significant. The analyses were made by STATA ® 15.

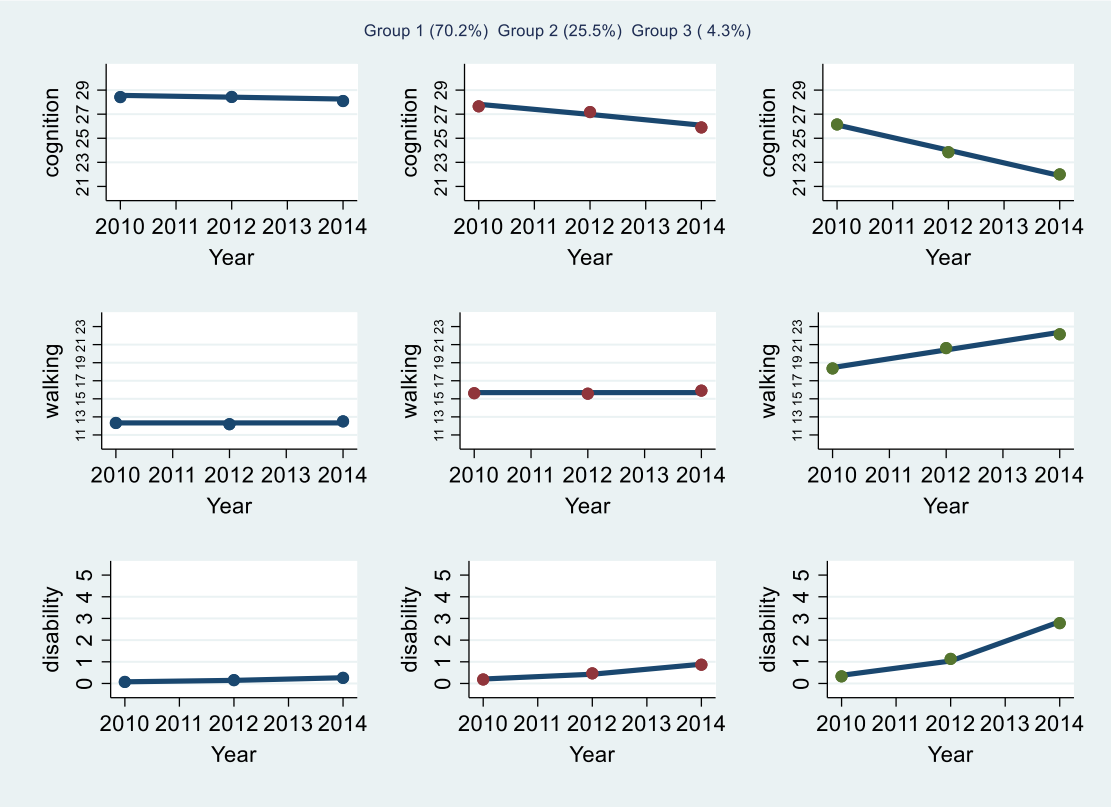
Results

Among the 993 subjects without dementia and with almost two measures for each outcome in the f-up time three subgroups of older adults with different patterns on cognitive, physical and disability trend were identified. More precisely the first group showed the better trend from all of the outcomes during the f-up time as reported in figure 1 and included the largest number of elderly (n=703). The 53.2% of elders with the best trajectories were males, the mean age at baseline was 72.7 years (± 1.4 years) and, in average, they took 3.3 drugs regularly (± 2.2). The second group (n=253) showed the "middle" trend due to a decline in cognition (MMSE score from 28 points to 25), but with constant speeding to walk for 10 metres around 16 seconds. This group with "middle" trajectories consisted of about 65% females, with a mean age, at baseline, of 72.9 years (±1.4 years). Concerning the disability, people classified in this group needed a help in one of daily living activities at the end of the study and took regularly, in average, 4.2 drugs (±2.3). The last group displayed a strong decline in cognitive function, with an average annual decrease of 1 point in MMSE score, an increase in time to perform the walking test and more than one difficulties in daily living activities. This group involved a small number of elderly (n=42) and was composed in large part by women (76%), slightly older than previous two groups (mean age 73.4 ± 1.2 years) and with similar use of drugs 4.3 (±2.2). The percentages of demented people were very different across three groups: increasing from the best group (1%) to the middle (5.7%) to the worst group (36%). Concerning the distribution of the ApoE-ε4 carrier status, there were no differences across the three groups (p=0.821). The profile of subjects in the three groups was different for age, gender, educational level and obesity. The risk to be older was higher only in the group with the worst trajectories in comparison with the best group, more precisely was 1.4 times greater (CI95%: 1.1-1.8 p=0.005). In the worst group the risk to be a female was 2.9 (CI95%: 1.4-6.2 p=0.005) times higher in comparison with the best group and 1.8 times greater (CI95%: 1.3-2.5 p<0.001) than in the middle group. Also education level was heterogeneous among groups: if the level of education was higher, the risk of being in the worst group was lower than that of being in the best group (0.63 [CI95%: 0.53-0.75] p<0.001), similarly the risk of being in the middle group in comparison with that of being in the best group was inversely related to level of education (0.84 [CI95%: 0.79-0.89] p<0.001). About the clinical profile, only the obesity showed a relevant difference across the three groups with different trajectories on cognitive status, disability and functional activity. Precisely, the risk to be obese in the group with worst trajectories and in the middle groups was 3.5 times greater (CI95%: 1.7-7.1 p=0.001) and 2.1 times (CI95%: 1.5-3.1 p<0.001) respect to those found in the best group, respectively.

Conclusion

Three groups with different trajectories of cognitive status, physical function, and disability in daily life activities were identified among an Italian elders' population living in the norther part of Italy, homogenous in ages (71-75 years). The large part of the people had a stable and good cognitive status, good function activity and no impairment on daily activities. People with the worst trajectories presented an important decline in cognitive function, disability and functional activity: the cognitive decline probably drives an increase in disability, bringing a worse quality of life. The proportion of people with dementia was higher, around 1/3, in the worst group in comparison with the other two groups. In the worst and in the middle group also there was a high presence of people with obesity. This finding agrees with the published evidences on relationship between obesity and cognitive health decline [5]. Similarly the association with well-known protective factor such as the education level was confirmed [6].

Figure 1. *Trajectories among elderly Italian people: the group 1 trajectories are depicted in the graph with blue dots, the group 2 trajectories in those with brown dots, and group 3 in graphs with green dots.*



Bibliography

[1] United Nations, Department of Economic and Social Affairs PD (2017). World Population Ageing 2017. *Department of Economic and Social Affairs* 2017.

[2] Alzheimer's-Association. Dementia Types | Symptoms, Diagnosis, Causes, Treatments. https://www.alz.org/alzheimers-dementia/what-is-dementia/types-of-dementia. Accessed April 26, 2019.

[3] Guaita A, Colombo M, Vaccaro R, et al. Brain aging and dementia during the transition from late adulthood to old age: Design and methodology of the "invece.Ab" population-based study. *BMC Geriatrics* 2013; 13(1):1. doi:10.1186/1471-2318-13-98.

[4] Nagin DS. *Group-based Modeling Development*.; 2005.

[5] Dye L, Boyle NB, Champ C, Lawton C. The relationship between obesity and cognitive health and decline. *Proceedings of the Nutrition Society* 2017; 76(04):443–454. doi:10.1017/s0029665117002014.

[6] Meng X, D'Arcy C. Education and dementia in the context of the cognitive reserve hypothesis: a systematic review with meta-analyses and qualitative analyses. *PloS one* 2012; 7(6):e38268. doi:10.1371/journal.pone.0038268.

A COMPARISON OF DIFFERENT STATISTICAL APPROACHES TO DEAL WITH MODEL MISSPECIFICATION AND MISSING OUTCOME DATA

Sciannameo Veronica[1], Fadini Gian Paolo[2], Bottigliengo Daniele[1], Avogaro Angelo[2], DARWIN-T2D Network[3], Baldi Ileana[1], Gregori Dario[1], Berchialla Paola[4]

[1] *Unit of Biostatistics, Epidemiology and Public Health, University of Padova, 35128 Padova, Italy*

[2] *Department of Medicine, University of Padova, 35128 Padova, Italy*

[3] *See the Acknowledgement section for composition of the network*

[4] *Department of Clinical and Biological Sciences, University of Turin, Italy*

Introduction

In observational biomedical research, non-randomized treatment assignment and missingness in the outcome measures contribute to bias the treatment effect estimate.
Moreover, treatment groups differ in baseline characteristics and an appropriate statistical adjustment for confounding variables is necessary, to obtain an unbiased causal inference. Rosenbaum and Rubin introduced propensity score (PS) methods, which allow to estimate the treatment mechanism [1]. PS is defined as the probability to be assigned to a treatment Z, conditionally to baseline covariates, denoted P (Z = 1|W). PS allows to balance baseline covariates between two different treatment groups, via matching, inverse probability of treatment weighting (IPTW) or stratification [2], [3].However, PS methods can be sensitive to misspecification on both treatment and outcome models [4].
Another challenge in PS methods is missing values. The high percentage of missing values, both in covariates and outcome data, is an issue of Real-World Data (RWD) [5], [6]. When the missingness is on the covariates values, it has been shown that the most advantageous method is multiple imputation (MI) [5]. MI replaces missing values with estimates computed on observed information. However, MI is not recommended if missingness is on the outcome variable [7]. In this case, the most common approach is a complete case (CC) analysis [8]. On the other hand, this strategy is valid only if missing values are missing completely at random (MCAR), because only in this situation the subjects with complete information are a representative subsample [7]. Otherwise, if excluded subject differs systematically from those included, CC analysis could entail biased estimates [7].
In the last decades, several double-robust methods has been developed for handling the impact of model misspecification [9] [10]. In 2006, Van Der Laan and Rubin introduced the Targeted Maximum Likelihood Estimator (TMLE), which allows to deal simultaneously with confounding, model misspecification and missing outcome data.
TMLE computes initial estimates of the outcome regression E (Y| Z, W) and of the propensity score model PS P(Z = 1| W). Then, the initial outcome regression is updated, targeting the parameter of interest, via the PS estimate on a logit scale. Both the initial and the updated estimates could be estimated via a machine learning ensemble method called Super Learner (SL) that finds the optimal combination of a collection of algorithms, in order to minimize the cross-validated risk. A mathematically demonstrated theorem states that SL performs asymptotically as well as the oracle selector, i.e. the best candidate between learner algorithms [11], [12]. TMLE returns also the coefficients that reflect each prediction algorithm's contribution in SL [12].
Among the advantages of the TMLE, is the ability to deal with missing data on outcome variable via the estimation of the missingness mechanism, $\pi (\Delta = 1 \mid Z, W)$, where $\Delta = 1$ means that the outcome is observed and $\Delta = 0$ that it is missing [13].

Aim

The aim of this study is to compare different statistical approaches to limit the size of the bias in the estimate of the treatment effect, due to model misspecification, non-randomized treatment assignment and missing not at random (MNAR) binary outcome measures, in different scenario settings, both analyzing RW and simulated data.

Methods

In this study, we compared covariates adjustment, PS adjustment, PS matching, IPTW and TMLE, in terms of bias and standard error. Then, comparisons were performed simulating two different scenarios: with 20% and 40% MNAR mechanism on the binary outcome. In both scenarios, both treatment and outcome models were misspecified and sensibility to sample size was addressed.
Logistic regression (LR) and PS analyses were performed following the CC approach. Two TMLE analyses were performed. In the first (TMLE1), SL includes only default algorithms (main terms logistic regression (LR), stepwise forward and backward model selection, main terms LR and interaction terms). In the second (TMLE2), SL includes the following algorithms: main terms LR, stepwise forward and backward model selection, main terms LR and interaction terms, stepwise forward and backward model selection with interaction, generalized additive models (Gam), Random Forest (RF) and recursive partitioning and regression trees (Rpart). TMLE analyses were performed both under CC and IPW approach to deal with missing data.

RWD case study. All the methods were used to analyze RWD from the Italian multicenter retrospective DApagliflozin Real Word evidence in Type 2 Diabetes (DARWIN-T2D) study.
DARWIN-T2D is a retrospective multicenter nationwide descriptive study conducted in 46 Italian diabetes outpatient clinics. The main aim is to describe the baseline clinical features in patients initiated on dapagliflozin as compared to patients initiated on comparator glucose lowering medications (GLM) as DPP-4 inhibitors, glucagon-like peptide-1 receptor agonist (GLP-1RA). The secondary aim is to compare the changes in glycemic efficacy parameters, between patients initiated on different medications [14].
Based on previous RCTs, some GLP-1RA medications are recommended for secondary cardiovascular prevention in T2D [15] [16]. The STENO-2 study, makes clear that the simultaneous control of multiple risk factors, as the reductions in the glycated hemoglobin (HbA1c), body weight (BW), and systolic blood pressure (SBP), could drastically improve micro- and macrovascular outcomes in T2D patient [17]. In [18] T2D patients initiated on dapagliflozin or GLP-1RA medication were compared evaluating the proportion of patients with simultaneous reduction of HbA1c >0.5%, of BW >2 kg, and of SBP >2 mm Hg.
Multiple Imputation by Chained Equation algorithm was applied to obtain a MI of missing covariate data, obtaining 5 imputed datasets. Only covariates with less than 40% of missing values were included as predictors in the imputation process, including observed outcome values [19] [20]. Outcome variable was not imputed. In each imputed dataset we computed PS considering the following baseline covariates: age, gender, duration of diabetes, BW, Body Mass Index (BMI), fasting plasma glucose (FPG), HbA1c, systolic and diastolic pressure (SBP and DBP), total and HDL cholesterol, triglycerides, eGFR, insulin and metformin therapy, microangiopathy and macroangiopathy. PS matching was performed with nearest 1:1 ratio without replacement and with a caliper of 0.15 standard deviations of the distribution of PS on the logit scale [21].
Outcome analyses were performed in each imputed database, then results were pooled together following the Rubin's Rules [5]. This method is referred to as the within method [22].

Simulation. Data were simulated defining a Directed Acyclic Graph (DAG), using the "simcausal" R package [23]; [24]. The variables included in the DAG were simulated according to the RWD pattern observed in

DARWIN-T2D in the following covariate: age, sex, low-density lipoprotein (LDL) cholesterol, macro-angiopathy, BMI, insulin use, binary treatment (Z) and binary outcome (Y).
Age was simulated following a gaussian distribution, with mean = 60 and standard deviation (SD) = 8; BMI with a gaussian distribution with mean = 35 and SD = 6; LDL Cholesterol with a gaussian distribution with mean = 90 for female and 97 for male and SD = 30; Sex with a Bernoulli distribution with probability 0.6; Insulin use with a Bernoulli distribution with probability of success given by plogis (- 2 + 0.05 * Age); Macro-angiopathy with a Bernoulli distribution with probability given by plogis (- 12 + 0.50 * BMI - Sex); Z with a Bernoulli distribution with probability plogis (-2 + 0.05 * BMI - 0.20 * Insulin use + 0.10 * Macro-angiopathy) and Y with a Bernoulli distribution with probability = plogis (-3 + Z - 0.05 * Age + 0.05 * LDL Cholesterol - 0.80 * Sex – 0.2 * Age * Sex), where plogis is the inverse logit function, denoted as 1/ log[p/(1-p)].
We simulated MNAR mechanism on the outcome, with two different frequency: 20 and 40%. More in detail, the probability of missingness on Y depends on the Y value itself: if Y = 1 the probability of missingness was 70%.
In the simulation study, we performed the comparisons in the most common real situation: with both treatment and outcome models misspecified. More in detail in the Z model Macro-angiopathy and Age were included as covariates; meanwhile LDL cholesterol and BMI were included in the Y model. So, we supposed that Insulin use, age, sex and the interaction term between age and sex were unobserved.
The true Odds Ratio (OR) value was computed on 5,000,000 observations generated through the DAG that and turned out to be equal to 1.66.
Sensibility to sample size was addressed, performing the analyses both with n = 1,000 and n = 5,000.
Overall, 1,000 simulations were performed in order to estimate the mean bias measure and the standard error (SE) size.
Statistical analyses were performed using R version 3.4.5.

Results

In the DARWIN-T2D study, from a population of 281,217 patients with T2D, longitudinal data were collected for 17,285 patients who started new GLMs. In particular, 2,484 initiated dapagliflozin and 2,247 initiated a GLP-1RA medication. A follow-up examination data was collected for 830 patients in the dapagliflozin group and for 811 patients in the GLP-1RA group. Composite outcome data was available for 473 patients initiated to Dapagliflozin and for 336 patients initiated to GLP-1RA. Therefore, the missingness on outcome variable was about 50%.
Results of the effect of the exposure of dapagliflozin compared to GLP-1RA, obtained via different statistical approaches do not underline any differences between the two treatments. On the other hand, TMLE2 shows an OR = 1.46, with a significant 95% confidence interval (1.10 - 1.95). So, results of RWD are very sensible to the statistical approach applied in the data analysis.
In Table 1 we can find the preliminary results about simulation study, with different percentages of MNAR mechanism on outcome and different sample size (n = 1,000 and n = 5,000).
Regardless the percentages of missingness on outcome and sample size, TMLE is the approach with the smallest bias and SE. Preliminary results shows that TMLE1 is better than TMLE2. Moreover, including missingness mechanism on outcome in TMLE improves the OR estimation. LR and PS matching are the methods that show the worst performance.

Table 1. *Results of simulation study with different scenarios.*

	Complete case analysis						***With missingness mechanism on outcome***	
	LR	**PS matching**	**PS covariate**	**IPTW**	**TMLE1**	**TMLE2**	**TMLE1**	**TMLE2**
SCENARIO 1: 20% MNAR on Y, n = 1,000								
Mean OR	1.89	1.90	1.88	1.84	1.79	1.79	1.75	1.75
Mean Bias	0.23	0.25	0.23	0.18	0.13	0.13	0.09	0.09
SE	0.67	0.84	0.68	0.66	0.58	0.58	0.58	0.57
SCENARIO 2: 40% MNAR on Y, n = 1,000								
Mean OR	2.06	2.07	2.05	2.00	1.97	1.98	1.92	1.94
Mean Bias	0.40	0.41	0.39	0.34	0.31	0.32	0.26	0.28
SE	1.30	1.31	1.30	1.23	1.20	1.18	1.15	1.17
SCENARIO 3: 20% MNAR on Y, n = 5,000								
Mean OR	1.78	1.77	1.77	1.74	1.70	1.70	1.66	1.66
Mean Bias	0.12	0.11	0.11	0.08	0.04	0.04	0.002	0.006
SE	0.25	0.28	0.25	0.25	0.22	0.22	0.22	0.22
SCENARIO 4: 40% MNAR on Y, n = 5,000								
Mean OR	1.78	1.79	1.78	1.74	1.73	1.73	1.68	1.68
Mean Bias	0.12	0.13	0.12	0.08	0.07	0.07	0.02	0.02
SE	0.37	0.43	0.37	0.36	0.35	0.35	0.35	0.35

Conclusion

Preliminary results suggest that in all scenarios TMLE has the lowest bias in the treatment estimate compared to the other approaches. Moreover, the TMLE variant that also models the missingness mechanism on the outcome is preferable with respect to the variant that follows the CC approach.

References

[1] Rosenbaum P. R. and Rubin D. B., The central role of the propensity score in observational studies for causal effects, *Biometrika*, 1983, pp. 70:41-55.

[2] Guo S. and Fraser M. W., Propensity Score Analysis: Statistical Methods and Applications. Second ed. , 2015, Los Angeles: SAGE.

[3] Lunceford J. K. and Davidian M., Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study, *Stat Med,* 2004, p. 23(19):2937-2960,.

[4] James G., Witten D. , Hastie T. et al., An introduction to statistical learning: with applications in R., 2013, New York: Springer-Verlag.

[5] Rubin D., Multiple imputation for nonresponse in surveys.,1987, New York: John Wiley and Sons.

[6] Van Buuren S., Flexible imputation of missing data, Boca Raton, 2012, FL: Chapman & Hall / CRC,.

[7] Sterne J. A., White I. R. and Carlin J. B., Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls., *BMJ, 2009,* p. ;338:b2393.

[8] White I. R. and Thompson S. G., Adjusting for partially missing baseline measurements in randomized trials. Stat, *Stat. Med* , 2005, p. 24(7): 993–1007.

[9] Van Der Laan M. and Rose S., Targeted Learning: Causal Inference for Observational and Experimental Data, 2011, New York: Springer Series in Statistics.

[10] Neugebauer R. and Van Der Laan M., Why prefer double robust estimators in causal inference?, *Journal of Statistical Planning and Inference.* , 2005, p. 129(1):405-426.

[11] Van der Laan M. J. and Dudoit S., Unified cross-validation methodology for selection among estimators and a general cross-validated adaptive epsilon- net estimator: Finite sample oracle inequalities and examples., Technical report, 2003, Division of Biostatistics, University of California , Berkeley.

[12] Gruber S. and Van der Laan M., tmle: An R Package for Targeted Maximum Likelihood Estimation, *Journal of Statistical Software,* 2012, pp. 13:1-35.

[13] Van der Laan M. J. and Rubin D., Targeted maximum likelihood learning, in *U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper, 2006, 213.*

[14] Fadini G. P., Zatti G. , Consoli A. et al, Rationale and design of the DARWIN-T2D (DApagliflozin Real World evIdeNce in Type 2 Diabetes): A multicenter retrospective nationwide Italian study and crowdsourcing opportunity, *Nutrition, Metabolism and Cardiovascular Diseases, 2017,* pp. 27 (12) 1089-1097.

[15] Davies M. J.,. D'alessio D. A, Fradkin J. et al, Management of Hyperglycemia in Type 2 Diabetes, *Diabetes Care, 2018,* pp. 41:2669-2701.

[16] Holman R. R., Bethel M. A., Mentz R. J. et al. , Effects on once-weekly Exenatide on Cardiovascular Outcomes in Type 2 Diabetes, *N Engl J Med, 2017,* pp. 377:1228-1239.

[17] Gaede P., Vedel p., Larsen N. et al, Multifaactorial intervention and cardiovascular disease in patients with type 2 diabetes, *N Engl J Med,2003,* pp. 348:383-393.

[18] Fadini G. P., Sciannameo V., Franzetti I. et al., Similar effectiveness of dapagliflozin versus GLP-1 receptor agonists on combined endpoints in routine clinical practice. A multicenter retrospective study.,*Diabetes obesity Metab,* 2019.

[19] van Buuren S. and Groothuis-Oudshoorn K., mice: Multivariate Imputation by Chained Equations in R, *Journal of Statistical Software, 2011,* p. 45 (3).

[20] Moons K. G., Donders R. A. , Stijnen T. et al, Using the outcome for imputation of missing predictor values was preferred, *J Clin Epidemiol, 2006,* pp. 59: 1092-1101.

[21] Austin and Austin P. C., An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies, *Multivariate Behav Res, 2011,* pp. 46:399-424.

[22] Penning De Vries B. L. and GroenWold R. H., A comparison of approaches to implementing propensity score methods following multiple imputation, *Epidemiology biostatistics and public health,* 2017.

[23] Sofrygin O., Van der Laan M. and Neugebauer R., simcausal Package: Technical Details and Extended Examples of Simulations with Complex Longitudinal Data.

[24] Luque- Fernandez M. A., Schomaker M., Rachet B. and Schnitzer M. E., Targeted maximum likelihood estimation for a binary treatment: a tutorial, *Statistics in Medicine, 2018,* pp. 37 (16), 2530-2546.

OVERSAMPLING AND REPLACEMENT STRATEGIES IN PROPENSITY SCORE MATCHING: A CRITICAL REVIEW FOCUSED ON SMALL SAMPLE SIZE IN CLINICAL SETTINGS

Bottigliengo D.[1], Baldi I.[1], Lanera C.[1], Bejko J.[2], Bottio T.[3], Tarzia V.[3], Carrozzini M.[3], Gerosa G.[3], Berchialla P.[4], Gregori D.[1]

1 Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, and Vascular Sciences and Public Health, University of Padova, Padova, Italy

2 Department of Clinical and Experimental Sciences, University of Brescia, Italy

3 Department of Cardiac, Thoracic and Vascular Sciences and Public Health, University of Padova, Italy

4 Department of Clinical and Biological Sciences, University of Torino, Torino, Italy

Introduction

Inferences on the effects of treatments or exposures are increasingly found by using observational studies [1]. In such situations, the lack of randomization does not ensure the overall balance of individual baseline characteristics. Thus, statistical methods that are able to detect the treatment effect on an outcome of interest while controlling for potential confounders are needed. Propensity score (PS) methods are among the most used approaches in the medical literature for addressing the impacts of therapies or exposures. In particular, the use of propensity score matching (PSM) is widespread in clinical studies because of its ability to mimic a randomized clinical trial (RCT) in which the effect of a therapy is evaluated by comparing the outcomes of treated and control subjects belonging to the matched sample [1].
Among the many branches of medicine, PSM methods have become very popular in cardiothoracic surgery, as recent studies have demonstrated [2–5]. In these settings, there are two main issues that greatly impede the inference process: the treatment assignment is highly dependent on the subject characteristics and the small sample size. A low number of individuals is a common feature of cardiothoracic surgery studies as sample sizes often range from less than 100 subjects to a few hundred subjects [6–8]. Constructing a matched sample using standard PSM methods, such as 1:1 matching without replacement, could further reduce the initial sample size of the study, thus leading to an unfeasible comparison of different surgical procedures among matched subjects. Some studies in the medical literature addressed this issue using PSM with oversampling, i.e., matching more than one control to each treated variable or more than one treated variable to each control [9,10]. The main idea behind this approach is to create a matched set of individuals with a larger size than the one that would be obtained using classical 1:1 matching to increase the statistical power that is needed to detect the potential effect of interest. Furthermore, the use of matching with replacement may be useful for finding all the matched units from the control group, which is defined by the level of oversampling, e.g., 5 control units.

Objectives

This study evaluates whether performing PSM with replacement and oversampling can eventually ease the “curse” of a small sample size. We carried out the investigation using Monte Carlo simulations, and we applied the proposed approaches to a real case study. As a motivating example, we use the data from the study of Bejko et al. (2018) [11], in which the outcomes of different continuous-flow left ventricular assist devices are compared.

Methods

PSM is a PS-based method that allows for the creation of a matched set in which treated and control individuals share similar pretreatment characteristics. Because of the similar PS values, matched pairs will have similar baseline characteristics. Thus, each treated individual can be thought of as paired with him or herself since the matched controls share similar characteristics. Classical 1:1 matching without replacement may not be the most suitable approach when the number of individuals who are enrolled in a study is small and when an imbalance in terms of treated and untreated subjects is present. In such settings, such as in cardiothoracic surgery, it is almost impractical to detect the presence of a potential benefit that could be derived by using some new therapies or procedures using classical 1:1 matching. A PSM approach might have the potential to construct a matched sample that is larger than the one that would be formed using the classical PSM method in order to make the evaluation of the effects of therapies or treatments less challenging. The use of matching with replacement and with oversampling can be helpful in similar situations. By matching treated subjects with more than one control subject and using each control subject multiple times, one may be able to create a matched set with more observations than the initial set.
The performances of PSM with different combinations of replacement and oversampling were compared using Monte Carlo simulations. We considered the setting with a binary outcome and a binary treatment for two reasons. First, it is the most common setting in the medical literature and, second, it mimics the situation of our motivating example. PS was computed using a logistic model and matching was then performed using the nearest neighbor algorithm with a caliper and random matching order, which was shown to perform well in many situations [12,13]. The caliper was set to 0.2 of the standard deviation of the logit of the PS distribution, which was shown to result in good performances in several settings [14]. Within every simulated dataset, we ran a total of 10 matching procedures over all the combinations of replacement (with and without replacement) and levels of oversampling, and we varied the number of paired subjects from 1 to a maximum of 5. The balance of the individual baseline characteristics was evaluated using the standardized mean differences (SMDs) of the covariates between the compared groups and the variance ratios (VRs) of the baseline covariates of the compared groups. Moreover, we provided an overall measure of the balance of the matched set using the L^1 norm distance, which evaluates the differences between the multivariate empirical distributions of the baseline covariates of the treated and control groups [15]. For each matched set of subjects, we evaluated the impact of the treatment by computing the ATT (the only measure the PSM allows one to compute) as the risk difference of the outcome occurrence between the two compared groups. The performance of PSM with replacement and oversampling was evaluated using the following five criteria: the bias, the empirical standard error (ESE), the root mean squared error (RMSE), the median absolute error (MAE) and the nominal coverage (NC).
The methods were applied to a study that aimed to compare two different LVADs that were implanted at the cardio surgery “V. Gallucci” center at the University of Padova in Italy and that were used to treat patients with end stage heart failure that is nonresponsive to medical and conventional surgical therapy. In our study, we focused our attention on the occurrence of both early and late driveline infections DLI.

Results

We first examine the results of Monte Carlo simulations. Considering the treatment group, the lower the sample size and the proportion of subjects in the treatment group were, the higher the imbalance. The overall balance remained constant after 2 oversampled units from the opposite group after a small decrease of the considered measure. Matching without replacement seemed to result in small benefits in terms of the balance only when the sample size was higher than 500. Switching to the control group, small differences were observed between the matching with and without replacement for different treatment assignment mechanisms.

The treatment effect was systematically underestimated for the treatment group, whereas a systematic overestimation of the treatment effect was observed for the control group. Generally, the most biased estimates were observed when the sample size was low and when the treatment assignment was strong.
Regarding NC, when the focus was on the treatment group, the overall higher NC was achieved when using PSM with replacement, especially with higher proportions of treated and weak treatment assignments. The NC worsened when the level of oversampling increased. A different situation emerged for the control group. A better NC occurred when the treatment assignment was weak and replacement was used. Furthermore, we found that PSM with replacement resulted in an NC closer to 0.95 when the sample size was 1000 than for the PSM without replacement. Here, no relevant pattern was observed when the level of oversampling increased except for a sample size of 1000 and a weak treatment assignment, where we found that the increased level of oversampling produced a lower NC.
Regarding the results of the case study, the overall imbalance in the matched sets was high. No significant differences on the driveline infection occurrence between the HVAD and JARVIK 2000 arose. The estimates rarely were greater than 0.1 points of risk difference, and they were characterized by high variability with a 95% CI that always included 0.

Conclusions

In summary, the treatment effect estimation in small sample size settings remains an open issue. If a researcher aims to gain greater balance in the matched set, we suggest applying PSM with replacement and oversampling. In these situations, 2 or 3 oversampled subjects are enough since the overall balance for a higher level of oversampling remains stable. When PSM is conducted with replacement and oversampling, the individuals in the matched sets should always be weighted proportionally with respect to the times they appear in each matched pair. Nevertheless, we highly encourage applied researchers to be very careful about bias due to incomplete matching, which can easily arise in situations with small sample sizes and were the treatment assignment is strongly affected by the individual pretreatment characteristics. In such situations, no incremental gain in the overall balance can address the issue of biased estimates of the treatment effect.

Bibliography

[1] Austin P.C., An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivar. Behav. Res.* 2011, *46*, 399–424.

[2] Ensminger S., Fujita B., Bauer T., et al., Rapid Deployment Versus Conventional Bioprosthetic Valve Replacement for Aortic Stenosis. *J. Am. Coll. Cardiol.* 2018, *71*, 1417–1428.

[3] Fraser C.D., Zhou X., Palepu S., et al., Tricuspid Valve Detachment in Ventricular Septal Defect Closure Does Not Impact Valve Function. *Ann. Thorac. Surg.* 2018.

[4] Park D.-W., Seung K.B., Kim Y.-H., et al., Long-Term Safety and Efficacy of Stenting Versus Coronary Artery Bypass Grafting for Unprotected Left Main Coronary Artery Disease: 5-Year Results From the MAIN-COMPARE (Revascularization for Unprotected Left Main Coronary Artery Stenosis: Comparison of Percutaneous Coronary Angioplasty Versus Surgical Revascularization) Registry. *J. Am. Coll. Cardiol.* 2010, *56*, 117–124.

[5] Bejko J., Toto F., Gregori D., et al., Left ventricle assist devices and driveline's infection incidence: a single-centre experience. *J. Artif. Organs Off. J. Jpn. Soc. Artif. Organs* 2018, *21*, 52–60.

[6] Colli A., Carrozzini M., Francescato A., et al., Acute DeBakey Type I aortic dissection without intimal tear in the arch: is total arch replacement the right choice? *Interact. Cardiovasc. Thorac. Surg.* 2018, *26*, 84–90.

[7] Colli A., Carrozzini M., Galuppo M., et al., Analysis of early and long-term outcomes of acute type A aortic dissection according to the new international aortic arch surgery study group recommendations. *Heart Vessels* 2016, *31*, 1616–1624.

[8] Shi W.Y., Moreno-Betancur M., Nugent A.W., et al., Long-Term Outcomes of Childhood Left Ventricular Non-Compaction Cardiomyopathy: Results from a National Population-Based Study. *Circulation* 2018.

[9] Toumpoulis I.K., Anagnostopoulos C.E., Katritsis D.G., et al., The impact of preoperative thrombolysis on long-term survival after coronary artery bypass grafting. *Circulation* 2005, *112*, I351-357.

[10] Boening A., Friedrich C., Hedderich J., et al., Early and medium-term results after on-pump and off-pump coronary artery surgery: a propensity score analysis. *Ann. Thorac. Surg.* 2003, *76*, 2000–2006.

[11] Bejko J., Pittarello D., Falasco G., et al., A pilot study on the efficacy and safety of a minimally invasive surgical and anesthetic approach for ventricular assist device implantation. *Int. J. Artif. Organs* 2017, 0.

[12] Austin P.C., Statistical Criteria for Selecting the Optimal Number of Untreated Subjects Matched to Each Treated Subject When Using Many-to-One Matching on the Propensity Score. *Am. J. Epidemiol.* 2010, *172*, 1092–1097.

[13] Austin P.C., Stuart E.A., Estimating the effect of treatment on binary outcomes using full matching on the propensity score. *Stat. Methods Med. Res.* 2017, *26*, 2505–2525.

[14] Austin P.C., Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm. Stat.* 2011, *10*, 150–161.

[15] Iacus S.M., King G., Porro G., Multivariate Matching Methods That Are Monotonic Imbalance Bounding. *J. Am. Stat. Assoc.* 2011, *106*, 345–361.

BIVARIATE COPULA METHOD FOR AGE ESTIMATION IN FORCED MIGRANTS.

Faragalli Andrea[1], Bucci Andrea[1], Skrami Edlira[1], Cameriere Roberto[2], Gesuita Rosaria[1], Dajko Marianxhela[1], Carle Flavia[1], Ferrante Luigi[1].

[1] *Centro di Epidemiologia, Biostatistica e Informatica Medica, Università Politecnica delle Marche, Ancona.*

[2] *Istituto di Medicina Legale, Università di Macerata, Macerata.*

Introduction

Age estimation for living individuals and human remains is a common problem in forensic medicine and its importance is growing due to the increasing number of subjects without identity document, forced to flee their home country. In that field, several approaches, based on regression analysis and relying on dental [1-4] and skeletal age [5-6] as maturity measures, have been devised. The most frequent methods treat chronological age as dependent variable which is estimated utilizing dental and skeletal maturation as predictors. Nevertheless, these approaches conduct to bias estimates of chronological age [7]. In order to reduce this bias, Lucy and Pollard [8] suggested to use classical calibration or "inverse regression" technique, where the predictor is specified in function of age. However, classical calibration could conduct to large values of Confidence Interval's limits, that could be even infinite. To overcome this issue, Ferrante et al [2] introduced a Bayesian calibration method focused on dental maturation as unique predictor of chronological age. Within the project entitled "Age estimation in forced migrant populations: a proposal for accurate e friendly-to-use method", we developed, in the Bayesian context, a new method that considers two age's predictors (wrist maturation and dental maturation indices) by applying the Copula instrument.

Targets

In this study we aimed to:

- Introduce a new Bayesian age estimation method via Copula instrument;
- Analyse the dependence structure of its predictors;
- Assess the predictive capability of the proposed method;
- Compare our estimation results with a model with independent predictors.

Methods

Let us define S as the dental maturity index, calculated by sum of open apices, and W as the hand/wrist maturation index, calculated as observed mineralization of the carpal area [6]. We consider a sample of n subjects with known ages, $y = \{y_1, \dots, y_n\}$, observed values $S = \{s_1, \dots, s_n\}$ and $W = \{w_1, \dots, w_n\}$. Our purpose is defining the joint distribution of predictors and using it to estimate the unknown chronological age (u). In order to do so, the Copula theory can be used to link the distribution of S and W. Copula is a statistical tool that can isolate and capture the full structure of dependence contained in every joint distribution function. It is based on the assumption that every joint distribution function can be decomposed into marginal distribution functions and copula [10].

In bivariate case, for example, this implies that:

$$F_{S,W}(S,W) = C\big(F_S(S), F_w(W)\big)$$

or

$$f(S,W) = f(S) \cdot f(W) \cdot c\big(F_S(S), F_W(W)\big)$$

where $c\big(F_S(S), F_W(W)\big)$ represents the density function of copula, evaluated in $F_S(S), F_W(W)$.. In this work, we focused on Clayton, Gumbel and Gaussian copulas that catch respectively lower, upper and no tail dependence. Rotated versions $(90°, 180°, 270°)$ of Gumbel and Clayton copulas were also considered. We applied a Bayesian calibration combined with a non-linear model for the dental maturity index, S [2; 13], and a linear model for wrist maturity index, W. To define the probability model in Bayesian calibration, a researcher should consider the characteristics of data and the entailed predictors.

Let us define $p(s|y, \theta_s)$ and $p(w|y, \theta_w)$ as conditional density function of our variables to age, where θ_s and θ_w represent respectively the vectors of their parameters.

By using the copula framework, we define the joint distribution as follows:

$$p(s, w|y, \theta) = p(s|y, \theta_s) \cdot p(w|y, \theta_w) \cdot c(F_S(s), F_W(w)|y, \theta_c). \quad (1)$$

Assuming that observations are independent and identically distributed, the likelihood can be factorized as:

$$\prod_{i=1}^{n} p(s_i|y_i, \theta_s) \cdot p(w_i|y_i, \theta_w) \cdot c(F_S(s_i), F_W(w_i)|y_i, \theta_c). \quad (2)$$

Given age u and vector of parameters $\boldsymbol{\theta} = \{\theta_s, \theta_w, \theta_c\}$, whose joint prior distribution is $h(\boldsymbol{\theta})$, the new observations s_{obs} and w_{obs} are independent of the observed data and follow the same probability model. Based on these assumptions, the posterior distribution of $\boldsymbol{\theta}$ may be written as:

$$h(\theta|y, s, w) = \frac{h(\theta) \prod_{i=1}^{n} p(s_i|y_i, \theta) \cdot p(w_i|y_i, \theta) \cdot c\big(F_S(s_i), F_W(w_i)|y_i, \theta\big)}{\int_\Theta h(\theta) \cdot \prod_{i=1}^{n} p(s_i|y_i, \theta) \cdot p(w_i|y_i, \theta) \cdot c\big(F_S(s_i), F_W(w_i)|y_i, \theta\big) d\theta}. \quad (3)$$

Then the calibrating distribution is:

$$f(u|s_{obs}, w_{obs}, y, s, w) = \frac{p(u)\phi(s_{obs}, w_{obs}|u, y, s, w)}{\int_0^{+\infty} p(u)\phi(s_{obs}, w_{obs}|u, y, s, w) du}. \quad (4)$$

Where $\phi(s_{obs}, w_{obs}|u, y, s, w)$ is the joint predictive distribution:

$$\phi(s_{obs}, w_{obs}|u, y, s, w) = \int_\Theta p(s_{obs}, w_{obs}|u, \theta) h(\theta|y, s, w)\, d\theta \quad (5)$$

and $p(u)$ is the prior distribution of age.

The predictive distribution $\phi(s_{obs}, w_{obs}|u, y, s, w)$ cannot be analytically solved but it can be approximated to the sample mean:

$$\frac{1}{M} \sum_{m=1}^{M} p(s, w|u\theta_m) \quad (6)$$

where $\theta_m, m = 1, \dots, M = 500$ are posterior draws from $h(\theta|y, s, w)$.

We considered that the maturity indices both followed either a Normal or Asymmetric Laplace Distribution (ALD). The location parameters of distributions of S and W were modelled respectively as broken line [13;9] and simple linear functions of time.

The algorithm used is the following:

1. Create the posterior distribution for the vector of parameters, $\boldsymbol{\theta}$, sampling from equation (3) through Metropolis-Hastings (M-H) algorithm**;**
2. Create the calibrating distribution (4);
3. Generate a random sample $\mathbf{u} = (\mathbf{u_1}, \dots, \mathbf{u_m})$**,** with $m = 500$, from equation (3);
4. Calculate the joint predictive distribution (5) by using equation (6).

We analysed data regarding the orthopantomography and wrist x-ray of Italian children and adolescents aged between 5 and 15 years. We randomly subdivided the whole dataset into a training sample, with N_{tr} subjects, and a test sample, with N_{te} subjects. We used the training sample to develop the models. The most suitable version of Gumbel and Clayton copula was selected by minimizing the Bayesian Information Criteria (BIC).
We used the test sample to assess the performance of the proposed method. In this study the mode of calibrating distribution, $\hat{u}_i$, was used for point estimates.
Model performance is assessed by:

- Mean Absolute Error (MAE),
  $MAE = \frac{1}{N_{te}} \sum_{i=1}^{N_{te}} |e_i|$ where $e_i = y_i - \hat{u}_i$;

- Root Mean Squared Error (RMSE),
  $RMSE = \sqrt{\frac{1}{N_{te}} \sum_{i=1}^{N_{te}} e_i^2}$;

- the Interquartile Range of distribution of the error (IQR_{ERR}).
- the average size of the quantile-based 95% Bayesian Confidence Interval ($MCI_{95\%}$) of the calibrating distribution was used as precision measure.
-

Results

We analysed 235 orthopantomography and wrist x-ray from 114 males and 121 females in which $N_{tr} = 160$ observations were used as training sample and $N_{te} = 75$ as test sample. The predictive performance of the considered models is reported in Table 1. In Panel A and B, we compared models' prediction accuracy when the age predictors were distributed in two different ways: Normal and asymmetric Laplace. In our case, Rotated Gumbel (270°) (RG) and Rotated Clayton (90°) (RC) copulas were selected. In Panel A, the RC model achieved the best performance in terms of MAE, RMSE and $MCI_{95\%}$, but not for the IQR_{ERR} measure where it performed slightly worse than independent one. This is motivated by the fact that the residuals' distribution of RC model had heavier left tail than the independent model. Both RC and independent model did not show any significant bias in estimated residuals. The results showed in Panel B confirmed the good performance of copula instrument also with a different age predictors' distribution. In fact, in that case, all our new models outperformed the independent one in each of the measures considered. Comparing the outcomes among panels, we could notice that the assumption of Normal distribution for our two maturation indices led to more accurate prediction of chronological age than assuming them distributed as asymmetric Laplace. In addition, the results showed in Panel B were mostly comparable across copulas, while the assumption of independent age predictors did not provide accurate estimates of chronological age.

Table 1. *Performance of considered models of age estimation with observed data*

Model	MAE	RMSE	IQR_{ERR}	$MCI_{95\%}$
Panel A: S ~ N, W ~ N				
Independent	0.98	1.22	(0.015;1.170)	2.18
Gaussian Copula	1.94	2.43	(0.025;3.190)	1.52
R-Gumbel Copula (270°)	1.60	2.08	(-0.130;2.200)	2.03
R-Clayton Copula (90°)	0.86	1.07	(-0.475;0.970)	2.05
Panel B: S ~ ALD (τ = 0.45), W ~ ALD (τ = 0.5)				
Independent	1.16	1.45	(-0.935;0.955)	6.79
Gaussian Copula	1.05	1.30	(-0.905;0.985)	2.52
R-Gumbel Copula (270°)	1.07	1.31	(-0.930;0.930)	2.52
R-Clayton Copula (90°)	1.07	1.31	(-0.940;0.935)	2.56

Conclusion

Accurately estimating chronological age in forensic and legal medicine assumes a crucial role for solving a variety of legal question like criminal liability, majority status and the identification of both living and death individuals. Extending the approaches suggested by Ferrante et al. [2] and Bucci et al [9], we introduced a new method that links together two different maturity indices, S (dental maturation) and W (hand/wrist maturation), via copula instrument. These maturations were calculated by using radiology images that are easy to obtain from a non-invasive clinical exam, therefore suitable to be used to forced migrant population when age estimation is needed. Our study seeks to combine the proposed methodology and the use of Bayesian calibration. This allows the researcher to remove the bias [9] in the estimation results and to capture the dependence structure existing between the age predictors used. We considered that our two variables both followed either a Normal or asymmetric Laplace distribution. In both cases, the copulas' model outperformed the independent model in which the age predictors were assumed independent. In conclusion the copula's models appear to be more accurate and robust than the competing independent model since they exhibited the lowest IQR_{ERR} and $MCI_{95\%}$. Our method also allows to exceed the typical bias inherent in the regression model approach. Further investigation is needed, with particular regard on the choice of a proper age predictors distribution. In addition, the copula model could be extended using several predictors.

Bibliography

[1] Cameriere R., De Luca S., Alemán I. et al., Age Estimation by pulp/tooth ratio in lower premolars by orthopantomography, Forensic Science International, 2012 Jan; 214: 105-112.

[2] Ferrante L., Skrami E., Gesuita R. et al., Bayesian calibration for forensic age estimation, Statistics in Medicine, 2015 May; 34(10):1779-17790.

[3] Agelakopoulos N., De Luca S. Veli a Palacio L. A. et al., Third molar maturity (I3M) for assessing age of maturity: study of a black South African sample, Int J Legal Med. 2018 Sep;132(5):1457-1464.

[4] Cameriere R., Vel ia Palacio L. R., Pinares J.et al., Assessment of second (I2M) and third (I3M) molar indices for establishing 14 and 16 legal ages and validation of the Cameriere's I3M cut-off for 18 years old in Chilean population., Forensic Sci Int. 2018 Apr;285:205.

[5] Tanner J. M., Whitehouse R.H., Marshall W.A. et al., Assessment of skeletal maturity, prediction of adult height (TW2 method). Arch Dis Child. 1976 Jun; 51(6): 488.

[6] Cameriere R., Ferrante L., Age estimation in children by measurement of carpals and epiphyses of radius and ulna and open apices in teeth: a pilot study. Forensic Sci Int. 2008 Jan;174(1):60-63

[7] Aykroyd R. Lucy D. Pollard, A. et al. Technical note: regression analysis in adult age estimation. Am J Phys Anthropol. 1997 Oct;104(2):259-65.

[8] Lucy D., Pollard A. Further comments on the estimation of error associated with the Gustafson dental age estimation method. J Forensic Sci. 1995 Mar;40(2):222-7.

[9] Bucci A. Skrami E. Faragalli A. et al., Segmented Bayesian Calibration Approach for Estimating Age in Forensic Sciences, Biometrical Journal, submitted.

[10] Sklar A. Fonctions de repartition à n dimension et leurs marges. 1959, Inst. Statist. Univ. Paris, 8, pp. 229-231, 1959.

[11] Fisher R. A., Statistical Methods for Research Workers. 1932

[12] Rosenblatt M., Remarks on a Multivariate Transformation. 1952, Ann. Math. Statist. Volume 23(3), 470-472.

[13] Muggeo V. M. R., Estimating regression models with unknown break-points. Stat Med. 2003 Oct; 22(19):3055-71.

VARIABLE SELECTION FOR MULTIPLE CORRELATED EXPOSURES IN HEALTHCARE RESEARCH

Pesenti Nicola[1], Zambon Antonella[1], Massimo Scacchi[2], Colicino Elena[3], Quatto Piero[4]

[1] *Department of Statistics and Quantitative Methods, Division of Biostatistics, Epidemiology and Public Health, University of Milano-Bicocca, Milan, Italy*

[2] *Division of Endocrinology and Metabolic Diseases, IRCCS Istituto Auxologico Italiano, Piancavallo-Oggebbio, Italy.*

[3] *Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA*

[4] *Department of Economics, Management and Statistics, University of Milano-Bicocca, Milan, Italy*

Introduction

The study of the association of multiple co-occurring exposures on a specific health outcome, which is of critical concern in epidemiology, have had a boost with the rapid evolution and adoption of big data by industry and the advent of “omics", that have led to an increase in the amount of available information. Therefore novel variable selection methods, able to handle jointly multiple correlated exposures which may have non-linear and non-additive relationships with the health outcome, are needed. Standard dimensionality reduction methods have drawbacks. For example, methods such as principal components analysis and factor analysis focus on the transformation of the original variables leading to a drop in interpretability. Penalized regression methods such as lasso, ridge regression or elastic net [1][2][3] in the presence of a group of highly correlated predictors tend to select only one predictor from the group [4].
In the recent years, a number of innovative approaches for estimating the health effects of several predictors, that allow for non-linear and non-additive relationships while adjusting for potential confounding factors and including variable selection within the regression framework, have been proposed. Among these methods stand out two bayesian semiparametric models: the bayesian kernel machine regression (BKMR) [5], which models the health outcome through the use of a kernel function of the exposure variables, and the bayesian semiparametric regression (BSR) [6], which flexibly models the relationships between the predictors and the outcome employing natural splines.

Aims

The aim of this work is to study the selection ability of these innovative methods and comparing them with the LASSO, which, among the penalized regression methods, uses the strongest penalty, leading to the most parsimonious model.

Methods

The BKMR model for a continuous outcome has the form:

$$Y_i = h(z_{i1}, ..., z_{iM}) + \boldsymbol{x}_i'\boldsymbol{\beta} + \epsilon_i \qquad (1)$$

where Y_i is the outcome for individual i (i = 1,. . . ,n), z_{im} is the mth exposure variable, $\boldsymbol{x_i}$ contains a set of potential confounders, h denotes the unknown exposure-response function to be estimated, $\boldsymbol{\beta}$ represents the effect of the covariates and

$$\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

BKMR uses a kernel machine representation to model the h function, which may include a large number of exposures of interest and the relationship between these exposures and the health outcome can be

complex. The BSR model is similar of BKMR (1), but the exposure-response function *h* is modeled through the use of natural spline basis functions.
Lastly, the LASSO is a technique for improving the OLS estimates, which imposes an L1 penalty on the regression coefficients, shrinking or setting them to zero:

$$\hat{\boldsymbol{\beta}}_{lasso} = arg\min_{\beta} \left[\sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p-1} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right]$$

The *λ* parameter controls the amount of shrinkage that is applied to the estimates.

We conducted a simulation study to evaluate the performance of the three proposed methods. We simulated *K*=100 datasets consisting of *n*=100 observations and *M*=50 predictors. The predictors were generated from a multinormal distribution with a known covariance matrix *R,* where Z_1, Z_5 and Z_6 were set to be highly correlated ($r > 0.7$). The other Z_i independent and identically distributed, $i = 2,3,4,7,...,50$. The outcome was generated as a linear function of the first two predictors and a quadratic function of the third predictor, no confounders were considered:

$$Y_i = 2Z_1 + 0.8Z_2 + Z_3^2 + \epsilon_i$$

where $\epsilon_i \sim N(0,1)$

Therefore, the variables of interest in the simulation are Z_1, Z_2 and Z_3 because related to the outcome and Z_5 and Z_6 which are correlated with Z_1. The average posterior inclusion probabilities (PIPs), which provide measures of variable importance for each exposure, were computed for BKMR and BSR for each predictor and the Monte Carlo's mean PIPs and standard errors are presented.
The LASSO Monte Carlo's mean inclusion probability $\hat{p}$ was calculated as the number of times each variable had non-zero coefficient over the 100 simulations, and the standard error estimated by:

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{K}}$$

The LASSO model was computed using 10-fold cross validation to choose the regularization parameter *λ*.

A case study motivating the proposed research are the data from Istituto Auxologico Italiano (FUOBAUXO project), a cohort study of patients hospitalized for weight loss at Piancavallo Center (VB). The aim of this research is to study the association between anthropometric, clinical and laboratory variables on weight loss during hospitalizations of same length. We compared the performance of the three techniques on *n*=200 patients and *K*=40 continuous variables of interest, choosed with institute's clinicians support.

Results

Simulation study. Compared models' mean PIPs and standard errors for the five variables of interest are presented in Table 1. Both novel methods were able to correctly identify the quadratic relationship with the outcome (Z_3) with 99% and 98% probability respectively. The strong linear relationship (Z_1) was perfectly detected by BKMR (100%) while with a lower probability from the BSR (76%). Variables Z_5 and Z_6, which were highly correlated with Z_1, were correctly excluded by both methods, though BSR mean PIPs were between 50 and 60%. Finally, the weak linear relationship (Z_2) was properly identified by the two models with mean PIPs between 70 and 80%. The LASSO model instead, was only able to identify the two linear associations with *Y* with 100 and 71% probability of inclusion, and not the quadratic relationship.

Regarding the non-associated variables, BKMR selected the variables Z_{17}, Z_{12} and Z_{27} with probabilities 64, 60 and 53%, while BSR selected the variable Z_{12} with probability 54% and LASSO retained Z_{39}, Z_8 and Z_{42} with 70, 68 ad 52% probabilities. All other variables had mean inclusion probability less than 50%.

Table 1. *BKMR, BSR and LASSO models' mean inclusion probabilities (Mean% (SE%)) for the five predictors of interest.*

Predictor	BKMR	BSR	LASSO
Z1	100.0 (0.0)	75.6 (2.8)	100.0 (0.0)
Z2	78.2 (2.8)	71.6 (3.0)	71.0 (4.5)
Z3	98.8 (0.9)	97.9 (0.4)	35.0 (4.8)
Z5	17.9 (1.1)	52.7 (2.9)	48.0 (5.0)
Z6	18.4 (1.0)	59.0 (3.0)	11.0 (3.1)

Case study. Applying the proposed approaches to identify the relationship between the 40 anthropometric, clinical and laboratory predictors and weight difference between hospitalization and discharge, we found that the variables *bmi* and *free fat mass* have been selected by all three methods with probability ≥ 98% for BKMR and BSR, while with non-zero coefficients for the LASSO method. BKMR and BSR identified other three laboratory variables as important, with posterior probability between 66 and 98%. Finally, BSR detected other four variables with PIPs ≥ 75% not identified by BKMR. The LASSO did not find any other predictors as important in describing the relationship with weight loss.

Discussion

Based on simulations, considering mean PIPs greater then 70%, BKMR and BSR models perform well on detecting linear and non linear relationships with the outcome, even in presence of highly correlated predictors, whereas LASSO methods lack in the identification of non linear associations. Regarding the case study all three models identified two variables which were expected in explaining the outcome variations. Both BKMR and BSR were able to detect other important predictors, with BKMR resulted as more parsimonious. Future work concerns the application of these methods on real applications and simulations, including confounder's variables and categorical predictors, and the assessment of computational performances and models' robustness when changes the number of iterations, subjects and variables.

References

[1] Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Ser B 1996, 58:267.

[2] Zou H, Hastie T (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society.Series B (Statistical Methodology), 67(2), 301-320.

[3] Bradley Efron, Trevor Hastie, Computer Age Statistical Inference: Algorithms, Evidence, and Data Science, Cambridge University Press, New York, NY, 2016.

[4] Kyung, M Gill, J Ghosh, Mand Casella, G 2010. Penalized regression, standard errors, and Bayesian lassos. Bayesian Analysis 5: 369{ 412.

[5] Bobb JF, et al. Bayesian kernel machine regression for estimating the health e_ects of multi-pollutant mixtures. Biostatistics. 2014. doi:10.1093/biostatistics/kxu058.

[6] Antonelli, J., Mazumdar, M., Bellinger, D., Christiani, D. C., Wright, R.,and Coull, B. A. (2017). Estimating the health e_ects of environmental mixtures using bayesian semiparametric regression and sparsity in-ducingpriors.ArXiv preprint arXiv:1711.11239.

SOCIOECONOMIC POSITION AND GASTRIC CANCER RISK IN THE STOMACH CANCER POOLING (STOP) PROJECT

Rota Matteo[1], Alicandro Gianfranco[2,3], Pelucchi Claudio[2], Bonzi Rossella[2], Bertuccio Paola[2], Lunet Nuno[3,4], Boccia Stefania[5,6], Negri Eva[7], Ferraroni M[2], La Vecchia Carlo[2] on behalf of the StoP Project study group.

[1] *Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy.*

[2] *Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy.*

[3] *Italian National Institute of Statistics, Rome, Italy.*

[3] *Departamento de Ciências da Saúde Pública e Forenses e Educação Médica, Faculdade de Medicina da Universidade do Porto, Porto, Portugal.*

[4] *EPIUnit – Instituto de Saúde Pública, Universidade do Porto, Porto, Portugal.*

[5] *Sezione di Igiene, Istituto di Sanità Pubblica, Università Cattolica del Sacro Cuore, Roma, Italia.*

[6] *Department of Woman and Child Health and Public Health - Public Health Area, Fondazione Policlinico Universitario A.Gemelli IRCCS, Roma, Italia.*

[7] *Department of Biomedical and Clinical Sciences, University of Milan, Milan, Italy.*

Introduction

Low socioeconomic position (SEP) has been linked to noncommunicable diseases, including several cancers. SEP reflects the availability of cultural, material and social resources that translate into advantages in terms of decision making, social network, lifestyle habits and also access to health services. SEP can be measured by a series of indicators, including education, occupation and income. These indicators are correlated but each of them measures different aspects of the socioeconomic stratification [1]. Education captures the intellectual assets of individuals besides the socioeconomic conditions in childhood and adolescence and also represents the opportunity to access to higher level jobs. Occupation reflects the privileges related to social standing, material resources and job-related risk factors; income is related to material resources, better living conditions and healthy environment.
The Stomach Cancer Pooling (StoP) Project is an international consortium of case-control (CC), including nested CC, studies on gastric cancer (GC) started in 2012 [2]. This consortium joins together scientists from several areas of the world aimed to study the role of lifestyle, dietary habits and genetic determinants in GC aetiology, through an individual participant data meta-analysis (IPD-MA). The latest StoP dataset release (version 3) includes 33 studies for a total of 12,753 stomach cancer cases and 30,682 controls. Its uniquely large sample size and access to raw patient-level data allowing centralized data harmonization are among the major strengths of the StoP Project [2].
In 2017, the center of Milan proposed to the Scientific Committee of the Stop Project a study on SEP and GC risk.

Aims

The first phase of the study aimed at i) quantifying the association between low SEP, as measured by education and income, and GC risk in the StoP Project, ii) studying the relation according to cancer subsite, histological subtype and in strata of geographic area or macroeconomic measure of income inequality of the country where the study was conducted. In a second phase, we will evaluate the mediation effect of some

risk factors for GC that are more frequent in people with low SEP (i.e. *Helicobacter pylori* infection, tobacco smoking, heavy alcohol drinking, obesity and diet rich in salt and processed meat).

Methods

A two-stage approach was adopted. We firstly estimated study-specific odds ratios (ORs) and the corresponding 95% confidence intervals (CIs) using multivariable unconditional logistic regression models. Polytomous unconditional logistic regression models were fitted when analyzing the association by cancer subsite and histological type. We applied multiple imputations using full chained equations to avoid data loss due to sporadically missing values in study-specific covariates. Under the missing at random assumption, five imputed datasets were generated for each study, with missing values filled in with a set of plausible values drawn from the posterior predictive distribution of the missing data, conditional on the observed data. The imputation models were congenial with the analysis models and included the same set of covariates plus the case–control status. Study-specific regression coefficients and their standard errors were obtained through the Rubin's rule. In the second stage, summary (pooled) effect estimates were computed using a random-effect model.
Education was standardized across studies using the International Standard Classification of Education (ISCED 2011) of the UNESCO, an international reference classification that facilitates comparisons of education systems across countries. We defined three categories: (i) *low education level*, including early childhood and primary education (ISCED 0–1); (ii) *intermediate education level*, including secondary education (lower and upper) and postsecondary non-tertiary education (ISCED 2–4); (iii) *high education level*, including tertiary vocational education, often designed to provide participants with professional knowledge, skills and competencies and education leading to a university degree (ISCED 5–6).
Household income was harmonized starting from study questionnaires by grouping comparable levels into 4 categories, i.e. low, lower middle, upper middle and high.
To facilitate comparison with results from different studies, we estimated the relative index of inequality (RII) for both education and household income. The RII is a unique regression-based summary measure of social inequality that allows comparisons across countries with different distributions of the socioeconomic variables. It considers the size of the population in each socioeconomic level and their relative position in the socioeconomic scale. It represents the GC risk of subjects at the highest level of the socioeconomic hierarchy as compared to those in the lowest one.
The RII was defined as follows:

1. Within each study, for each of the k ordered levels (i = 1, …, k) of the SEP variable (i.e., education or household income), let c_i be the proportion of study subjects in class i or lower.
2. Then, for each class i = 1, …, k, let define $x_i = \frac{c_i + c_{i-1}}{2}$ as the mean rank, that is, the midpoint between the proportion of study subjects in class i (c_i) and those in the previous one (c_{i-1}).
3. The RII was then estimated by including the mean rank x_i as explanatory variable in the models used to derive the ORs instead of the original SEP variable.

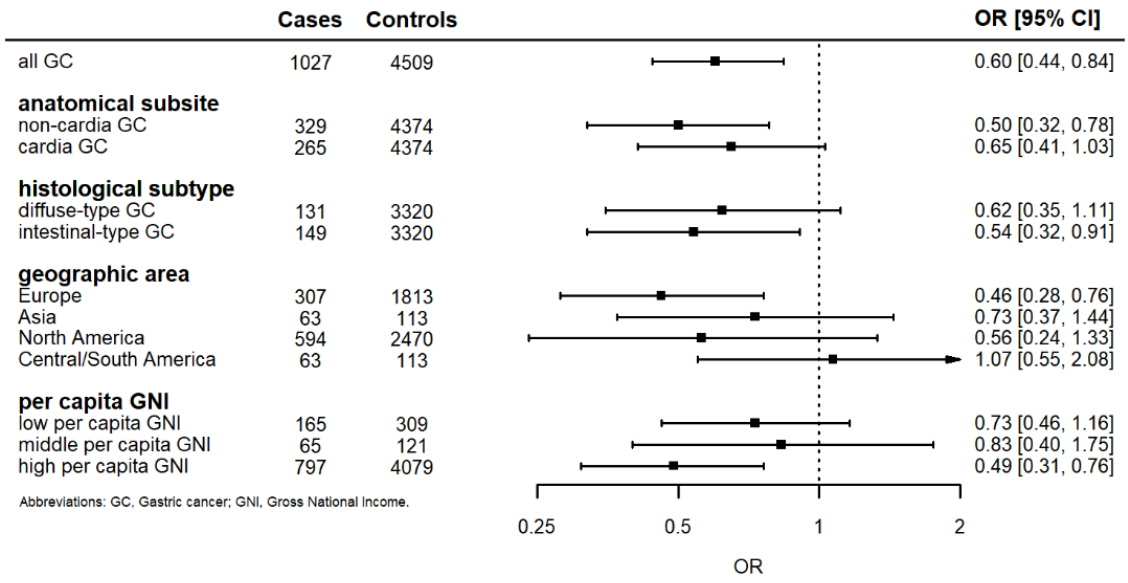
Models included terms for age, sex, alcohol drinking (never, ≤1 drink per day, >1 to ≤4 drinks/day and >4 drinks/day), tobacco smoking (never, former, current ≤10 cigarettes/day, >10 to ≤20 cigarettes/day, and >20 cigarettes/day), race/ethnicity (White, Hispanic/Latino, Black/African American, other), fruit and vegetable consumption (study-specific tertiles) and study centre (for multicentric studies).

Results

We used data from the second release of the StoP project dataset. A total of 25 case-control studies with available data for education or income (11 from European countries, 6 from Asian countries, 3 from North and 5 from Central/South American countries) for a total of 10,290 GC cases and 26,153 controls were considered [3]. Compared to low educational level (ISCED 0–1), both intermediate (ISCED 2–4) and high

(ISCED 5–6) educational levels were significantly associated with reduced GC risk, being the ORs from the fully adjusted models 0.68 (95% CI, 0.55–0.84) and 0.60 (95% CI, 0.44–0.84), respectively. The corresponding pooled RII was equal to 0.45 (95% CI, 0.29–0.69). Similar results emerged when using household income as a proxy for the SEP, with a significantly reduced GC risk in the highest as compared to the lowest household income category (OR 0.65, 95% CI, 0.48–0.89). The corresponding RII was 0.40 (95% CI, 0.22–0.72).
Results according to anatomical subsite, histological subtype, geographic area and study per capita GNI for highly (ISCED 5-6) as compared to less educated subjects (ISCED 0-1) are presented in Figure 1. In the analysis by cancer subsite, a strong inverse association was observed both for noncardia (highest vs. lowest level education: OR 0.50, 95% CI, 0.32–0.78) and cardia GC (OR 0.65, 95% CI, 0.41–1.03). Higher level of education was inversely associated with both diffuse (OR 0.62, 95% CI, 0.35–1.11) and intestinal-type (OR 0.54, 95% CI, 0.32–0.91) GC risk. The association was null when considering the studies from Central/South America (highest vs. lowest level education: OR 1.07, 95% CI, 0.55–2.08). There was a stronger significant inverse relationship between educational attainment and GC risk in studies from countries with high per capita GNI (OR 0.49, 95% CI, 0.31–0.76) as compared to those with middle (OR 0.83, 95% CI, 0.40–1.75) and low (OR 0.73, 95% CI, 0.46-1.16) per capita GNI.

Figure 1. *Selected results from stratified analyses for highly (ISCED 5-6) as compared to less educated subjects (ISCED 0-1).*



Conclusions

This uniquely large individual participant data meta-analysis provides a precise estimate of the strong inverse relationship between SEP and GC risk. We found a decreased GC risk among individuals with intermediate and high education levels as compared to those in the lowest level. Similar results emerged when we used household income as a proxy for the SEP. Our data call for public health interventions to reduce GC risk among the more vulnerable groups of the population. From the medical statistician perspective, the StoP project offers several opportunities of research related to methodological needs, including the management of systematically missing confounding factors, i.e. variables not collected in one

or more studies. In the next phase of the study we will estimate the direct and indirect effects of SEP on GC risk via other GC risk factors through causal mediation analysis. This will give further insights into the mechanisms underlying the socioeconomic disparities in GC risk.

References

[1] Geyer S, Hemstrom O, Peter R *et al*., Education, income, and occupational class cannot be used interchangeably in social epidemiology. J Epidemiol Community Health., 2006 Sep; 60(9):804–10.

[2] Pelucchi C, Lunet N, Boccia S *et al.*, The Stomach Cancer Pooling (StoP) project: study design and presentation. Eur J Cancer Prev., 2015 Jan; 24(1):16-23.

[3] Rota M, Alicandro G, Pelucchi C *et al.*, Education and gastric cancer risk – An individual participant data meta-analysis in the StoP project consortium. Int J Cancer. 2019 Mar 28 [Epub ahead of print]. DOI: 10.1002/ijc.32298.

ASSOCIATION BETWEEN DIETARY PATTERNS AND ARTERIAL STIFFNESS IN A CONTEMPORARY ADULT ITALIAN POPULATION: THE RISK OF CARDIOVASCULAR DISEASES AND ABDOMINAL AORTIC ANEURYSM IN VARESE (ROCAV) STUDY

Lasalvia Paolo[1], Gianfagna Francesco[1,2], Tarallo Antonino[3], Tozzi Matteo[3,4], Iacoviello Licia[1,5], Grandi Anna Maria[4], Ferrario Marco Mario[1], Zambon Antonella[6], Veronesi Giovanni[1]

1Centro Ricerche Epidemiologia e Medicina Preventiva, Dipartimento di Medicina e Chirurgia, Università degli Studi dell'Insubria, Varese;

2Mediterranea Cardiocentro, Napoli;

3Chirurgia Vascolare, ASST Sette Laghi, Varese;

4Dipartimento di Medicina e Chirurgia, Università degli Studi dell'Insubria, Varese;

5IRCCS Istituto Neurologico Mediterraneo Neuromed, Pozzilli; 6Dipartimento di Statistica e Metodi Quantitativi – Università di Milano-Bicocca, Milano

Introduction

A healthy diet is a well-established protective factor with respect to the onset of cardiovascular (CV) and other chronic diseases. The specific pathogenetic mechanisms for the onset of CV diseases are the modulation of lipid metabolism, inflammation and alterations of the micro-circulation and micro-vascular function, particularly in hypertensive patients. The pathway related to macro-circulation through the loss of elasticity of the large arterial vessels is less explored, few studies investigated specific foods rather than a complete dietary pattern [1, 2]. In these studies, the amount of low-fat dairy, fermented dairy and cheese seemed to have a beneficial effect on arterial stiffness [1, 2] as well as on subclinical atherosclerosis. [2]

Aims

The aim of this study was to investigate the cross-sectional association between dietary patterns, estimated both from Principal Components Analysis (PCA) and using a standard score of adherence to the Mediterranean diet, and arterial stiffness (diameter of the abdominal aorta in the infrarenal tract and the carotid-femoral Pulse Wave Velocity), in the context of a large population-based study recruited between 2013 and 2016 in the Varese city.

Methods

The ROCAV (Risk Of Cardiovascular diseases and abdominal aortic Aneurysm in Varese) study is a population based sample of men (50-75 years old) and women (60-75 years old) residents in the Varese city. Overall, the study summed-up 3777 subjects, with a satisfactory participation rate of 63.8% [3]. At enrolment, a vascular surgeon performed an abdominal aortic ultrasound scan to measure the diameter of the abdominal aorta in the infrarenal tract, which represents an indicator of a reduced arterial wall elasticity. In addition, a random half of participants underwent carotid-femoral Pulse Wave Velocity (cfPWV) assessment, a proxy of arterial stiffness. Lifestyle habits, including detailed cigarette smoking, were investigated through questionnaire, while anthropometric and clinical measures were ascertained by trained physicians. A blood sample was also collected. In addition, participants completed a computerized Food Frequency Questionnaire derived from the EPIC study on their habitual food consumption in the year prior to the survey [4]. First a detailed Data Quality Assessment was performed, including data completeness;

internal consistency through extremely high daily energy intake; seasonality assessment; external validity, through temporal (2003 EPIC cohort of residents in Varese) and spatial (with Moli-sani study, a 2005-2010 population-based cohort of the Molise population) comparisons. Then, from daily consumption data (in grams) for 41 food groups, food patterns were determined using PCA, by means of the scree-plot and the Kaiser method (eigenvalues>1) to retain the PCs. In addition, an adherence score to Mediterranean diet (MedS [5]) was derived according to gender-specific sample median daily intake of 9 food groups. The original score was then categorized as poor (0-3), medium (4-5) and optimal (6-9) adherence. The association between the retained PCs and the MedS with infrarenal diameter was estimated through nested linear regression models, adjusting for age, sex and average daily intake (Model 1), and further adding cigarette smoking, body mass index and diastolic blood pressure (Model 2). The same approach was adopted for cfPWV as the outcome, on the 1284 subjects (795 men and 489 women) with measured cfPWV. SAS 9.4 release was used for the analyses.

Results

From the data quality assessment, 238 (5%) subjects were excluded due to missing or unreliable dietary data. No significant seasonality trend was observed in daily energy intake (ANOVA-test p=0.15). The percentage of daily energy intake from macronutrients was consistent with the 2003 EPIC study from the same population. The spatial comparison confirmed a well-known North-South geographic gradient in average daily energy intake, with an average difference of about 200 kcal/day in both men and women, as well as in specific nutrients consumption, including daily alcohol intake. The percentage of energy intake from lipids was less than 50% in both men and women, suggesting a good adherence of this northern Italian population to the Mediterranean Diet. For the current analysis, 883 subjects with chronic degenerative diseases were further excluded, to attenuate the potential reverse bias, leaving a final sample of 2656 (1617 men, 1039 women). According to the mentioned criteria, four PCs were retained, corresponding to recognizable food patterns: high protein and animal fats (PC1); mediterranean-like diet (PC2); high carbs diet (PC3); and a poor diet characterized by low carbs and red meat, and high white meat, milk and coffee (PC4). All together, these explained 24% of the overall variance, in line with similar analyses [6]. The dietary patterns had expected association with the socio-demographic, the metabolic, the inflammatory, and the anthropometric subjects' profile (data not reported). In Model1, for 1 point increase in PC1 and PC4 the infrarenal diameter increased by 0.14 mm (95%CI: -0.02 to 0.31; p=0.08) and by 0.11 mm (95%CI: 0.00 to 0.22; p=0.06), respectively. In Model2, these coefficients were reduced to 0.10 and 0.09, respectively, not statistically significant. No association was observed for the remaining PCs, nor for the MedS. PC1 was associated with an increase in the cfPWV of 0.33 m/s (95%CI: 0.12 to 0.54; p=0.002), while PC2 led to a decrease in cfPWV by -0.16 m/s (95%CI: -0.33 to 0.01; p=0.07). The coefficient for PC1 remained substantially un-modified in Model 2 (0.34; 95%CI: 0.13 to 0.54; p=0.002); while PC2 coefficient was partially attenuated (-0.12; 95%CI: -0.29 to 0.05; p=0.16). No association was observed for PC3 and PC4, as well as for the MedS.

Conclusions

In our contemporary sample of Northern Italian adults free of chronic disease, a dietary pattern characterized by high intake of protein and animal fats was associated with increased cfPWV, a proxy of arterial stiffness. The association remained after adjustment for major confounders. Our data also suggest a potential protective role for a Mediterranean diet-like pattern, although its beneficial effects may be accounted for by other lifestyles and anthropometric risk factors. Finally, our data suggest no association between dietary habits and infrarenal diameter, potentially excluding a strong role for diet on the loss of large vessel functionality.

References

[1] Ribeiro A.G., Mill J.G., Cade N.V. et al., Associations of Dairy Intake with Arterial Stiffness in Brazilian Adults: The Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). Nutr., 2018; 10(6):701

[2] Recio-Rodriguez J.I., Gomez-Marcos M.A., Patino-Alonso M.C., et al., Association between fat amount of dairy products with pulse wave velocity and carotid intima-media thickness in adults. Nutr J., 2014; 13:37

[3] Gianfagna F., Veronesi G., Bertù L. et al., Prevalence of Abdominal Aortic Aneurysms and its relation with cardiovascular risk stratification: protocol of the Risk of Cardiovascular diseases and abdominal aortic Aneurysm in Varese (RoCAV) population based study. BMC Cardiovasc Disord., 2016; 16:243

[4] Pala V., Sieri S., Palli D. et al., Diet in the Italian EPIC cohorts: presentation of data and methodological issues. Tumori, 2003; 89:594-607

[5] Bonaccio M., Di Castelnuovo A., De Curtis A. et al., Moli-sani Project Investigators. Adherence to the Mediterranean diet is associated with lower platelet and leukocyte counts: results from the Moli-sani study. Blood, 2014; 123:3037-44

[6] Smith A., Emmett P., Newby P. et al., Dietary patterns obtained through principal components analysis: The effect of input variable quantification. Br J Nutr., 2013; 109(10):1881-1891

DIETARY FLAVONOIDS AND LUNG FUNCTION IN COPD PATIENTS

Mattioli Veronica[1], Garcia-Aymerich Judith[2], Zanolin Maria Elisabetta[1], Benet Marta[2], Cazzoletti Lucia[1], Garcia-Larsen Vanessa[3]

[1] *Unit of Epidemiology and Medical Statistics, Department of Diagnostics and Public Health, University of Verona, Verona, Italy*

[2] *ISGlobal, Barcelona, Spain*

[3] *Program in Human Nutrition, Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA*

Introduction

COPD (Chronic Obstructive Pulmonary Disease) is one of the leading cause of death in the world. The main risk factors are tobacco smoke, air pollutants and occupational dust, vapors and fumes. [1] The disease is characterized by airflow limitation mainly caused by restriction of small airways, obstruction, augmented mucus secretion and emphysema. [2] The pathogenesis of COPD begins with activated macrophages that start the inflammatory process, which in turn leads to an increase of oxidative stress and injury to the lung tissue. [3] An increase in the level of inflammation and oxidative stress determine a worsening of COPD symptoms and physiological deterioration, defined as exacerbation, that often require hospitalization. [4] Microorganisms, cigarette smoke and exposure to noxious gases trigger exacerbations as they increase the inflammation in the lungs. [5]
Flavonoids are a class of substances that belongs to polyphenols and they are synthesized by plants. They are mainly contained in fruits, vegetables and other beverages and plant-derived foods, such as dark chocolate, tea and wine. A large body of evidence showed that flavonoids have antioxidant, anti-inflammatory and anti-allergic properties. [6, 7] Experimental studies showed that flavonoids could have a beneficial effect on COPD thanks to their ability to counteract inflammation and oxidative stress. [8]
Epidemiological studies suggested a possible beneficial role of flavonoids on lung function, symptoms in COPD patients or prevalence of COPD, but the evidence is limited to a few subclasses of flavonoids. [9, 10]

Objective

The main aim of the present study is to assess the association between dietary flavonoids intake (total flavonoids and the main subcategories: flavanones, anthocyanins, flavan-3-ols, flavonols, flavones, proanthocyanidins) and post bronco-dilation Forced Expiratory Volume in One Second (FEV1) % predicted, Forced Vital Capacity (FVC) % predicted and FEV1/FVC ratio in Chronic Obstructive Pulmonary Disease (COPD) patients. A secondary aim is to study the association between flavonoids intake and the risk of hospital readmission for COPD exacerbation.

Methods

The analysis was performed in the frame of PAC-COPD (Phenotype and Course of COPD) study, which involves 9 university hospitals in Spain. PAC-COPD is prospective study designed with the aim of characterize the different phenotypes of COPD, its natural course and to identify the potential risk factors of exacerbations. [11] Subjects at their first hospital admission for COPD were recruited between January 2004 and March 2006 and the follow-up period was at least 5 years when possible. The diagnosis of COPD was confirmed if post-bronchodilator FEV1/FVC ratio was ≤0.70, in stable condition at least after 3 months from recruitment. For all participating hospitals, the protocol was approved by the respective Ethics Committees and all patients provided the written informed consent.

Among all the subjects participating in the PAC-COPD study, the 80% (n=274) completed a validated 122-items Food Frequency Questionnaire (FFQ) on the dietary habits of the last 2 years. The FFQ was administered telephonically from 9 to 12 months after recruitment. [12] The daily amount of each food (gr/day) was estimated from the answers provided by each participant. Nutrient intake was assessed using the Centre for Superior Studies in Nutrition and Dietetics (CESNID) Spanish food composition table.
Daily dietary flavonoids intake (total flavonoids and the subclasses flavanones, anthocyanins, flavan-3-ols, flavonols, flavones, and proanthocyanidins) were estimated combining information provided by the patients through the FFQ and composition tables from the US Department of Agriculture (USDA). A list of 54 vegetable foods was considered for the flavonoids intake estimation.
For each subject pulmonary parameters (FEV1, FVC and FEV1/FVC) were measured according to the procedures of the Spanish Society of Pulmonology and Thoracic Surgery (SEPAR); [13, 14] the bronchodilator test was performed through the administration of 400 µg of salbutamol, according to the GOLD recommendations. [15] The predicted values for FEV1 and FVC were estimated using the references values for forced spirometry. [16] A national administrative database, the Minimum Basic Dataset (CMBD), was used to obtain information on hospital readmission. Standardized questionnaires were used to collect information on sociodemographic characteristics, lifestyle, respiratory symptoms and drug treatment. For each site, a pulmonologist estimated the Charlson index of comorbidity by reviewing medical history of each participant. [17]
Unadjusted and adjusted linear regression models were used to assess the association between each flavonoids variable and lung function parameter (post-bronchodilator FEV1% predicted, FVC% predicted and FEV1/FVC), one at a time. Total flavonoids and the main sub-categories (flavanones, anthocyanins, flavan-3-ols, flavonols, flavones, and proanthocyanidins) were analyzed as continuous exposure for a variation of 1 Standard Deviation (SD). Beta coefficients (β-coefficients) are reported together with the 95% Confidence Interval (95% CI). The potential confounders considered were sex, age, Body Mass Index (BMI), socio-economic status, physical activity (kcal/week, assessed using Yale Questionnaire in the validated version for Spanish adults), total energy intake. [18, 19]
The association between dietary flavonoids intake and risk of hospital readmissions was assessed using negative binomial regression, adjusting for sex, age, body mass index, Charlson Index of Comorbidity (dichotomous variable: 1 comorbidity or at least 2 comorbidities), FEV1, socio-economic status (skilled or unskilled manual workers are classified as subjects with low socioeconomic status), total energy intake.

Results

The 93% of subjects were male and mean age was 68 years (SD 8). Mean post-dilation FEV1% predicted was 53 (SD 16), mean FVC% predicted was 72 (SD 16) and mean FEV1/FVC % was 54 (SD 12). According to the GOLD criteria, 119 subjects (43.4%) had from severe to very severe COPD (stages III and IV). The median intake of total flavonoids was 252.7 mg (Interquartile Range (IQR) 156.7-390.7). Patients were followed for a median period of 3.4 years (IQR: 2.7-3.9).
The results of the unadjusted and adjusted regression analyses are reported in Table 1. In the adjusted analysis, an increase of 1SD of anthocyanins intake was associated with an increase of 2.19 in FEV1% predicted (95% Confidence Interval (95% CI): 0.29;4.10); anthocyanidins intake was also associated with an increase of 1.60 (95% CI: 0.22;2.98) in FEV1/FVC% in the unadjusted analysis and an increase of 1.52 (95% CI: 0.20;2.84) in FEV1/FVC% in the adjusted analysis. Flavones intake was negatively associated with FEV1/FVC% (β coefficient -1.43, 95% CI: -2.81;-0.05) in the unadjusted model, but the association became not significant after adjustment.
In the unadjusted model, dietary intake of flavones was associated with a reduced risk of readmission for COPD (Incidence Rate Ratio (IRR) 0.80, 95% CI: 0.64;0.99) and after the adjustment the association remained statistically significant (IRR 0.72, 95% CI 0.54;0.97).

Conclusion

Experimental studies already suggested a potential antioxidant and anti-inflammatory effect of flavonoids in lung inflammation diseases. [8] The results of our analyses suggest that flavonoids contained in fruits and vegetables might help to counteract the negative effects of the inflammation in lungs. Dietary anthocyanins intake, in particular, was associated with better lung function in COPD patients and flavones intake was associated with a reduced risk of hospitalization for COPD related reasons.

Table 1. *Unadjusted and adjusted* association between flavonoids intake (increase of 1 Standard Deviation) and post-bronco dilation respiratory function (FEV_1% predicted, FVC% predicted, FEV_1/FVC%). Beta coefficients (β) are reported together with the 95% Confidence Interval (95% CI). Significant results are reported in bold.*

	FEV_1% pred.		**FVC% pred.**		**FEV1/FVC**	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Total flavonoids	0.78	1.30	-0.67	-0.13	0.01	0.01
	(1.09;2.66)	(-0.69;3.28)	(-2.59;1.26)	(-2.26;2.00)	(-0.00;0.02)	(0.00;0.03)
Flavanones	0.29	0.32	0.76	0.71	-0.00	-0.00
	(-.58;2.17)	(-1.59;2.22)	(-1.16;2.68)	(-1.33;2.74)	(-0.02;0.01)	(-.02;0.01)
Anthocyanins	1.85	**2.19**	0.18	0.80	**0.02**	**0.02**
	(-.01;3.71)	**(0.29;4.10)**	(-1.74;2.11)	(-1.25;2.86)	**(0.00;0.03)**	**(0.00;0.03)**
Flavan-3-ols	0.63	0.84	-0.62	-0.80	0.01	0.01
	(-.24;2.51)	(-1.10;2.78)	(-2.54;1.31)	(-2.88;1.27)	(-0.00;0.02)	(0.00;0.03)
Flavonols	-1.34	-1.41	-1.50	-1.75	-0.01	-0.01
	(-3.21;0.53)	(-3.45;0.63)	(-3.41;0.42)	(-3.93;0.42)	(-0.02;0.01)	(0.02;0.01)
Flavones	-1.87	-1.19	-0.68	-0.82	**-0.01**	-0.01
	(-3.73; 0.01)	(-3.15;0.77)	(-2.60;1.25)	(-2.93;1.28)	**(-0.02;-0.00)**	(0.02;0.01)
Proanthocyanidins	-0.30	0.10	-1.30	-0.80	0.00	0.01
	(-.18;1.57)	(-1.85;2.05)	(-3.22;0.62)	(-2.88;1.28)	(-0.01;0.02)	(-.01;0.02)

References

[1] Mannino D. M.,Buist A. S. Global burden of COPD: risk factors, prevalence, and future trends. 2007; 370(765-73)

[2] Barnes P. J., Shapiro S. D.,Pauwels R. A. Chronic obstructive pulmonary disease: molecular and cellular mechanisms. 2003; 22(672-88)

[3] de Boer W. I., Sont J. K., van Schadewijk A., et al. Monocyte chemoattractant protein 1, interleukin 8, and chronic airways inflammation in COPD. 2000; 190(619-26)

[4] Perera W. R., Hurst J. R., Wilkinson T. M., et al. Inflammatory changes, recovery and recurrence at COPD exacerbation. 2007; 29(527-34)

[5] Viniol C.,Vogelmeier C. F. Exacerbations of COPD. 2018; 27

[6] Rice-Evans Catherine, Miller Nicholas,Paganga George Antioxidant properties of phenolic compounds. 1997; 2(152-159)

[7] Ferrandiz M. L.,Alcaraz M. J. Anti-inflammatory activity and inhibition of arachidonic acid metabolism by flavonoids. 1991; 32(283-8)

[8] Kim H. P., Lim H.,Kwon Y. S. Therapeutic Potential of Medicinal Plants and Their Constituents on Lung Inflammatory Disorders. 2017; 25(91-104)

[9] Arts I. C.,Hollman P. C. Polyphenols and disease risk in epidemiologic studies. 2005; 81(317S-325S)

[10] Zhai T., Li S., Hu W., et al. Potential Micronutrients and Phytochemicals against the Pathogenesis of Chronic Obstructive Pulmonary Disease and Lung Cancer. 2018; 10

[11] Garcia-Aymerich J., Gomez F. P., Anto J. M., et al. [Phenotypic characterization and course of chronic obstructive pulmonary disease in the PAC-COPD Study: design and methods]. 2009; 45(4-11)

[12] Martin-Moreno J. M., Boyle P., Gorgojo L., et al. Development and validation of a food frequency questionnaire in Spain. 1993; 22(512-9)

[13] Manual SEPAR de Procedimientos. Módulo 3. Procedimientos de evaluación de la función pulmonar. 2002

[14] Manual SEPAR de Procedimientos. Módulo 4. Procedimientos de evaluación de la función pulmonar-II. 2004;

[15] Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD). 200

[16] Roca J, Burgos F, Sunyer J, et al. References values for forced spirometry. Group of the European Community Respiratory Health Survey. 1998; 11(1354-1362)

[17] Charlson M., Szatrowski T. P., Peterson J., et al. Validation of a combined comorbidity index. 1994; 47(1245-51)

[18] Dipietro L., Caspersen C. J., Ostfeld A. M., et al. A survey for assessing physical activity among older adults. 1993; 25(628-42)

[19] De Abajo S., Larriba R.,Marquez S. Validity and reliability of the Yale Physical Activity Survey in Spanish elderly. 2001; 41(479-85)

THE RELATIONSHIP BETWEEN CANCER AND DEMENTIA: ASSOCIATION OR CAUSATION? RESULTS FROM SWEDISH POPULATION BASED COHORT STUDIES

Damiano Cecilia[1], Bellocco Rino[1,2], Rizzuto Debora[3,4], Grande Giulia[3]

1 Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milano, Italy

2 Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

3 Aging research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, Stockholm, Sweden

4 Stockholm Gerontology Research Center, Stockholm, Sweden

Introduction

Older people are often affected by several comorbid conditions and by an increasing risk of death that arises with aging. Previous studies examining the association of cancer with dementia in older adults have usually used standard approaches without taking into account the competing risk of mortality. However, ignoring mortality may not provide valid estimates of risk of dementia, because cancer is strongly associated with the competing risk of death.

Objectives

Aim is to study the association between cancer and the onset of dementia in the older population. The competing risk methodology is used, to illustrate the appropriate statistical methods for competing risks, their correct application and interpretation of the results, having death as the competing event.

Methods

The study considers people over-72 years old from two Swedish population-based longitudinal studies: the Kungsholmen Project (KP) and the Swedish National Study on Aging and Care project, conducted in the Kungsholmen district of the city of Stockholm (SNAC-K project). Participants have been considered cancer patients if they had a cancer diagnosis prior to the start of the study, or a diagnosis reported by physicians at clinical examinations during the follow-up period. The presence of dementia has been investigated through clinical and neuropsychological assessments conducted by doctors and psychologists, using a three-step diagnostic procedure. The variables of interest include socio-demographic factors (age, sex, level of education) and lifestyle factors (body mass index, smoking and alcohol habit, physical activity, mental activity and social activity). Only participants with complete data for all the variables of interest have been considered, for a total of 910 (57.4 %) in KP and 1,363 (74.9%) in SNAC-K. Analyses have been conducted for each single cohort and for the combined cohort. For this reason, cohorts have been harmonized for the variables of interest. Three models have been run: Model 1 treats cancer as a time-independent variable; Model 2 treats cancer as a time-dependent variable; Model 3 provides the Fine and Gray approach to control for the competing risk of mortality. A further analysis has investigated the effect of the covariates when death is the outcome of interest, in the perspective of the time-dependent model.

Results

With dementia as the outcome of interest, using Model 1 cancer seems to be protective of dementia in KP cohort and combined cohort (HR 0.63; CI 0.47-0.85). With Model 2, there is no significantly protective effect for any of the cohorts (combined cohort HR 0.85; CI 0.63-1.14). Model 3 shows a subdistribution hazard

ratio non-significant for all three cohorts (combined cohort SHR 0.78; CI 0.58-1.05). When death is the outcome of interest and cancer diagnosis is treated as a time-dependent covariate, people with cancer have an increased hazard of death (combined cohort HR 1.30; CI 1.13-1.50).

Discussion

The study shows that when using models with the right specifications that control for mortality, cancer is not protective for dementia in the older population. Considering dementia as the outcome of interest, treating the exposure as a time-independent covariate leads to observe a biased inverse association between the two age-related diseases. When cancer is treated as a time-dependent covariate, more reliable estimates are obtained and the inverse association between cancer and dementia is not supported. When accounting for death as a competing risk, mortality selection is also an element that should be considered. The possible inverse relationship that can be observed is a result attributable to mortality selection, or to the increase in mortality in cancer patients, rather than to biological reasons. When death is the outcome of interest, results confirm what is previously known form the literature. The roles of the main risk factors are confirmed, in particular the presence of cancer causes a clear increase in the risk of death and the risk increases as well with increasing age. Any analysis of the association between pathological medical conditions in the elderly should correctly define the models, contemplate statistical considerations of competing risks and report results for both cause-specific and subdistribution hazard models. To better understand the association of interest between cancer and dementia, additional studies conducted in different populations are needed and also association between various types of cancer and dementia should be investigated. Data should come from high quality population databases and should be prospective.

Bibliography

[1] Koller M.T., Raatz, H., Steyerberg, E.W. et al., Competing risks and the clinical community: irrelevance or ignorance? Statistics in Medicine, 2012; 31:1089–1097.

[2] Hanson H.A., Smith K.R., Rasmussen K.M. et al., Is Cancer Protective for Subsequent Alzheimers Disease Risk? Evidence From the Utah Population Database. The Journals of Gerontology: Series B, 2016; 72(6):1032–1043.

[3] Putter H., Fiocco M., Geskus R.B. et al., Tutorial in biostatistics: competing risks and multistate models. Statistics in Medicine, 2007; 26(11):2389–2430.

[4] Fratiglioni L., Grut M., Forsell Y. et al., Prevalence of Alzheimer's disease and other dementias in an elderly urban population. Neurology, 1991; 41.

[5] Rizzuto D., Orsini N., Qiu C., Lifestyle, social factors, and survival after age 75: population based study. BMJ : British Medical Journal, 2012; 345.

[6] Calderón-Larrañaga A., Santoni G., Wang H. X. et al., Rapidly developing multimorbidity and disability in older adults: does social background matter? J Intern Med, 2018; 283:489– 499.

[7] Latouche L., Allignol A., Beyersmann J. et al., A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. Journal of Clinical Epidemiology, 2013; 6:648–653.

OVERWEIGHT IN CHILDHOOD, PUBERTY OR EARLY ADULTHOOD: CHANGING THE ASTHMA RISK IN FUTURE GENERATIONS?

Lucia Calciano[1*], Ane Johannessen[2,3*], Marianne Lønnebotn[2,3*],Bryndis Benediktsdóttir [4,5] , Randi Jacobsen Bertelsen[6], Lennart Bråbäck[7], Shyamali Dharmage[8], Karl A. Franklin[9], Thorarinn Gislason[5], Mathias Holm[10], Christer Janson[11], Deborah Jarvis[12], Rain Jõgi[13], Jeong-Lim Kim[14], Jorunn Kirkeleit[1], Caroline Lodge[15], Andrei Malinovschi[16], Jesus Martinez-Moratalla[17], Roy Miodini Nilsen[18], Antonio Pereira-Vega[19], Francisco Gomez Real[6,20],Vivi Schlünssen[21], Cecilie Svanes[22,3**], Simone Accordini[1 **].

*Equal contribution as first authors.

**Equal contribution as last authors.

1 Unit of Epidemiology and Medical Statistics, Department of Diagnostics and Public Health, University of Verona, Verona, Italy

2 Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

3 Department of Occupational Medicine, Haukeland University Hospital, Bergen, Norway

4 Faculty of Medicine, University of Iceland, Reykjavik, Iceland

5 Department of Sleep, Landspitali University Hospital, Reykjavik, Iceland

6 Department of Clinical Science, Haukeland University Hospital, Bergen, Norway

7 Occupational and Environmental Medicine, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden.

8Allergy and Lung Health Unit, School of population and Global Health, University of Melbourne, Melbourne, Australia

9Department of Surgical and Perioperative Sciences, Surgery, Umeå University, Umeå, Sweden

10 Department of Occupational and Environmental Medicine, Sahlgrenska University Hospital, Gothenburg, Sweden

11 Department of Medical Sciences, Respiratory, Allergy & Sleep Research, Uppsala University, Uppsala, Sweden

12 Faculty of Medicine, National Heart and Lung Institute, Imperial College, London, United Kingdom

13 Lung Clinic, Tartu University Hospital, Tartu, Estonia

14 Occupational and Environmental Medicine, Department of Public Health and Community Medicine, The Sahlgrenska Academy, University of Gothenburg, Sweden

15 Department of Population Health, University of Melbourne, Melbourne, Australia

16 Department of Medical Sciences, Uppsala University, Uppsala Sweden

17 Servicio de Neurología del Complejo Hospitalario Universitario de Albacete (CHUA), Servicio de Salud de Castilla-La Mancha (SESCAM), Albacete, Spain

18 Department of nursing, Bergen University College, Bergen, Norway

19 Pneumology Section, Hospital Juan Ramón Jiménez, Huelva, Spain

20 Department of Obstetrics and Gynecolog, Haukeland University Hospital, Bergen, Norway

21 Department of Public Health, Aarhus University, Aarhus, Denmark

22 Centre for International Health, Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

Background

Overweight is a risk factor for asthma [1], but it is still debated whether parental overweight may also increase asthma risk in the next generation. In fact, mothers' health behavior shortly before and during pregnancy affects children's health, but the research on the multi-generational effects of risk factors for asthma is limited and has been mainly focused on cigarette smoke [2].
The aim of our study is to assess the effect of parental overweight on offspring's asthma, taking into account different susceptibility time windows throughout parents' preconception lifespan, as well as to evaluate the potential mediating role of offspring's own overweight.

Methods

We evaluated 6,347 adult offspring (aged 18-52 years) of 2,044 fathers (37-66 years) and 2,549 mothers (39-65 years) from the multicentre RHINESSA generation study (Respiratory Health in Northern Europe, Spain and Australia, www.rhinessa.net), which examines offspring of initial participants in the ECRHS (European Community Respiratory Health Survey) [3]. Associations of parental overweight at age 8, puberty and age 30 (exposures of interest), with offspring's overweight at age 8 (potential mediator) and offspring's asthma with/without nasal allergies (outcome), was analyzed using two-level logistic regression and two-level multinomial logistic regression models, respectively. A counterfactual-based mediation analysis was performed to establish whether observed associations were causal direct or indirect effects (mediated through offspring's own overweight). In the mediation analysis, the estimate of the natural effects was obtained by the latent response variable mediator approach with probit link, theta parameterization and weighted least squares means and variance adjusted (WLSMV) estimators [4]. Non-bias-corrected bootstrap confidence intervals (10,000 resamples) were obtained.

Results

In the paternal line, fathers' overweight at age 8 was associated with offspring's childhood at age 8 (odds ratio OR 2.23; 95%CI 1.45-3.42), but not with the outcome. Fathers' overweight in puberty was associated with offspring's asthma without nasal allergies (relative risk ratio RRR=2.36; 95%CI: 1.22-4.29), but not with the potential mediator. No significant exposure-mediator and exposure-outcome associations were found for asthma with nasal allergies in offspring.
In the maternal line, mothers' overweight at age 8 and in puberty were strongly associated with offspring's overweight at age 8 (OR 2.45; 95%CI 1.86-3.22 and OR 2.13; 95%CI 1.26-3.60, respectively), but these susceptibility windows were not significantly associated either with the potential mediator or with the outcome.
Therefore, the mediation analysis was carried out only for offspring's asthma without nasal allergies (vs no asthma) within the paternal line. We found a slight but statistically significant indirect effect of fathers' overweight at age 8 on offspring's asthma without nasal allergies mediated by offspring's overweight at age 8 (indirect-only mediation). The effect of fathers' overweight in puberty on offspring's ever asthma without nasal allergies (OR 2.23; 95%CI 1.06-4.02) did not involve offspring's overweight at age 8 (direct-only non-mediation).

Conclusion

Our findings suggest that factors long before conception may increase the risk for asthma, and that male puberty is a time window of particular importance for future generations.

References

[1] Accordini S, Calciano L, Johannessen A, et al. A three-generation study on the association of tobacco smoking with asthma. International journal of epidemiology. 2018.

[2] Eder W, Ege MJ, von Mutius E. The asthma epidemic. The New England journal of medicine. 2006;355(21):2226-35.

[3] Burney PG, Luczynska C, Chinn S, Jarvis D. The European Community Respiratory Health Survey. Eur Respir J. 1994;7(5):954-60.

[4] Bandalos DL. Relative Performance of Categorical Diagonally Weighted Least Squares and Robust Maximum Likelihood Estimation. Structural Equation Modelling: A Multidisciplinary Journal. 2014;21(1):102-16.

L'USO DEI CORTISONICI PER AEROSOL NEI BAMBINI IN ETÀ PRESCOLARE: UN'ANALISI DELL'APPROPRIATEZZA PRESCRITTIVA NELLA REGIONE LAZIO

Alagna Michela[1], Lallo Adele[1], Reali Laura[2], Cortis Elisabetta[3], Di Martino Mirko[1], Davoli Marina[1]

per la Commissione Regionale del Farmaco (Co.Re.Fa.).

[1] *Dipartimento di Epidemiologia del Servizio S.S.R – Regione Lazio – ASL Roma 1*

[2] *Pediatra di famiglia – Regione Lazio*

[3] *Ospedale S. Eugenio, ASL Roma 2*

Introduzione

In Italia, l'utilizzo dei farmaci a base di cortisone per aerosol è largamente diffuso tra i bambini con età compresa tra 0 e 5 anni [1]. Tali farmaci vengono essenzialmente prescritti per il trattamento delle patologie delle alte vie respiratorie e per il controllo del sintomo 'tosse' ad esse correlato, sebbene non esistano evidenze circa la loro efficacia [2, 3]. L'utilizzo abituale dei cortisonici inalatori nelle infiammazioni delle prime vie respiratorie nei bambini è stata indicata al primo posto nella classifica delle cinque pratiche ad alto rischio di inappropriatezza stilata dall'Associazione Culturale Pediatri (ACP) [4]. In un recente studio, condotto su 520 bambini in età prescolare con infezioni delle vie aeree superiori, non è emersa alcuna differenza nei tempi di scomparsa dei sintomi tra i soggetti esposti a cortisone per via aerosolica e soggetti esposti a placebo. Inoltre, la terapia cortisonica non ha modificato in modo significativo la probabilità di successive visite pediatriche, accessi in Pronto Soccorso o ricoveri ospedalieri [5].

Obiettivi

- Misurare l'esposizione ai cortisonici inalatori nei bambini in età prescolare nella regione Lazio.
- Misurare la variabilità nelle modalità prescrittive tra ASL, distretti sanitari e pediatri di libera scelta (PLS).
- Quantificare quanta parta della variabilità è attribuibile ai diversi Provider del Servizio Sanitario Regionale (SSR) e identificare gli assi prioritari di intervento.

Metodi

È stato condotto uno studio di farmaco-utilizzazione dopo aver integrato, in un'unica "banca dati di popolazione" diversi Sistemi Informativi della Regione Lazio: Anagrafe degli Assistiti, Registro di Mortalità, Flusso Informativo della Farmaceutica Territoriale, Registro delle Esenzioni Ticket, Anagrafe dei Pediatri di Libera Scelta.
La popolazione in studio è costituita da tutti i bambini di età compresa tra 0 e 5 anni, assistiti e residenti nella regione Lazio al 31/12/2016 (*data indice*).
Il periodo di *follow-up* è di 365 giorni, a partire dalla data indice.

Sono stati esclusi dall'analisi:

- i bambini che, nell'intervallo di osservazione, presentavano almeno tre prescrizioni *proxy* di broncospasmo. In particolare, sono stati considerati i seguenti principi attivi/associazioni precostituite: Salbutamolo (ATC: R03AC02), Ipratropio bromuro (ATC: R03BB01), 'Salbutamolo + Beclometasone' (ATC: R03AK13), 'Salbutamolo + Ipratropio bromuro' (ATC: R03AL02);

- i bambini con un'esenzione attiva per asma (codice esenzione: 007.493) nell'anno di osservazione;
- i bambini che non risultavano sempre assistiti dallo stesso pediatra e i bambini assistiti da pediatri che hanno cambiato distretto o interrotto il servizio nell'anno di *follow-up*.

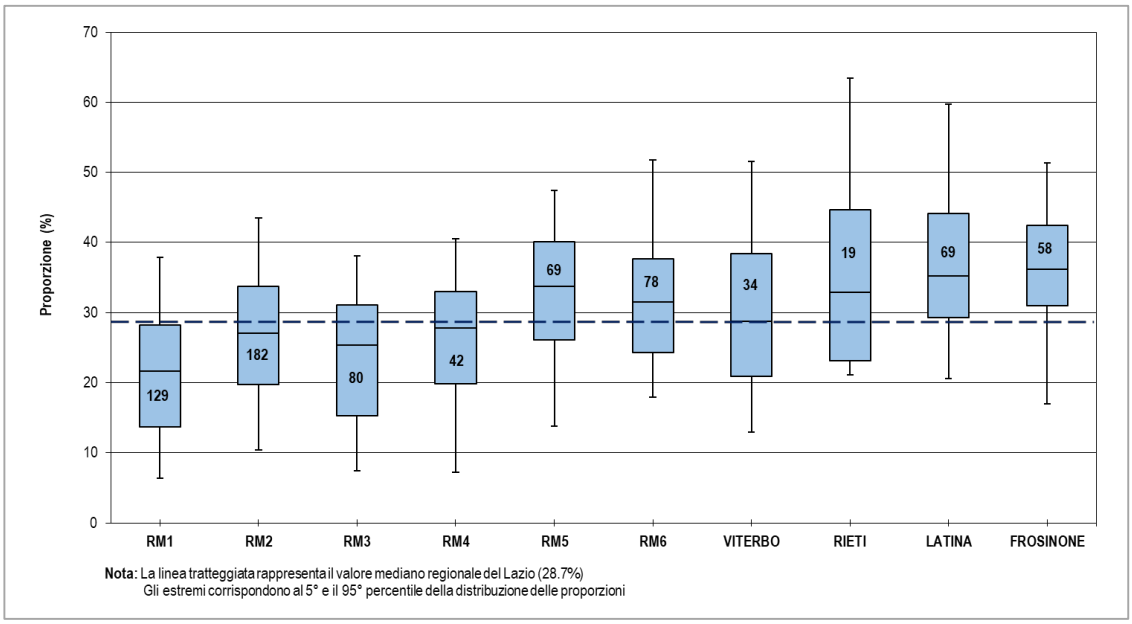
Per la definizione dell'esposizione ai cortisonici inalatori sono stati considerati i seguenti principi attivi: Beclometasone dipropinato (ATC: R03AB01), Budesonide (ATC: R03AB02), Flunisolide (ATC: R03AB03). L'esposizione ai farmaci in studio è stata misurata attraverso la prevalenza di prescrizione, espressa come numero di bambini con almeno una prescrizione di cortisonici per aerosol nel periodo di *follow-up*, per 100 individui. La prevalenza di prescrizione è stata analizzata rispetto al genere, alla tipologia di cortisonico e alla stagionalità. Gli stessi criteri di arruolamento ed esclusione e la stessa misura di esposizione sono stati applicati al periodo 2007-2017, allo scopo di ottenere un trend temporale delle modalità di prescrizione.
Lo studio della variabilità nella prevalenza di prescrizione è stata misurata per singolo PLS, per distretto e per ASL di assistenza. Al fine di quantificare quanta parte della variabilità risulti attribuibile ai vari "livelli" del SSR ed individuare gli assi prioritari di intervento è stato implementato un modello di regressione logistica multilivello, con intercetta casuale, a quattro livelli (paziente-PLS-distretto-ASL), controllando per il sesso del bambino (variabile esplicativa di I° livello), il sesso e gli anni di esperienza del pediatra (variabili di II° livello). Per evitare fenomeni di sovra-dispersione sono stati esclusi dall'analisi i pediatri con meno di 10 assistiti e i distretti con meno di 5 pediatri di libera scelta. L'approccio ha permesso di 'scomporre' la variabilità totale in componenti indipendenti e additive, attribuibili a ciascun 'livello' del SSR. Le componenti della varianza sono state espresse in termini di *Median Odds Ratio* (MOR), che quantifica la variabilità tra gruppi. La misura è sempre maggiore o uguale a 1. Un MOR uguale a 1 indica assenza di variabilità, all'aumentare del MOR aumenta la variabilità tra gruppi. A ciascuno dei Provider analizzati (PLS, distretto sanitario di assistenza e ASL) è stato associato un MOR. In questo contesto, i MOR possono essere interpretati come un sistema dei pesi. Ciascun valore è direttamente proporzionale all'impatto del corrispondete Provider sulla probabilità di prescrivere cortisonici inalatori. Si sottolinea come, dopo aver escluso dalla coorte i bambini con broncospasmo, la prescrizione del cortisonico per aerosol possa essere considerata ad elevato rischio di inappropriatezza.

Risultati

Complessivamente, la popolazione in studio è costituita da 251 705 bambini con età compresa tra 0 e 5 anni. Nei 365 giorni di *follow-up*, il 28.8% dei bambini ha ricevuto almeno una prescrizione di cortisonici inalatori. Dalle analisi è emersa una minore esposizione delle femmine alla terapia inalatoria: nel 2017, per il 27.9% delle bambine in età prescolare si è registrata almeno una prescrizione di cortisonici inalatori, contro il 29.7% dei bambini. Dei farmaci in studio, il Beclometasone risulta quello maggiormente prescritto e rappresenta il 68.9% del totale delle prescrizioni erogate nel 2017, seguito dal Budesonide, con il 27.9% e dal Flunisolide, con il 3.2%.
La prevalenza di prescrizioni di cortisonici per aerosol è stata valutata nel decennio 2007-2017. A seguito di una crescita progressiva della proporzione di bambini trattati con cortisone per aerosol tra il 2007 e il 2012, si evidenzia una prevalenza d'uso in costante diminuzione a partire dall'anno 2013. Per quanto riguarda la distribuzione delle prescrizioni per tipologia di farmaco, in ciascuno degli anni analizzati, il Beclometasone si conferma il farmaco maggiormente prescritto, mentre si osserva una graduale riduzione delle prescrizioni di Fluinisolide durante i dieci anni di osservazione.
Un primo segnale di potenziale inappropriatezza si evince dall'analisi della stagionalità delle prescrizioni: la prevalenza è maggiore nei mesi autunnali e invernali, quando i bambini sono maggiormente esposti alle infiammazioni delle prime vie respiratorie, e decisamente minore in quelli estivi. Nel 2017, la prevalenza mensile variava dallo 0.7% del mese di agosto al 5.6% del mese di gennaio.
Una grande variabilità nelle modalità prescrittive si riscontra anche 'a livello territoriale'.

Nel 2017, la proporzione di bambini esposti a cortisone per via aerosolica, tra i distretti sanitari della regione Lazio, variava da un minimo del 17.9% ad un massimo del 41.0%.
La variabilità nella 'propensione' alla prescrizione di Beclometasone, Budesonide e Flunisolide è ancora più elevata considerando il singolo pediatra di libera scelta: in questo caso, la prevalenza varia dallo 0.4% al 65.3%.
La figura 1 mostra la variabilità tra pediatri di libera scelta nella prevalenza prescrittiva di cortisonici per aerosol, distinta per ASL di assistenza.
Dall'analisi dei box-plot emerge una notevole variabilità sia tra le ASL che entro le ASL. Latina e Frosinone rappresentano le ASL con una maggiore proporzione di bambini esposti ai cortisonici inalatori. Rieti e Viterbo spiccano, invece, per la variabilità interna, tra i pediatri di libera scelta.

Figura 1. *Variabilità nella proporzione di bambini tra 0 e 5 anni con almeno una prescrizione di cortisonici inalatori. Box-plot per pediatra di libera scelta (Lazio, 2017)*



Il modello multilivello ha prodotto, per ciascuna componente della varianza, stime statisticamente significative. Esiste pertanto una variabilità significativa tra i pediatri di libera scelta che afferiscono "al medesimo" distretto, tra i distretti sanitari appartenenti alla medesima ASL e tra le ASL. Il MOR più elevato è stato stimato per il Provider "PLS", che ha il maggior impatto sulla probabilità di esporre i propri assistiti al cortisonico per aerosol (MOR = 1.64, IC 95%: 1.60 – 1.69, p < 0.001). I MOR stimati per gli altri Provider sono sostanzialmente inferiori: livello "ASL", MOR = 1.27 (IC 95%: 1.15 – 1.49, p =0.028); livello "distretto sanitario", MOR = 1.16 (IC 95%: 1.11 – 1.25, p = 0.007).
Gli *Odds Ratio* (OR) stimati per le variabili esplicative di I° e II° livello inserite nel modello sono tutti statisticamente significativi. L'OR relativo al genere del bambino è pari a 0.91 (*reference category*: maschio, IC 95%: 0.89 – 0.92, p < 0.001): la probabilità di essere esposti al cortisonico inalatorio nelle bambine è del 9.0% inferiore rispetto ai bambini. Gli OR relativi al genere e agli anni di esperienza del pediatra, pari, rispettivamente, a 1.14 (*reference category*: maschio, IC 95%: 1.04 – 1.24, p = 0.003) e a 0.99 (IC 95%: 0.989 – 0.997, p = 0.001) indicano che la probabilità di esporre al cortisonico per aerosol i propri assistiti è

maggiore nel caso di pediatri di libera scelta di sesso femminile e diminuisce all'aumentare degli anni di esperienza del medico.

Conclusioni

L'estrema variabilità intra-regionale nell'utilizzo di cortisonici inalatori sottolinea l'assenza di protocolli terapeutici condivisi a livello regionale e nazionale e solleva problemi di equità nell'accesso alle cure ottimali. I risultati del modello multilivello individuano nei pediatri di libera scelta l'asse prioritario di intervento. Le direzioni sanitarie delle ASL e i direttori di distretto dovrebbero coinvolgere i pediatri di libera scelta in interventi di audit volti ad attenuare la sovra-prescrizione di cortisonici per aerosol nella popolazione prescolare e promuovere un utilizzo più appropriato e razionale di questi farmaci.

Bibliografia

[1] Clavenna A., Sequi M., Bonati M., Drug prescribing by Italian paediatricians: an exception? Acta Paediatr, 2010 May; 99(5):754-757.

[2] Castro-Rodriguez J.A., Pedersen S., The role of inhaled corticosteroids in management of asthma in infants and preschoolers. Curr Opin Pulm Med, 2013; 19(1):54-59.

[3] Ducharme F.M., Lemire C., Noya F.J., Preemptive use of high-dose fluticasone for virus-induced wheezing in young children. N Engl J Med, 2009 Jan; 360:339-353.

[4] Associazione Culturale Pediatri (ACP), Le cinque pratiche a rischio d'inappropriatezza di cui medici e pazienti dovrebbero parlare. Quaderni ACP, 2015; 22(3):153-154.

[5] Clavenna A., Sequi M., Cartabia M., et al., Effectiveness of nebulized Beclomethasone in preventing viral wheezing: an RCT. Pediatrics, 2016 May; 31(5):531-532.

SUDDEN CARDIAC DEATH IN SPORT: STUDY OF INCIDENCE AND EVALUATION OF THE EFFECTS OF THE BALDUZZI LAW

Di Rocco Arianna[1], Farcomeni Alessio[1], Alunni Fegatelli Danilo[1], Vestri Annarita[1]

[1] *Department of Public Health and Infectious Disease, Sapienza – University of Rome*

Introduction

Sudden cardiac death (SCD) occurs unexpectedly within 1h of the onset of symptoms or unwitnessed within 24h of the deceased being observed in good health. Some patients die instantly, but most of them have some form of prodromes [1].
Despite considerable progress in terms of prevention and treatment, SCD remains one of the main causes of mortality, which about half of all deaths from cardiovascolar diseases are linked to [2]. The overall and regional incidence of SCD varies greatly depending on three main factors: definition, criteria for excluding subjects from studies, and methodology of data collection.
Rare events of SCD, often reported by the media for its emotional impact on the public, are those associated with sport practice. The incidence rate of SCD in sports is currently unknown. National registers are not available and the magnitude of the SCD may be considerably understimated. In Italy, all surveys are derived from the count of events reported by the press in newspaper or on the web [3].
The main cause of death in sports is sudden cardiac arrest (SCA), causing about 75% of events [4]. In SCA the contractile capacity of the heart is lost and the brain is the first organ to suffer. Cerebral anoxia produces reversible lesions after 4-6 minutes of absence of circulation, and irreversible ones after about 10 minutes. Public access defibrillation, which places semi-automatic defibrillators (AEDs), has the potential to be the single greatest advance in the treatment of cardiac arrest since the development of CPR. Early defibrillation is the strongest determinant of survival following out-of-hospital SCA with survival rates declining to 7% - 10% for each delayed minute of defibrillation. [5]
The Italian "Balduzzi" law, established in 2012, was finally approved on 30 June 2017. Since then, all sports clubs must be equipped with an AED. The law provides detailed guidelines on the provision and use of AEDs, it also requires the presence of personnel trained for the execution of the Basic Life Support - Defibrillation (BLS-D).

Aims & Methods

The aim of the study was to obtain an unbiased estimate of the incidence of SCD during sporting activities, and to assess the effect of the Balduzzi law one and half years after its release.
The data collection was performed from 1 January 2015 to 31 December 2018. All cases of SCD due to probable cardiovascular origin, published on the web by any national and local press, have been prospectively collected. The dataset includes events which took place in sports facilities, outside organised sports events or in school gyms. Deaths in outdoor facilities, such as ski slopes, were also considered, provided they were clearly defined and manned. For each event, characteristics of study population (such as age, gender, sport practiced and place in which the death occurred, and the type of aid provided before the arrival of the ambulance) were collected.
In order to assess the effect of the Balduzzi law, SCDs that occurred during sports activities carried out outside the sports facilities, or at least where the defibrillator is cannot be reasonable reached, were used as a control group.
After collecting data, each event has been screened again, in order to assess the number of sources reporting it.

To estimate the number of deaths due to sport-related SCD during the observation period we then used a capture-recapture approach with unobserved heterogeneity. This corresponds to estimation of the number of deaths never reported on the web.
The Generalized Chao (GC) estimator proposed by Bohning [6] was finally used. Such estimator combines the use of covariates as a source of observed hererogeneity, reducing the bias of the classic Chao one [7], with the guarantees of Chao estimator with respect to unobserved heterogeneity. In order to select the optimal model, the biased Focused Information Criterion (bFIC) proposed by Farcomeni [8] was used. Among all the possible models, bFIC selects the one most likely to minimize the mean-squared error (MSE) of the estimate of the sample size.
A limitation of the GC-bFIC approach devised above is the fact that estimates might be unstable in the presence of small sample sizes or separable covariate-outcomes. To tackle this issue, since GC estimators are essentially based on logistic regression, we replaced it with Firth bias-adjusted logistic models. This lead to substantially more stable and reliable estimates.
To calculate the incidence rates, the denominators were derived from an ISTAT survey, called “The practice of sport in Italy” (19 October 2017), and projected using a polynomial regression model.
The evaluation of the effect of the Balduzzi law was carried out using a causal approach, the difference in difference model (DiD). DiD allows us to assess the causal effect of a policy by comparing a treated group before/after the treatment with a control group before/after the treatment. It is not based on the common assumption that all confounders have been measured, but only on the assumption of independence of hidden confounders and treatment. That is to say, the observation period was divided into two time intervals, t0 and t1, by the release of the Balduzzi law on 30 June 2017. Furthermore, we assumed that the Balduzzi law could have no effect on preventing events on the road or far from a defibrillator. Let k1 be the group in which death occurred in a cardio-protected area, and k0 the group whose events occurred on the road. The DiD estimator for the causal effect of the Balduzzi law was obtained through the following formula: (nt1 – nt0)k1 - (nt1 – nt0)k0. In order to assess significance and to obtain a confidence interval for the causal effect, we used a non-parametric bootstrap strategy.

Results

The collected sample was made up of 425 subjects who died during sporting activities. Of these, 249 occurred in places and conditions that currently fall within the cardio-protection criteria imposed by the Balduzzi law (we called it “AED area"). 97% of deaths were in men with a male/female ratio of 29:1; this percentage also remains unchanged in the treated subgroup.
Median age was 52 [IQR 43 - 60], 49 [IQR 39 – 57] for the treated; among subjects who died in a cardio-protected area, 66.7% was between 35 and 65 years of age, while 22.1% was under 35.
The estimate obtained using the GC estimator is 434 [IC 95% 418 – 449], taking into account the complete sample and 257 [IC 95% 245 – 270] considering only deaths occurred in AED area. It can be noted that due to the strong emotional consequences of a SCD in sports, under reporting is minor and only few cases were not included in our list.
Analyses will also be presented on the stratified sample for the main characteristics: age class (young ≤ 35, middle age >35 and <65, elderly ≥ 65), gender, practiced sport and year of death.
For the calculation of the incidences of SCD, three denominators were used, referring to the italian population over three years old: the total, people who have declared to practice sport, and people who have declared to practice sport continuously.
Finally, the DiD approach showed that Balduzzi law lead to 22 fewer deaths in the 18 months after its promulgation.

Conclusions

We have highlighted a positive causal effect of the issue of the Balduzzi law, which has significantly reduced the number of deaths in sports facilities after June 30, 2017.
The estimated incidence of SCD is comparable with those reported up until now in literature. However, we are the first to provide not only the estimates of the incidence, but also the 95% confidence intervals, whitout limiting ourselves to the pure description of the phenomenon.
In further work, we will try to investigate why the male sex is more at risk of SCDs during the practice of physical activity.

References

[1] Bayés de Luna A, Elosua R, Sudden death, Rev Esp Cardiol. 2012; 65(11): 1039 -1052

[2] Myerburg R.J, Castellanos A, Cardiac Arrest and sudden cardiac death (in Braunwald's heart disease: a textbook of cardiovascular medicine) Saunders, Philadelphia (2007)

[3] Zorzi A, Susana A, Spadotto V, et al. L'arresto cardiaco negli impianti sportivi nel 2015: possibili implicazioni epidemiologiche dell'applicazione della "Legge Balduzzi", G Ital Cardiol. 2016; 17(11): 903-907

[4] Harmon KG, Asif IM, Klosser D, et al. Incidence of sudden cardiac death in National Collegiate Athletic Association athletes, Circulation. 2011; 123: 1594 – 1600

[5] Niemann JT, Stratton SJ, Cruz B, et al. Outcome of out of hospital postcountershock asystole and pulsless electrical activity versus primary asystole and pulseless electrical activity, Crit Care Med, 2001; 29(12): 2366 – 2370

[6] Bohning D, Vidal-Diez A, Lerdsuwansri R, et al. A Generalization of Chao's Estimator for Covariate Information, Biometrics. 2013; 69: 1033 – 1042

[7] Chao A. Estimating the population size for Capture-Recapture Data with unqual catchability, Biometrics 1987; 43: 783 – 791

[8] Farcomeni A. Fully general Chao and Zelterman estimators with application to a whale shark population, Appl. Statist (2018); 67(1): 217 – 229

LEVELLING OF CANCER RATES IN THE VERY OLD?

Carioli Greta[1], Hashim Dana[2], Malvezzi Matteo[1], Bertuccio Paola[3], Waxman Samuel[2], Negri Eva[3], La Vecchia Carlo[1], Boffetta Paolo[2,4]

[1] *Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milan, Italy.*

[2] *Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.*

[3] *Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milan, Italy.*

[4] *Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy*

Introduction

Cancer incidence and mortality increase with age and as life expectancy lengthens; with more people surviving to older ages [1], the number of cancer cases and cancer deaths is projected to increase [2].
It is however unclear how aging contributes to variations in cancer rates. Differences in exposure to risk factors over time may modify cancer risk in birth cohorts, making it difficult to distinguish between whether the oldest old were simply exposed, throughout the lifespan, to fewer risk factors for certain types of cancers [3], being less susceptible to cancer, or whether this result is an artifact due to less intense diagnostic screening or testing in old age [4]. On the other hand, the theory of mutation accumulation should be also considered, i.e. genomic damages accumulate with age during life, increasing the risk of cancer [5, 6].

Objective

To understand the state of cancer burden among the elderly, we described cancer incidence and mortality trends by age and calendar period, with emphasis on individuals 85 and older.

Methods

Mortality data. We retrieved official death certification data from the World Health Organization (WHO) database, available on electronic support (WHOSIS) [7] for all cancers combined and several cancer sites: oral cavity and pharynx (C00-C14), esophagus (C15), stomach (C16), colorectum (C18-C21), liver (C22), pancreas (C25), larynx (C32), lung (C33-C34), breast (C50), uterus (cervix and corpus) (C53 and C54), ovary (C56), prostate (C61), bladder (C67), kidney (C64-C65), non-Hodgkin lymphoma (C81-C90, C96), multiple myeloma (C90), and leukemia (C91-C95). We obtained mortality figures from 2000 through 2014 for 8 major countries worldwide, i.e. those with over 20 million inhabitants and over 95% population death certification coverage [7]: the United States of America (USA), Japan, Australia, the United Kingdom (UK), Germany, France, Italy and Poland. During the calendar period considered, in the WHO database, the 10th Revision of the International Classification of Diseases (ICD) [8] was mainly used (in Italy for 2000-2002 the 9th Revision of the ICD [9] was used, however, since coding differences between various revisions were in general minor, we recoded all cancer sites deaths according to the 10th).
Incidence data. We used Surveillance, Epidemiology, and End Results Program (SEER) 9 Registries [9] to retrieve incidence data on older age groups for the USA. We selected the period 2000-2014 and the same cancer sites considered for mortality data, with some minor differences: we added data on brain and central nervous system tumors (C70-72) for both sexes and thyroid (C73) for men. We did not consider larynx for female because of paucity of data and we were able to distinguish uterine cervix from endometrium.
Population data. We obtained estimates of the resident population for the corresponding calendar periods, based on official censuses, from the WHO and, for the USA, from the Pan American Health Organization (PAHO) databases [10]. Since the main focus of this analysis was on older population groups (85-89, 90-

94, and 95+ years), and since data for these age groups was missing in the WHO and PAHO databases, we retrieved population figures in the older age groups from United Nations Census Bureau database (UN) [11]. Data validity and consistency were cross-checked between the databases. In the cancer incidence analysis, since the SEER 9 Registry does not include population data for 85-89, 90-94, 95+ years age groups, we calculated pseudo-SEER populations using the proportion between UN population and SEER population for the 85+ years age group, stratifying by sex and calendar period. We applied the obtained proportions to the older age groups from the UN population.

Statistical analysis. Using the matrices of certified deaths, cancer cases and resident population, we calculate age-specific mortality and incidence rates for each 5-year age group (from 85 to 95+ years) and the 2000-2014 period. For comparison, we also computed rates for 5-year age groups between 65 and 84. We used joinpoint regression models [12], allowing for up to two joinpoints, to identify age groups in which there were significant changes in mortality trends for the 8 countries and the 17 cancer sites under study plus all cancers combined. We computed the estimated annual percent change (APC) for each of the trends identified by the joinpoint model and the average annual percent change (AAPC) over all age groups considered [13, 14] as a summary measure.

Results

Mortality. Mortality trends over the period 2000-2014 from all cancer combined and for both sexes tended to decrease or flatten at the older age groups (after the age 90-94 for most considered countries and after 85-89 for Poland). In contrast, mortality rates for the USA and France also increased after age 85-89 (Figure 1). Mortality rates were consistently higher in men compared to women.

With the exception of France (0 joinpoints), in men, we identified significant changes in mortality trends at the 80-84 age group and, for the UK and Germany, at 85-89. After the identified joinpoint, the UK showed a tendency to level off (APC of -0.8, non significant), Germany and Poland to decrease (APCs of -5.7 and -12.6, respectively, non significant). AAPCs in men, over all age groups considered, ranged between 21.4 (Germany) and 33.6 (Australia). In women, we identified joinpoints at the 75-79 age group in the USA, and at 85-89 for other countries. Italy showed a tendency to flatten (APC of 0.3), while the UK and Poland to decline (APCs of -1.6 and -12.8, respectively), but none significantly. AAPCs in women ranged between 11.2 (Poland) and 36.1 (Japan).

The highest mortality rates among the very old (85+ years) were for cancers of colorectum, lung, breast, and prostate. For Japan only, stomach cancer mortality was higher than colorectal cancer mortality. Colorectal cancer trends reached a peak at age 90-94 to then decrease for most countries and both sexes, except for the USA and France (rises across all age groups) and Poland (peak at earlier age). In men, lung cancer mortality rates started to decline around the age group 80-84 and 85-89, while in women trends across the various ages were less favourable. Breast cancer mortality increased with age, except for Germany, Italy and Poland, where rates stabilized after age 90-94. Mortality trends for prostate cancer increased across all age groups in the USA, Japan and France, while in the other countries declined after age 90-94 (in Poland after 85-89).

Incidence. Total cancer incidence in men from the USA increased up to the age group 80-84, flattened up to 90-94, to decrease thereafter. In women, after age 80-84, the trend declined.

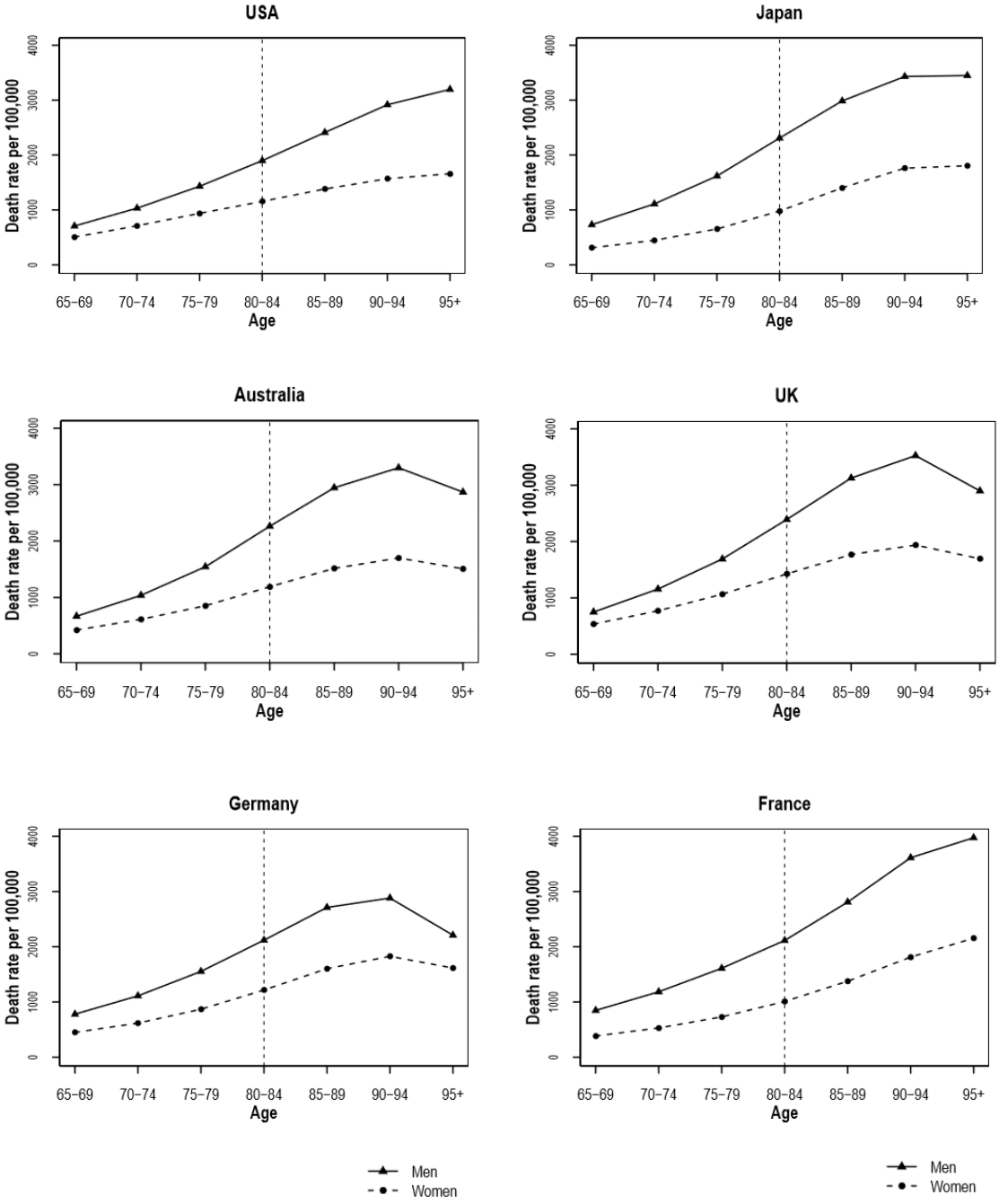
Considering specific cancer sites, the highest incidence rates among US persons aged 85+ were again for cancers of colorectum, lung, breast, and prostate. Colorectal cancer incidence showed declines after the 90-94 age group for men, and for women a plateau between 85-89 and 90-94 age groups, to decline thereafter. Incidence rates for lung cancer increased up to the age 80-84 in men and up to 75-79 in women, to then become favourable. Breast cancer incidence rates in the elderly continued to decline since the age 75-79. Incidence rates for Prostate cancer decreased up to the age 90-94, to then flatten.

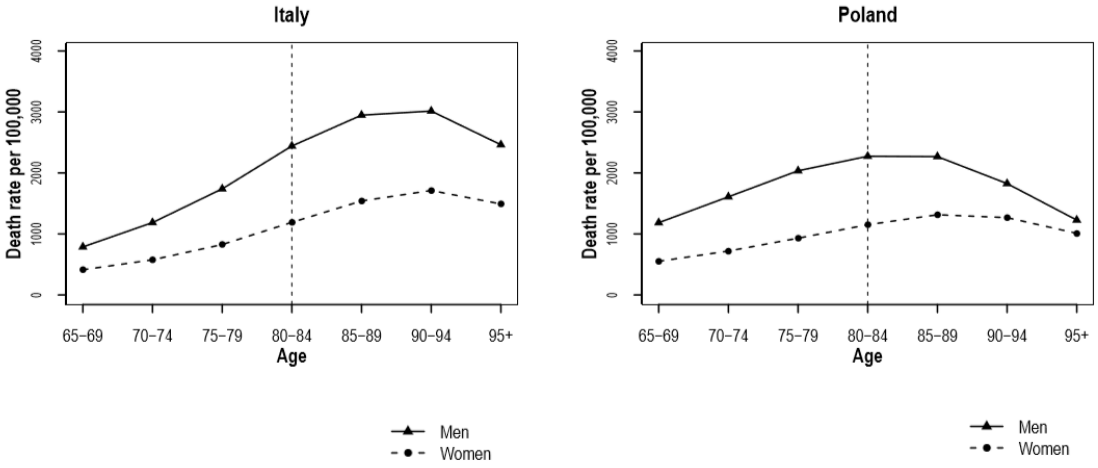
Conclusions

Overall cancer mortality decreased at older age groups, mostly after 90 years of age. Incidence trends showed declines at earlier age compared to those of mortality trends.
Considering specific cancer site, elderly colorectal cancer patients tend to have a more advanced disease stage, they also have more comorbidities and are treated less aggressively compared their younger counterparts. Advanced colorectal cancer often requires major surgery as part of treatment, the comorbidities heavily influence surgical eligibility and potential benefit from other treatment options [15]. This is probably why we found less favorable trends in older age compared to other cancer sites. However, the combination of comorbidity conditions and the avoidance of higher risk therapies, that may cause non-cancer death, could explain the declines observed in the oldest age group.
Lung cancer mortality and incidence rates for the USA have similar trends across age groups, this is due to the low survival for this cancer site. Female smoking patterns reached similarity with those of men worldwide [16], this probably caused the observed differences in trends between the two sexes.
While some studies have proposed that older women may present more aggressive breast cancer or may be at a higher risk due to being outside the screening coverage, others studies have documented that elderly women have a less aggressive disease and a more favourable tumour marker profile [17]. Further, the decreased levels of sex hormones in older adults contributes to a hormonal environment that is less likely to stimulate the growth of breast cancer [18].
The favourable trends in mortality and incidence rates for prostate cancer are probably mainly due to a reduction in overdiagnosis, particularly in the USA (screening was stopped for men over age 75 in 2008 and for all men in 2013) [19].
A paucity of data regarding cancer mortality trends in persons aged 85+ has been noted. Health conditions in the elderly are complex and heterogeneous (comorbidities, multiple drug use, etc.), causing difficulties in cancer care. As a consequence, older people are commonly excluded from clinical trials or aggregated as 85+ years in population-based analyses [20-22].
This work examined cancer mortality and incidence among individuals aged 85 years or more, by major cancer sites and countries using accurate estimates that are based on the quality of national incidence and mortality data of SEER, WHO and PAHO databases [7, 23]. While some studies have found that the population of the very old in the USA is overestimated, leading to underestimated rates for those over 85 years [24, 25], we have utilized validated PAHO/WHO data [26]. Incidence and mortality rates were also stratified by period trends to elucidate trends that may be due to changes in diagnostic criteria and treatment improvements.
However, identifying the underlying cause of death may be challenging in older adults who may present several non-cancer comorbidities; changes in criteria for cancer death diagnosis and inaccuracy in coding and classification likely affect mortality and incidence trends over time. Quantifying the consequences of uncertainties in cancer death certification on observed changes in rates remains difficult. Recent improvements in early diagnosis and cancer death classification have influenced incidence data, but their role on mortality trends is still not entirely clear. Another limitation of this analysis is that we were only able to calculate incidence trends among the very old in the USA, providing a limited scope of incidence trends among those 85+ years. Lastly, to maximize the accuracy of age-specific mortality trends, our study was limited to countries with large populations and 95% population death certification coverage [7].
Therefore, care must be taken not to extrapolate the interpretation of these findings, particularly to low- and middle- income countries with varying access to diagnostic and treatment resources.
This is the first study of mortality in old age that examines both period and cohort specific effects worldwide. Two different processes seem to have been working together in relation to the changes in cancer mortality during the period 2000-2014. Less exposure to risk factors and better cancer prevention, screening and treatment over the period have caused mortality rates to generally decrease.

Cancer incidence and mortality generally increase with age. However, beyond a peak at 80 to 85 years, mortality is likely determined by factors pertaining to lifestyle and risk factor exposure, diagnosis, and treatment that are site specific.

Figure1. *Cancer mortality trends for all cancer combined, in men and women, across 5-year age groups (from 65-69 to 95+ years of age), in 8 major countries worldwide, 2000-2014.*





References

[1] Chang AY, Skirbekk VF, Tyrovolas S et al. Measuring population ageing: an analysis of the Global Burden of Disease Study 2017. Lancet Public Health 2019; 4: e159-e167.

[2] Berger NA, Savvides P, Koroukian SM et al. Cancer in the elderly. Trans Am Clin Climatol Assoc 2006; 117: 147-155; discussion 155-146.

[3] Salinari G. Rethinking mortality deceleration. Biodemography and Social Biology 2018; 64:62, 127-138.

[4] Pedersen JK, Engholm G, Skytthe A et al. Cancer and aging: Epidemiology and methodological challenges. Acta Oncol 2016; 55 Suppl 1: 7-12.

[5] Medawar PB. An Unsolved Problem of Biology. 1952; H.K. Lewis & Co., London.

[6] Pedersen JK, Skytthe A, Christensen K. Cancer occurrence in offspring of long-lived siblings. In Hofman A (ed) Eur J Epidemiol., EuroEpi2013. Aarhus, Denmark: Springer 2013; pp. S11-S12.

[7] World Health Organization Statistical Information System. WHO mortality database Available at: http://www.who.int/healthinfo/statistics/mortality_rawdata/en/index.html (Last accessed April 2019).

[8] World Health Organization. International Classification of Disease and related Health Problems: 10th revision. Geneva: World Health Organization,1992.

[9] World Health Organization. International Classification of Disease: 9th revision. Geneva: World Health Organization,1977.

[10] Pan American Health Organization (PAHO). Health Statistics from the Americas, 2006 Edition, Chapter VI, 2-20. Available at: http://www.paho.org/English/DD/AIS/HSA2006.htm (Last accessed April 2019).

[11] United Nations DoEaSA, Population Division (2017). World Population Prospects: The 2017 Revision, DVD Edition.

[12] Kim HJ, Fay MP, Feuer EJ et al. Permutation tests for joinpoint regression with applications to cancer rates. (Erratum in: Stat Med 2001;20: 655). Stat Med 2000; 19: 335-351.

[13] Clegg LX, Hankey BF, Tiwari R et al. Estimating average annual per cent change in trend analysis. Stat Med 2009; 28: 3670-3682.

[14] National Cancer Institute. Joinpoint Regression Program, version 4.1.1. Available at: https://surveillance.cancer.gov/joinpoint/ 2014.

[15] Lemmens VE, Janssen-Heijnen ML, Houterman S et al. Which comorbid conditions predict complications after surgery for colorectal cancer? World J Surg 2007; 31: 192-199.

[16] Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. Adv Exp Med Biol 2016; 893: 1-19.

[17] Moss SM, Cuckle H, Evans A et al. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet 2006; 368: 2053-2060.

[18] Balducci L. New paradigms for treating elderly patients with cancer: the comprehensive geriatric assessment and guidelines for supportive care. J Support Oncol 2003; 1: 30-37.

[19] Grossman DC, Curry SJ et al. Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement. JAMA 2018; 319: 1901-1913.

[20] Hutchins LF, Unger JM, Crowley JJ et al. Underrepresentation of patients 65 years of age or older in cancer-treatment trials. N Engl J Med 1999; 341: 2061-2067.

[21] Lewis JH, Kilgore ML, Goldman DP et al. Participation of patients 65 years of age or older in cancer clinical trials. J Clin Oncol 2003; 21: 1383-1389.

[22] Townsley CA, Selby R, Siu LL. Systematic review of barriers to the recruitment of older patients with cancer onto clinical trials. J Clin Oncol 2005; 23: 3112-3124.

[23] Surveillance Epidemiology and End Results Program (SEER). SEER Registries. Available at: https://seer.cancer.gov/registries/ (Last accessed May 2019).

[24] Boscoe FP. Subdividing the age group of 85 years and older to improve US disease reporting. Am J Public Health 2008; 98: 1167-1170.

[25] Hill ME, Preston SH, Rosenwaike I. Age reporting among white Americans aged 85+: results of a record linkage study. Demography 2000; 37: 175-186.

[26] Organization WH. Strengthening civil registration and vital statistics for births, deaths and causes of death: resource kit. World Health Organization. https://apps.who.int/iris/handle/10665/78917 In. 2013.

RISCHIO DI OSPEDALIZZAZIONE PER TROMBOEMBOLISMO VENOSO ED INQUINAMENTO ATMOSFERICO: UN'ANALISI CASE-CROSSOVER SUGLI EFFETTI A BREVE TERMINE DEL $PM_{2.5}$

Matteo Renzi[1], Massimo Stafoggia[1,2], Paola Michelozzi[1], Marina Davoli[1], Francesco Forastiere[3] and Angelo G. Solimini[4]

[1] *Department of Epidemiology, Health Authority Service, ASL Rome 1, Rome, Italy*

[2] *Karonlinska Instituet, Stockholm, Sweden*

[3] *King's College, London, UK*

[4] *Department of Public Health and Infectious Diseases, University of Rome "La Sapienza", Rome, Italy*

Background

L'esposizione ad inquinanti aerodispersi è stata collegata con un aumento del rischio di eventi cardiovascolari nel breve e nel lungo periodo. Tuttavia, sono ancora limitate le evidenze epidemiologiche sull'esistenza di una associazione tra inquinamento atmosferico e tromboembolismo venoso, una delle principali cause di ricovero con un carico globale di circa 10 milioni di casi ogni anno.

Obiettivi

Valutare se 1) l'esposizione a breve termine al particolato atmosferico fine ($PM_{2.5}$) sia associata al rischio di ricovero urgente per TVP o PE a diversi tempi di latenza e se 2) l'eventuale associazione sia modificata dal periodo dell'anno (periodo caldo o freddo), sesso, eta' e comorbidita'.

Metodi

Utilizzando il sistema informativo dei ricoveri ospedalieri, è stata condotta un'analisi case-crossover mediante regressione logistica dei ricoveri urgenti per trombosi venosa profonda (TVP) ed embolia polmonare (EP) nei pazienti >35 anni degli ospedali di Roma durante il periodo 2006-2017. L'effetto dell'esposizione al $PM_{2.5}$ è stato valutato diversi tempi di latenza (lag 0, 0-1, 0-5 e 2-5 giorni), aggiustando per temperatura atmosferica, trend a lungo termine e stagionale, periodo dell'influenza stagionale e cambiamento della popolazione residente durante i periodi di vacanza. La presenza di eventuali modificatori della misura d'effetto è stata valutata inserendo un termine di interazione con la variabile di esposizione.

Risultati

Durante il periodo di studio sono stati registrati 45368 casi di tromboembolismo venoso, dei quali il 74% era costituito da soggetti con TVP in una delle diagnosi di dimissione. Considerando l'intero periodo dell'anno, il rischio di ospedalizzazione per TVP ed EP aumenta in maniera non significativa con la concentrazione cumulativa di PM2.5 registrata nei 5 giorni precedenti il ricovero. La stagione modifica l'effetto del PM2.5 sulle ospedalizzazioni per EP (ma non per TVP) in maniera significativa con un aumento significativo del rischio nei mesi estivi del 19.1% (IC95% 8.3, 31.1) per incrementi di 10 □g/m3 di PM2.5 a lag 0-5 giorni. Questa associazione è risultata robusta quando il modello statistico è' stato aggiustato per la concentrazione di biossido di azoto e quando l'analisi è stata limitata ai casi con EP come diagnosi principale. L'interazione tra PM2.5 e le altre variabili testate è risultata non significativa.

Conclusioni

I nostri risultati suggeriscono la presenza di una associazione positiva durante il periodo estivo tra esposizione a breve termine al $PM_{2.5}$ ed il rischio di ricovero per embolia polmonare ma non per trombosi polmonare profonda.

QUALE INDICE DI VARIABILITÀ DELLA PRESSIONE ARTERIOSA PER MISURARE IL RISCHIO DI MORTALITÀ CARDIOVASCOLARE?

Soranna Davide[1], Parati Gianfranco[1,2], Bilo Grzegorz[1,2], Palomba Andrea[3], Antonella Zambon[3]

[1] *IRCCS Istituto Auxologico Italiano, Milano, Italia*

[2] *Dipartimento di Medicina e Chirurgia, Università Milano-Bicocca, Milano, Italia*

[3] *Dipartimento di Statistica e metodi quantitativi, Università Milano-Bicocca, Milano, Italia*

Introduzione

L'ipertensione arteriosa è universalmente riconosciuta come fattore di rischio della mortalità cardiovascolare [1]. La misurazione della pressione arteriosa può essere ottenuta sia con la tradizionale tecnica Office (media di tre misurazioni prese a distanza di 5 minuti) sia con la tecnica 24-h (media delle misurazioni rilevate durante le ore diurne - ogni 20 minuti - e delle misurazioni nelle ore notturne - ogni 30 minuti) [1,-4]. Ridurre il contributo informativo alla sola media pressoria può però essere limitante. Infatti, la pressione arteriosa (PA) è un fenomeno caratterizzato da forte variabilità temporale (VPA). Questa variabilità può essere classificata come: i) a breve termine (valutata nelle 24 ore); ii) a medio termine (valutata in un periodo di giorni) e iii) a lungo termine (valutata in un periodo di mesi o anni). La VPA è nota aumentare all'incremento dei valori di PA [5]. Alcuni studi primari e di tipo meta-analitico [6,7] hanno suggerito che un aumento della VPA potrebbe essere un predittore indipendente degli effetti clinici legati all'ipertensione. In particolare, l'aumento della VPA a breve termine è stato segnalato come fattore di rischio associato al danno d'organo ed a eventi cardiovascolari [7-9], prevalentemente nei giovani [10]. In letteratura esistono diversi indici che quantificano la VPA a breve termine ma non è chiaro quali di essi siano da preferirsi nei modelli di regressione che considerano anche l'effetto della PA a causa di una potenziale collinearità tra i due indicatori.

Obiettivi

Verificare l'impatto della collinearità tra PA e VPA (utilizzando diversi indici) sulle stime di mortalità cardiovascolare ottenute da un modello di regressione a rischi proporzionali di Cox e sulla relativa capacità predittiva del modello.

Metodi

Questo lavoro si basa sui dati del Dublin Outcome Study. Lo studio includeva i pazienti che si presentavano al Beaumont Hospital di Dublin in modo consecutivo nel periodo 01/06/1982-30/09/2002 (data indice) a causa di un'elevata pressione arteriosa. A tutti i pazienti è stata misurata la pressione "24-h" e venivano registrate le seguenti informazioni: genere, età, BMI, frequenza cardiaca, fumo (sì/no), diabete (sì/no), precedenti eventi cardiovascolari (sì/no). Sono stati esclusi i pazienti: i) con età inferiore ai 18 anni, ii) con meno del 70% delle misurazioni attese (direttiva dell'ESH position paper [4]) e iii) con valori anomali di pressione (pressione sistolica inferiore a 50 mmHg oppure superiore a 300 mmHg), iv) con meno di un anno di follow-up. Ogni membro della coorte finale ha accumulato anni-persona di follow-up dalla data indice fino al primo dei seguenti eventi: morte, migrazione o 30 Settembre 2002. L'evento di interesse era rappresentato dalla mortalità cardiovascolare (CV). Sono stati considerati i seguenti indici di variabilità: i) Standard Deviation (SD); ii) Standard deviation pesata (wSD – media ponderata per le ore diurne e notturne); iii) Day SD (SD delle sole ore diurne), iv) Average real variability (ARV – media delle differenze assolute tra letture consecutive di pressione arteriosa), v) Residua (potenza dello spettro pressorio

rimanente dopo la rimozione delle prime due componenti principali), vi) Coefficiente di variazione (CV); vii) Coefficiente di variazione pesato (wCV – calcolata con la wSD); viii) Variability indipendent of the mean (VIM); ix) Weighted Variability indipendent of the mean (wVIM – calcolato con la wSD). Gli indici di variabilità sono stati analizzati dopo opportuna standardizzazione. I pazienti inclusi sono stati stratificati in 4 categorie di età: <50 anni, 50-65 anni, 65-75 anni e > 75 anni. Per ognuno di questi strati, l'effetto della PA e della variabilità sul rischio di mortalità cardiovascolare è stato analizzato attraverso un modello di regressione a rischi proporzionali di Cox aggiustando per le seguenti covariate: genere, diabete, fumo, precedenti eventi cardiovascolari, BMI, la caduta notturna di pressione, frequenza cardiaca e media della pressione sistolica. Gli assunti di rischi proporzionali sono stati valutati attraverso i residui di Schoenfield [11]. Per ciascun modello vengono riportate le stime dei parametri di regressione della PA e degli indici VPA (con i relativi standard error e p-value) e le corrispondenti misure di capacità predittiva C-index di Harrell [12].

Risultati

Nell'analisi erano inclusi 9154 pazienti di cui il 47% di genere maschile e con un range di età che variava dai 18 ai 94 anni. Il tempo mediano di follow-up era di 6,3 anni durante il quale si sono verificati 429 decessi per eventi cardiovascolari. I soggetti più anziani erano caratterizzati da una maggior proporzione di donne così come una più alta prevalenza di diabete, precedenti malattie cardiovascolari, più basso tasso di fumo e più basso BMI (dati non riportati). La tabella riporta le stime dei parametri di regressione della PA sistolica e degli indici di VPA (con i relativi standard error e p-value) e la C-index dei diversi modelli stratificati per classi di età. Come si può osservare solo nella categoria di pazienti più giovani emerge un ruolo statisticamente significativo della VPA (indipendentemente dall'indice considerato). Il ruolo della PA risulta statisticamente significativo solo per gli indici CV e VIM nella forma pesata e non. Per questi indici, inoltre, lo standard error dell'effetto della PA sembra leggermente inferiore. Nello strato dei giovani gli indici VPA che mostrano valori più elevati della C di Harrell sono gli indici CV e VIM nella forma pesata.

	Pressione Sistolica				
	Variabilità		Media		
	β (SE)	P-value	β (SE)	P-value	C Harrell (SE)
Età < 50 anni					
SD	0.365 (0.165)	0.027	0.014 (0.010)	0.158	0.780 (0.04)
Weight SD	0.433 (0.117)	< 0.001	0.012 (0.010)	0.204	0.803 (0.04)
Day SD	0.456 (0.120)	< 0.001	0.014 (0.010)	0.150	0.808 (0.04)
ARV	0.288 (0.132)	0.029	0.016 (0.010)	0.101	0.773 (0.05)
Residua	0.370 (0.124)	0.003	0.013 (0.010)	0.176	0.778 (0.05)
CV	0.365 (0.164)	0.026	0.024 (0.009)	0.011	0.780 (0.04)
Weight CV	0.395 (0.115)	0.001	0.025 (0.009)	0.008	0.803 (0.04)
VIM	0.362 (0.163)	0.026	0.023 (0.009)	0.014	0.780 (0.04)

Weight VIM	0.395 (0.114)	0.001	0.025 (0.009)	0.009	0.803 (0.04)
50 <Età < 65 anni					
SD	-0.013 (0.100)	0.894	0.019 (0.005)	< 0.001	0.757 (0.02)
Weight SD	0.081 (0.081)	0.321	0.017 (0.005)	< 0.001	0.760 (0.02)
Day SD	0.091 (0.078)	0.245	0.017 (0.005)	0.001	0.759 (0.02)
ARV	0.093 (0.084)	0.271	0.017 (0.005)	0.001	0.760 (0.02)
Residua	-0.009 (0.088)	0.922	0.019 (0.005)	< 0.001	0.757 (0.02)
CV	-0.020 (0.099)	0.840	0.018 (0.005)	< 0.001	0.757 (0.02)
Weight CV	0.074 (0.081)	0.361	0.019 (0.005)	< 0.001	0.760 (0.02)
VIM	-0.019 (0.098)	0.844	0.018 (0.005)	< 0.001	0.757 (0.02)
Weight VIM	0.074 (0.080)	0.356	0.019 (0.005)	< 0.001	0.759 (0.02)
65 < Età < 75 anni					
SD	-0.017 (0.096)	0.862	0.011 (0.005)	0.014	0.656 (0.02)
Weight SD	-0.027 (0.086)	0.755	0.011 (0.005)	0.011	0.656 (0.02)
Day SD	0.013 (0.084)	0.881	0.011 (0.004)	0.015	0.658 (0.02)
ARV	-0.047 (0.085)	0.578	0.012 (0.004)	0.008	0.655 (0.02)
Residua	0.026 (0.086)	0.764	0.010 (0.005)	0.022	0.658 (0.02)
CV	-0.027 (0.094)	0.777	0.011 (0.004)	0.009	0.656 (0.02)
Weight CV	-0.025 (0.082)	0.759	0.011 (0.004)	0.009	0.656 (0.02)
VIM	-0.025 (0.093)	0.784	0.011 (0.004)	0.008	0.656 (0.02)
Weight VIM	-0.026 (0.082)	0.754	0.011 (0.004)	0.008	0.656 (0.02)
Età > 75 anni					
SD	0.102 (0.117)	0.385	0.012 (0.006)	0.043	0.637 (0.03)
Weight SD	0.112 (0.109)	0.306	0.011 (0.006)	0.052	0.640 (0.03)
Day SD	0.150 (0.104)	0.150	0.011 (0.006)	0.052	0.642 (0.03)

ARV	0.074 (0.107)	0.486	0.012 (0.006)	0.029	0.635 (0.03)
Residua	0.023 (0.111)	0.839	0.013 (0.006)	0.023	0.634 (0.03)
CV	0.115 (0.109)	0.289	0.014 (0.005)	0.008	0.638 (0.03)
Weight CV	0.119 (0.289)	0.232	0.014 (0.005)	0.009	0.641 (0.03)
VIM	0.112 (0.108)	0.302	0.014 (0.005)	0.009	0.638 (0.03)
Weight VIM	0.118 (0.100)	0.234	0.014 (0.005)	0.009	0.641 (0.03)

Conclusioni

Gli indici pesati e indipendenti dalla media per costruzione (CV e VIM) sembrano essere un buon strumento per la stima dell'effetto della VPA consentendo di controllare i problemi di collinearità nei modelli che considerano anche come covariata la PA. Il CV pesato potrebbe essere l'indicatore da preferire basandosi su un criterio di semplicità di applicazione.

Bibliografia

[1] Williams B, Mancia G, Spiering W, et al. 2018 ESC/ESH Guidelines for the management of arterial hypertension. Eur Heart J 2018; 39: 3021–3104.

[2] Mancia G, Facchetti R, Bombelli M, et al. Long-Term Risk of Mortality Associated With Selective and Combined Elevation in Office, Home, and Ambulatory Blood Pressure. Hypertension 2006; 47: 846–853.

[3] Parati G, Stergiou GS, Asmar R, et al. European Society of Hypertension guidelines for blood pressure monitoring at home: a summary report of the Second International Consensus Conference on Home Blood Pressure Monitoring. J Hypertens 2008; 26: 1505–1526.

[4] O'Brien E, Parati G, Stergiou G, et al. European Society of Hypertension position paper on ambulatory blood pressure monitoring. J Hypertens 2013; 31: 1731–68.

[5] Mancia G, Ferrari A, Gregorini L, et al. Blood pressure and heart rate variabilities in normotensive and hypertensive human beings. Circ Res 1983; 53: 96–104.

[6] Parati G, Ochoa JE, Lombardi C, et al. Assessment and management of blood-pressure variability. Nat Rev Cardiol 2013; 10: 143–155.

[7] Stevens SL, Wood S, Koshiaris C, et al. Blood pressure variability and cardiovascular disease: systematic review and meta-analysis. BMJ 2016; 354: i4098.

Galeone C., Pelucchi C., Levi F. et al., Onion and garlic use and human cancer. Am J Clin Nutr., 2006 Nov; 84(5):1027-32

[8] Parati G, Pomidossi G, Albini F, et al. Relationship of 24-hour blood pressure mean and variability to severity of target-organ damage in hypertension. J Hypertens 1987; 5: 93–8.

[9] Frattola A, Parati G, Cuspidi C, et al. Prognostic value of 24-hour blood pressure variability. J Hypertens 1993; 11: 1133–7.

[10] Mehlum MH, Liestøl K, Kjeldsen SE et al. Blood pressure variability and risk of cardiovascular events and death in patients with hypertension and different baseline risks. Eur Heart J. 2018 Jun 21;39(24):2243-2251.

[11] Schoenfeld D. Partial Residuals for The Proportional Hazards Regression Model. Biometrika 1982, 69:239-241

[12] Harrell FE, Califf RM, Pryor DB et al. Evaluating the yield of medical tests. Journal of the American Medical Association. 1982;247:2543–46.

DESIGN AND FEASIBILITY OF A MULTICENTRIC STUDY ASSESSING BLOOD MICROBIOTA AND DIET IN RELATION TO ADENOMA AND COLORECTAL CANCER RISK.

Marcello Cintolo[1], Roberto Penagini[2], Pietro Carnevali[3], Aldo Airoldi[4], Pietro Soru[2], Federica Mascaretti[5], Clorinda Ciafardini[2], Rossella Bonzi[5], Giovanni Ferrari[3], Carmelo Magistro[3], Barbara Oreggia[6], Marcello Vangeli[4], Maurizio Vecchi[2], Carlo La Vecchia[5], Simone Guglielmetti[7], Massimiliano Mutignani[1], Marta Rossi[5].

[1]*SC Endoscopia Digestiva e Interventistica, ASST Grande Ospedale Metropolitano Niguarda – Milan, Italy;*

[2]*Gastroenterology and Endoscopy Unit, Fondazione IRCCS Cà Granda, Ospedale Maggiore Policlinico; Università degli Studi di Milano, Italy;*

[3]*Chirurgia Generale Oncologica e Mininvasiva, ASST Grande Ospedale Metropolitano Niguarda – Milano*

[4]*S.C. Epatologia e Gastroenterologia, ASST Grande Ospedale Metropolitano Niguarda, Milano.*

[5]*Department of Clinical Sciences and Community Health, Università degli Studi di Milano – Milan, Italy;*

[6]*Chirurgia Generale, Fondazione IRCCS Cà Granda, Ospedale Maggiore Policlinico; Università degli Studi di Milano, Italy;*

[7]*Department of Food, Environmental and Nutritional Sciences (DeFENS), Division of Food Microbiology and Bioprocesses; Università degli Studi di Milano – Milan, Italy.*

Introduction

Colorectal cancer (CRC) is the 3rd more common cancer in males and the 2nd one in females in Italy (as in the European Union), and ranks as the 2nd cause of cancer death in both sexes, with about 48,000 new cases and 19,000 deaths in 2012 [1, 2].
Colonoscopy removal of adenomatous polyps is the standard care to efficiently reduce CRC mortality [3]. Therefore, population-wide screening programs currently uses fecal occult blood test to identify patients who should undergo colonoscopy. However, given its limited sensitivity and specificity, there is a need of non-invasive and more accurate screening tests [4, 5].
Environmental factors associated with intestinal adenoma (IA) and CRC include selected dietary habits (such as high red-meat consumption, low intakes of plant foods, fiber, calcium, and other dietary compounds), sedentary lifestyles, high body mass index (BMI) alcohol drinking, tobacco smoking, and history of type 2 diabetes [6-8]. Particularly, diet has been suggested to exert a dominant role over other environmental variables in shaping compositional features of the gut microbial ecosystem.
CRC derives from a sequential accumulation of genetic alterations that involves the transition from normal mucosa to pre-malignant lesions with progression to intestinal adenoma (IA) and invasive CRC [9]. The bacterial microbiota may influence colorectal carcinogenesis through various mechanisms inlcuding bacteria-derived genotoxin production, microbial-derived metabolism, the modulation of host defenses and inflammation pathways. Microbe-associated molecular patterns have been reported to sustain and exacerbate inflammation in colorectal tumours, mediating the loss of barrier function that can also lead to increased bacterial translocation into the lamina propria, where bacteria (or bacterial products) may further promote pro-inflammatory responses likely increasing tumorigenesis [10-12]. Plausibly, disruption of the epithelial barrier during CRC results in a leaky gut that permits the translocation of intestinal microbes or their components into bloodstream, where may lead to systemic inflammation. The presence of *Streptococcus bovis*, biotype 1 in blood has been proposed for many years as a predictor of IA or CRC cancer [13].

Objective

We designed a study to evaluate the existence of a diet-microbiota axis as a prominent factor predisposing IA and CRC and to investigate whether bacterial translocation from gastrointestinal tract to bloodstream is associated to the risk of IA and/or CRC (Figure 1). A double epidemiological and metagenomic approach will be used to address the specific project's objectives below:

1) the relation between the bacterial DNA load in blood and the risks of IA and CRC;
2) the relation between bacterial taxonomic groups (assessed by 16S rRNA profiling) in blood and the risks of IA and CRC;
3) the relation between dietary factors and the blood microbiota (bacterial DNA load and bacterial taxonomic groups) among controls, IA and CRC cases;
4) the effect of dietary patterns on the risks of IA and CRC.

These results will allow to identify potential predictive bacterial biomarkers for an early diagnosis of CRC, as well as to identify dietary guidelines associated to a risk reduction of CRC.

Methods

The *ongoing* project includes the conduction of a case-control study collecting epidemiological information and blood samples of 100 CRC cases (histologically confirmed), 100 matched IAs and 100 matched healthy controls (by gender and age: +/- 5 years) in two hospitals of Milan. Subjects are recruited among patients (aged 20-85) scheduled for a colonoscopy in one of the participant hospitals. Colonoscopy reports are revised by an expert endoscopist who determines CRC cases, IA cases and non-cancer/IA hereinafter (healthy) controls.

Ad hoc and trained interviewers contact eligible subjects during the colonoscopy preparatory appointment. Participants are face-to-face interviewed by interviewers through an *ad hoc* developed questionnare. The questionnaire includes information on socio-demographics, smoking habit, physical activity, occupational exposures, medical history, dental history, drug use (e.g. antibiotics), sleeping habits and family history of cancer. All participants sign the informed consent. A food frequency questionnaire (FFQ) was developed on the basis of a reproducible and valid FFQ [7] assessing patients' usual diet during the past as well as the recent patients' diet. Blood samples are also collected before colonoscopy from each subject by a physician and the blood are immediately stored at -80 ºC after extraction.

Subjects' past and recent intakes of selected foods, micro- and macro-nutrients and total energy from dietary information will be derived using *ad hoc* developed Italian food data composition tables and programs [14]. Similarly, total antioxidant capacity index and dietary inflammatory index will be estimated using published international data opportunely adapted for the Italian diet [15, 16].

The total bacterial load in blood samples will be determined through quantitative real-time PCR on extracted DNA with 16S rRNA gene [17].

The relation of various taxonomic levels with the risks of IA and CRC will be estimated through multiple logistic regression. Polytomous logistic regression will be used to estimate separate odds ratios (ORs) for IA and to test for heterogeneity between the two diseases. Additional analysis will be also performed to separately analyze colon and rectum, and cancer specific subtypes. The relation between various dietary factors and the relative abundance of bacteria at various taxonomic levels will be estimated trough t-test, analysis of variance (ANOVA), multivariate regression models, and permutation multivariate analysis of variance. Analyses will be separately performed among controls, cases of IA and cases of CRC. Technique of factor analysis (e.g. PCA, PCOA,) will be applied to reduce the redundancy of variability of both dietary and microbiome data. Subgroups analysis will be performed in strata of gender, as well as of age, BMI (Body Mass Index), diabetic individuals, and other factors. Particular attention will be paid to the creation of dietary patterns considering implications of all possible confounders (e.g. antibiotic use, smoking habits).

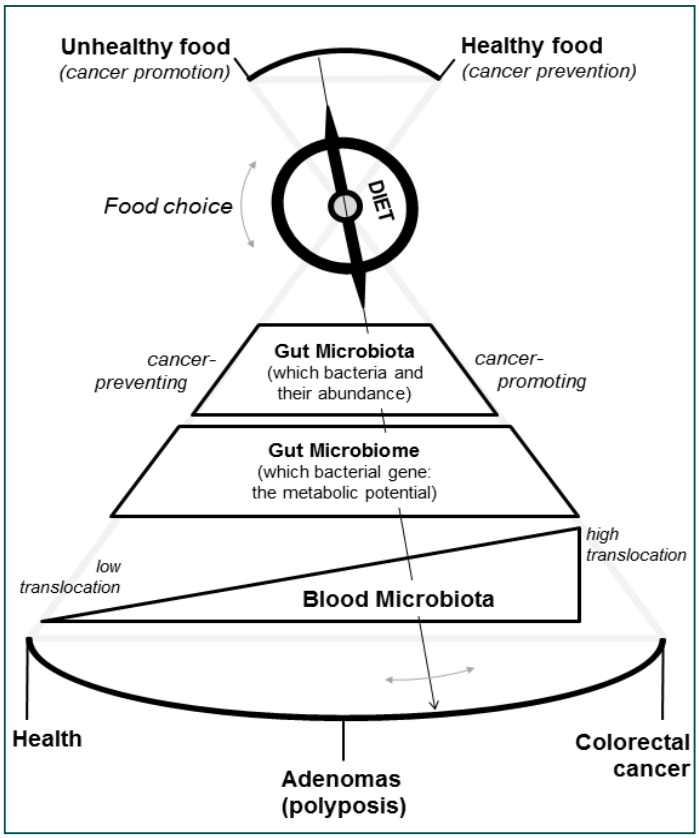
Results

Project received ethical committee approval and data recruitment started in May 2017, with involvement of CRC regional screening program. Three hundred and nine subjects, 136 women and 173 men were recruited, with a total of 81 three matched trios, to May 2019. A portion of subjects will be matched afterwards, since the definition of group (control, IA or colorectal cancer case) will be classified after colonoscopy and histology reports.
In order to keep the signal to noise ratio optimal and to reduce technical variability, possibly overcoming the biological variation between groups, all samples will be analyzed in the same time, with the same reagents batches and by the same manipulator at the end of data recruitment to perform DNA extraction, qPCR quantification and taxonomic profiling by Illumina MiSeq sequencing of the 16S rRNA gene copies from all samples in the same experiment.

Conclusion

An epidemiological and metagenomic approach is used to evaluate the association between colorectal carcinogenesis and specific bacteria translocated into the bloodstream. Confirmed associations would support evidence of the existence of a diet-microbiota axis as a prominent factor predisposing IA and CRC, and allow to identify microbial taxa involved in translocation in CRC and the potential microbial CRC targets in blood. This will allow to evaluate whether early diagnosis of CRC may be defined according to blood bacterial profiling.
The following questions can also be addressed: whether IA may cause some degree of bacterial translocation and/or the presence of bacterial taxa is associated to CRC cases; which mechanisms are behind the adenoma-carcinoma sequence in the CR development of carcinogenesis; which subgroups of patients are more likely to be associated with CR malignancy; which dietary guidelines can be specifically designed to prevent CRC. In addition, this project will give information regarding the overall entity of bacterial translocation occurring in (putative) healthy adult population.
Moreover, the project will offer a validated framework of procedures to standardize and elaborate microbiological data, and to design future epidemiological studies involving microbiota analysis. In fact, the project has been conceived in the context of a larger ambitious project that includes the study of the intestinal microbial ecosystem (through the fecal microbioma evaluation), and can serve as a pilot study to drag to this larger research.

Figure 1. *The food-microbial axis in colorectal carcinogenesis*



Bibliography

[1] Malvezzi M, Carioli G, Bertuccio P et al. European cancer mortality predictions for the year 2018 with focus on colorectal cancer. Annals of oncology: official journal of the European Society for Medical Oncology 2018; 29: 1016-1022.

[2] Ferlay J, Soerjomataram I, Dikshit R et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer 2015; 136: E359-386.

[3] Winawer SJ, Zauber AG, Ho MN et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. The New England journal of medicine 1993; 329: 1977-1981.

[4] Allison JE, Tekawa IS, Ransom LJ et al. A comparison of fecal occult-blood tests for colorectal-cancer screening. N Engl J Med 1996; 334: 155-159.

[5] Zeller G, Tap J, Voigt AY et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. Molecular systems biology 2014; 10: 766.

[6] Rossi M, Negri E, Parpinel M et al. Proanthocyanidins and the risk of colorectal cancer in Italy. Cancer Causes Control 2010; 21: 243-250.

[7] Rossi M, Negri E, Talamini R et al. Flavonoids and colorectal cancer in Italy. Cancer Epidemiol Biomarkers Prev 2006; 15: 1555-1558.

[8] Hans-Olov Adami DJH, Dimitrios Trichopoulos. Textbook of Cancer Epidemiology. Oxford University Press,2008

[9] Vogelstein B, Papadopoulos N, Velculescu VE et al. Cancer Genome Landscapes. Science 2013; 339: 1546-1558.

[10] Taverniti V, Stuknyte M, Minuzzo M et al. S-layer protein mediates the stimulatory effect of Lactobacillus helveticus MIMLh5 on innate immunity. Applied and environmental microbiology 2013; 79: 1221-1231.

[11] Guglielmetti S, Zanoni I, Balzaretti S et al. Murein lytic enzyme TgaA of Bifidobacterium bifidum MIMBb75 modulates dendritic cell maturation through its cysteine- and histidine-dependent amidohydrolase/peptidase (CHAP) amidase domain. Applied and environmental microbiology 2014; 80: 5170-5177.

[12] Ivanov II, Atarashi K, Manel N et al. Induction of intestinal Th17 cells by segmented filamentous bacteria. Cell 2009; 139: 485-498.

[13] Boleij A, Tjalsma H. The itinerary of Streptococcus gallolyticus infection in patients with colonic malignant disease. The Lancet Infectious diseases 2013; 13: 719-724.

[14] Rossi M, Bosetti C, Negri E et al. Flavonoids, Proanthocyanidins, and Cancer Risk: A Network of Case-Control Studies From Italy. Nutrition and Cancer-an International Journal 2010; 62: 871-877.

[15] La Vecchia C, Decarli A, Serafini M et al. Dietary total antioxidant capacity and colorectal cancer: a large case-control study in Italy. Int J Cancer 2013; 133: 1447-1451.

[16] Shivappa N, Hebert JR, Ferraroni M et al. Association between Dietary Inflammatory Index and Gastric Cancer Risk in an Italian Case-Control Study. Nutr Cancer 2016; 68: 1262-1268.

[17] Amar J, Lange C, Payros G et al. Blood microbiota dysbiosis is associated with the onset of cardiovascular events in a large general population: the D.E.S.I.R. study. PloS one 2013; 8: e54461.

THE IMPACT OF JOB CHANGE ON THE CARDIOVASCULAR HEALTH: EVIDENCE FROM THE BELGIAN IDEWE DATASET

Maniscalco L.[1], Godderis L.[2,3], Boon J.[2], Schouteden M.[2], Matranga D.[4]

[1] *Dipartimento di Biomedicina, Neuroscienze e Diagnostica Avanzata, Università degli Studi di Palermo, Italia;*

[2] *IDEWE, External Service for Prevention and Protection at Work, Interleuvenlaan 58, 3001 Heverlee, Belgium.*

[3] *Katholieke Universiteit Leuven, Centre for Environment and Health, 3000 Leuven, Belgium.*

[4] *Dipartimento di Promozione della Salute, Materno-Infantile, Medicina Interna e Specialistica di eccellenza "G. D'Alessandro", Università degli Studi di Palermo, Italia;*

Introduction

According to the World Health Organization (WHO), cardiovascular diseases (CVDs) are the main cause of death in the world. In 2016, the total number of world death for CVDs was 31%, of which 85% are due to heart attack and stroke [1], whereas in Europe was 45% [2] and in Belgium was 28% [3] Risk factors for CVDs are unhealthy diet, physical inactivity, tobacco use and harmful use of alcohol. Moreover, other determinants of CVDs include poverty, stress and hereditary factors. Several studies demonstrated that lifestyle and job-related behaviors as perceived work stress, low occupational physical activity, long working hours and shift work, are associated with cardiovascular diseases [4, 5, 6, 7]. Psychosocial risks have been discovered to be as one of the major risk factors for worker's occupational health and safety. Indeed, strong evidence indicate an association between work-related health complaints and exposure to psychosocial hazards. To our knowledge, the role of job change as psychosocial risk for CVD has been poorly explored, till now (8, 9).

Aim

The aim of this study was to investigate whether there exists an association between job change and CVD in employees working in the Healthcare sector, using information from the Belgian IDEWE dataset.

Methods

Study population. Data were obtained by the largest central repository collecting data on Belgian workers, the IDEWE data warehouse. IDEWE is a Belgian External Service for Prevention and Protection at Work. Its data warehouse contains data from the annual health check of Belgian workers executed by occupational nurses and physicians that recorded and encoded these data in an electronic format using international or national classification standard [10].
Variables. The final dataset contained 16709 observations and 24 variables. These observations correspond to 2387 individuals followed for 7 years. Indeed, for statistical reasons, we only retained subjects having exactly seven measurement-points. The covariates included in the analysis are: demographics (age, gender), risk factors (as diastolic blood pressure, BMI, smoking habits), occupational (as job change and skill levels) and risks related to job. The response variable "Medication use" is a binary variable with category "No user" indicating that a subject in a particular year did not use medication for cardiovascular problems, and category "User" otherwise. Using the ILO classification, skill levels are used as proxy of educational level. In detail, the first skill level corresponds to lower educational level (primary school) and it includes jobs of simple and routine physical or manual tasks. The second skill level involves jobs requiring little competence as reading and writing. Indeed, it includes jobs requiring the completion of the first stage of secondary education. Occupations in the third skill level require high level of literacy, numeracy and a well-

developed interpersonal communication skill so it corresponds to a higher educational level. The fourth skill level requires very high competencies and it involves the first or the master's degree and a possible advanced research qualification. The different classification of skill levels highlights the different educational levels and the several job tasks of mental or physical type. Some work-related risks were considered. These are: stress at work, burnout, three risks related to noise at work (lower exposure values of 80 decibel (dB), upper exposure values of 85 dB and exposure limit values of 87 dB) [11], physical load, manual handling of loads, manual lifting/holding/carrying, manual pulling and/ or pushing, manual repetitive tasks, static load, shift work with and without particular risks, night work with and without particular risks, psychosocial load.

Statistical methods. Continuous variables were summarized as mean (SD=standard deviation), median and range. Categorical variables were analyzed as counts and percentages. To assess the statistical significance of the difference, the Chi-squared test was used for categorical variables, and the ANOVA test was used for continuous variables. A logistic regression model for autocorrelated data with repeated measures was used to model the medication use for cardiovascular problem. Data were expressed as coefficients and standard errors and the category "no" was chosen as a reference for each categorical covariate. The backward selection procedure was used. Data were analyzed using R software (version 3.5.1). A p-value of less than 0.05 was considered statistically significant.

Results

A Logistic regression model for autocorrelated data was proposed to study the influence of time-dependent covariates on the marginal distribution of medication use for CVD problems in serially correlated binary data. Job change for people employed in the healthcare sector seemed not associated with medication use for CVD (RR=1.32, p<0.001). Analogously, smoking status and gender were not associated with medication use. Conversely, BMI and age were both important to explain variability of medication use for CVD and an increase of these two variables seemed to increase the probability to be a user. Skill levels were relevant for the response variable. The skill levels 3 or 4 played a protective role on cardiovascular medication use respect to worker with lower skill levels. Psychosocial load had an important role on the response variable (RR =1.20, p=0.004) (Table 1).

Table 1. *Logistic model for autocorrelated data: coefficients, standard errors and p-values.*

Coefficients	Value	Std. Error	p-value
Intercept	-5.821	0.416	**<0.0001**
BMI	0.030	0.007	**<0.0001**
Diastolic	-0.002	0.000	**0.012**
Smoking habits (Smoker vs No smoker)	-0.067	0.059	0.252
Sex (Female vs Male)	-0.236	0.152	0.120
Age	0.086	0.006	**<0.0001**
Job change (Yes vs No)	0.282	0.217	0.194
Skill level (2 vs 1)	-0.128	0.134	0.339
Skill level (3+4 vs 1)	-0.298	0.134	**0.026**
Psychosocial load (Yes vs No)	0.184	0.064	**0.004**
log (RR (Y_t; Y_{t-1}))	7.384	0.139	**<0.0001**

Conclusions

Results of our study showed the important role of skill levels and psychosocial load for CVD among Belgian healthcare workers, while change in work was not statistically significant. These results are in line with other literature. Indeed, psychosocial load was found as a significant determinant of medication use for cardiovascular diseases according to the Belgian Job Stress Project [8]. Higher skill level reduces the probability of medication use for cardiovascular diseases [12].

Bibliography

[1] WHO. The challenge of cardiovascular disease – quick statistics, 2016. http://www.euro.who.int/en/health-topics/noncommunicable-diseases/cardiovascular-diseases/data-and-statistics. Access 6 May 2019

[2] Townsend, N., Wilson, L., Bhatnagar, P. et al. Cardiovascular disease in Europe: epidemiological update 2016. European heart journal, 2016, 37(42), 3232-3245.

[3] WHO. Noncommunicable diseas es country profiles 2018. https://www.who.int/nmh/countries/bel-_en.pdf?ua=1

[4] Fishta, A., & Backe, E. M. Psychosocial stress at work and cardiovascular diseases: an overview of systematic reviews. International archives of occupational and environmental health, 2015, 88(8), 997-1014.

[5] Li, J., & Siegrist, J. Physical activity and risk of cardiovascular disease—a meta-analysis of prospective cohort studies. International journal of environmental research and public health, 2012, 9(2), 391-407.

[6] Kivimäki, M., Jokela, M., Nyberg, S. T. et al. Long working hours and risk of coronary heart disease and stroke: a systematic review and meta-analysis of published and unpublished data for 603 838 individuals. The Lancet, 2015, 386(10005), 1739-1746.

[7] Vyas, M. V., Garg, A. X., Iansavichus, A. V. et al. Shift work and vascular events: systematic review and meta-analysis. Bmj, 2012, 345, e4800.

[8] De Bacquer, D., Pelfrene, E., Clays et al. Perceived job stress and incidence of coronary events: 3-year follow-up of the Belgian Job Stress Project cohort. American Journal of Epidemiology, 2005, 161(5), 434-441.

[9] Haynes, S., Feinleib, M., & Kannel, W. B. The relationship of psychosocial factors to coronary heart disease in the Framingham Study. III. Eight-year incidence of coronary heart disease. American Journal of Epidemiology, 1980, 111(1), 37-58.

[10] Godderis, L., Mylle, G., Coene, M. et al. Data warehouse for detection of occupational diseases in OHS data. Occupational Medicine, 2015, 65(8), 651-658.

[11] European Agency for Safety, & Health at Work. Magazine: Magazine of the European Agency for Safety and Health at Work. Office for Official Publications of the European Communities, 1999, 1.

[12] Lemogne, C., Meneton, P., Wiernik et al. When blue-collars feel blue: depression and low occupational grade as synergistic predictors of incident cardiac events in middle-aged working individuals. Circulation: Cardiovascular Quality and Outcomes, 2017, 10(2), e002767.



Atti | Poster

DETECTING DIFFERENCES IN "CYTOTEC" SEARCH QUERIES USING THE FUNCTIONAL ANALYSIS OF VARIANCE APPROACH APPLIED TO GOOGLE TRENDS

Porreca Annamaria[1], Scozzari Francesca[1], Di Nicola Marta[2]

[1] *Department of Economic Studies, "G. d'Annunzio" University of Chieti-Pescara.*

[2] *Department of Medical, Oral and Biotechnologies Sciences, "G. d'Annunzio" University of Chieti- Pescara.*

Introduction

Functional data analysis (FDA) addresses problems in which the observations are described by functions rather than finite-dimensional vectors [1] [2]. This kind of data is very common in different fields, for example, environmental science, biology, medicine, meteorology, among many others [3]. The FDA approach assumes the existence of certain smooth functions (basic functions), which generate the observations. In real applications, functional data are often observed as a sequence of data points. The FDA approach provides several advantages for the analysis of data streams. First, it drastically reduces the dimensionality of data by converting discrete observations into functions; second, the use of additional functional tools, such as first and second derivatives, may sometimes reveal crucial information better than the original data; and third, FDA can tackle cases where data are not sampled at equally spaced time points [4].
In the era of Big Data, Internet search data can provide valuable information on disease models and population behaviour [5]. A tool that allows users to interact with Internet search data is Google Trends, a public service accessible for free online. Google Trends analyses a portion of Google's number of searches and provides data on geospatial and temporal models in search volumes for user-specified terms [6].
Since Google provides an enormous amount of data over time, traditional statistical techniques fail where a limited number of statistical units and a large number of variables or a large number of temporal observations are available. For this reason, the use of the (FDA has recently been proposed to analyse this type of data, which is comparable to the so-called datastreaming, which is characterized by data that flows continuously and in large quantities. The storage of these type of data, therefore, is very complicated, and thus the need to analyse them with sophisticated techniques arises.
In this study, we concentrate on Google trends regarding the worldwide search of drugs. In particular, we focus on the problem of abortion. Although abortion is illegal in some countries, many people try to terminate their pregnancy outside of a controlled clinical environment [7], [8]. Access to abortion is conditioned both by ethical-religious, legislative and socio-economic factors, but also by a lack of information and a low level of education. For these reasons, more and more people are consulting the Web to find information on self-abortion. Today, recent research suggests that it is easy to purchase on the WWW a combined treatment of mifepristone and misoprostol; these latter are usually used in early abortions in controlled clinical settings. In the practice of "non-surgical abortions", among costumers, the most common abortion drugs are the misoprostol, mainly in the form of the trade name Cytotec. Indeed, as recently stresses by international organizations (e.g. Pulitzer Center), Cytotec, "is used at lower doses to induce labor and slow post-partum bleeding, but has been widely used for self-induced abortions in Nigeria" [9]. Depending on the country and the legislation in force, these drugs are used alone or in combination based on the conditions for the supply, availability and use, with consequences on the effectiveness of the treatment [8].
In recent years, the use of drugs to practice abortion has expanded considerably, both in countries where the practice is legal and where it is not. In countries where access to abortion is limited, the use of medical abortion has become widespread and, Cytotec has acquired a certain "notoriety" for its abortion properties, whether administered orally or vaginally [8]. Even when health authorities do not officially recognize it or available through health services, women show interest in seeking information about self-abortion through informal networks or online pharmacies, where, the misoprostol is sold for use related to other conditions.

In Brazil, after its arrival on the market in 1986, the demand for Cytotec has undergone a drastic expansion [8]. The use of Cytotec has become widespread in other counties where applies restrictive legislation on abortion: Latin American countries [9] in Asian countries such as the Philippines [10] in Africa, e.g. Uganda, Ghana, and Nigeria [11], [12]. It has also gained ground in Spain and Italy by way of Latin American migrants who have spread the word of its abortifacient properties [13], [14]. In contexts where abortion is legal, such as Uruguay, Mexico City, France, Germany and others, medical abortion is replacing surgery-based methods and is now the main method used [15]. In some countries, such as France, medical abortion is under the strict control of health professionals, with two medical appointments necessary to obtain the drug, while in other countries, the practice is less regulated.

Objectives

The objective of this study is to propose an original application of FDA and functional analysis of variance (fANOVA) in the context of biostatistics. In particular, treating Google trends search queries on a specific drug over time as a functional datum, we aim understanding if the average trends of groups are statistically different in the considered time domain. To this purpose, the data regarding the search queries on "Cytotec" in 2018 are considered. As a grouping variable (factor), it is considered the legal possibility to abort in these countries.

Method

Google Trends is a tool that allows users to graph frequency searches for a term, a string of multiple terms or a sentence. The output data used to generate function graphs are resized to the average search traffic for the selected term and are also normalized on a relative, rather than absolute, basis. These graphs can also be manipulated to limit results to specific time intervals and geographic regions.
Functional data analysis addresses problems in which the observations are or (can be represented as) functions rather than vectors [1]. Thus, the first step in the FDA approach is to convert the original data into functions using a suitable smoothing technique.
If $y(x) \in L_2$, a basis function system is a set of K known functions $\theta_j(x)$ that are linearly independent. Hence, a function y(x) can be expressed by a linear combination of these basis functions as follows:

$$y(x) \approx \sum_{j=1}^{K} c_j \, \theta_j(x) \qquad (1)$$

Where c_j is the vector of coefficients defining the linear combination, and $\theta_j(x)$ is the vector of basis functions. The most common approach to converting a time series into a functional object is to use a linear combination of B-spline.
In this study, we can consider that starting from the original dataset, it is possible to obtain a function for each country, and thus a group of curves for each agglomeration of countries identified according to some categorical variable.
In this framework, to quantify the effects exerted at multiple levels on the functional observation by some factors, the fANOVA is proposed [1]. In the context of fANOVA, to understand if there are significant differences between groups of functional data, a pointwise F statistic is generally adopted [1], [16], [17].
This technique consists in testing for possible differences among mean curves under G different treatments over the whole functional domain. In this study, we propose this method to understand if two or more groups of countries are statistically distinguishable according to their Google queries. Therefore, we test the null hypothesis that the groups have the same functional mean against the alternative that there is some difference among them.

Hence, we assume that there is a single factor with V different groups (v = 1, 2, …, V) and $N = \sum_{i=1}^{V} n_v$ observations, with n_v observations within each group. Thus, the model for the i-*th* observation (i = 1, 2, …., N) in the v-*th* group can be expressed by:

$$y_{iv}(x) = \mu(x) + \gamma_v(x) + u_{iv}(x) \quad (2)$$

where $y_{iv}(x)$ is the functional response in the v-th group, $\mu(x)$ is the grand mean function, $\gamma_v(x)$ is the functional effect of being in a specific group, and $u_{iv}(x)$ is the residual function, i.e. the unexplained variation for the i-*th* observation within the v-*th* group.
The null hypothesis that the functional groups have the same functional mean, against the alternative hypothesis that there is some difference among them, can be expressed as follows:

H_0: $u_1(x) = u_2(x) = \cdots = u_V(x)$
H_1: $u_{v'}(x) \neq u_{v''}(x) \quad for\ at\ least\ one\ couple\ (v', v'')\ with\ v' \neq v''$

If SSY(x) represents the variance between groups in the functional context and SSE(x) is the functional within variance, we can express the pointwise F statistics as follows:

$$F(x) = \frac{SSY(x) - SSE(x)}{V-1} \Big/ \frac{SSE(x)}{N-V} \quad (3)$$

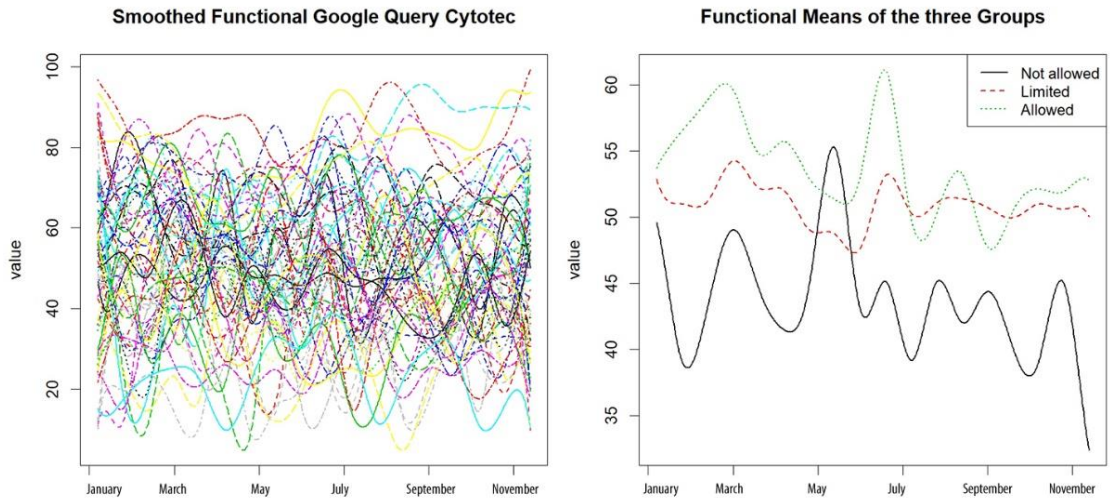
Equation 3 provides the observed F-statistic function obtained calculating the Fisher test statistic for each time instant. As in classical ANOVA, a high value of F means that the variance described by the model is higher than the non-explained variance. Differently from the classical ANOVA, in this framework, the F-Statistics is not a unique value but is a function of time.
In the classical ANOVA, the significance level is generally fixed for a single hypothesis; on the contrary, in the fANOVA framework, we need to test the hypothesis in a continuous domain. At this end, a viable solution is the application of the permutation test [1] that is the functional equivalent of the univariate F-test statistic. It allows us to estimate if there are any significant differences between the groups. The basic idea is to calculate Equation 3 for each piece of the domain. Nevertheless, to formally test the null, a unique test statistic is required. Adopting the maximum of the observed F-statistics function, the distribution of this one under the null hypothesis can be obtained by estimating the test-statistic several times making random permutations of the curves. In particular the idea is to assess the observed F-statistic function getting the "observed F" whose maximum is needed to estimate the p-value of the test. In the second move, we randomly re-label the curves with different curve numbers, without modifying the grouping structure; then for the set of re-labelled curves we calculate the F-statistics for each part of the domain and the maximum of these functions. This re-labeling procedure is replicated many times, and for each we estimate the pointwise F-statistic function and its maximum. In the third round, we find the pointwise 0.05 critical value of the null distribution at each point of the domain and calculate the 95th percentile of the F-statistic values corresponding to that point. The last action provides us the maximum 0.05 critical value of the null distribution, calculating the 95th percentile of the distribution obtained by the permutations of the second step. Taking the 95th percentile of the distribution we find the value of the critical threshold without using the statistical tables.

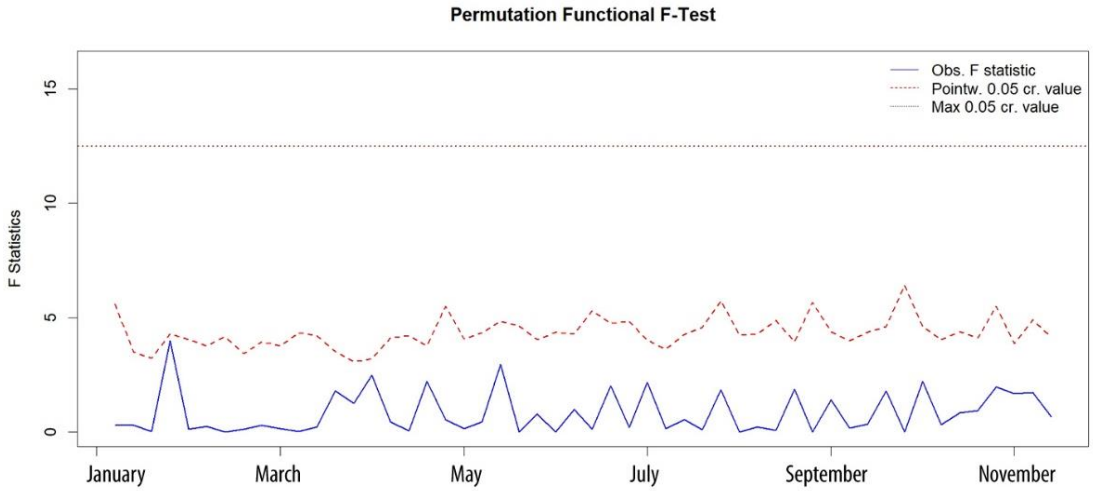
Results

Figure 1 is composed of three sub-figures: the smoothed Google trends of every country, the average functions of the three groups, and functional F test, respectively. In the first two graphs, the values of the

ordinate show how frequently a given search term is entered into Google's search engine relative to the site's total search volume over a given period of time.
The top-left sub-figure shows the smoothed version of the Google queries obtained using the B-spline approach over the period from the 07th of January 2018 to the 30th of December 2018; in total, we have collected a 52-instants time series and then used FDA to approximate the curves.
The top-right chart displays the functional means of the three groups. The countries selected for this study are those highlighted by Google Trends as the ones where the query of interest was more intense during the considered period.
We stress that the factor is the legal possibility to abort, and the treatments are "no possibility", "time-limited", and "always allowed". Group 1 ("no possibility") is composed of: Algeria, Saudi Arabia, Bolivia, Botswana, Brazil, Cameroon, Chile, Colombia, Congo - Kinshasa, Ivory Coast, Costa Rica, Ecuador, Egypt, El Salvador, United Arab Emirates, Philippines, Finland, Ghana, Jordan, Guatemala, Haiti, Honduras, Hong-Kong, Indonesia, Iran, Iraq, Ireland, Kenya, Lebanon, Malaysia, Morocco, Nicaragua, Nigeria, Pakistan, Panama, Paraguay, Peru, Poland, Dominican Republic, Senegal, Swaziland, Thailand, Trinidad and Tobago, Uganda, Venezuela, Zambia, and Zimbabwe. Group 2 ("time-limited") is composed of: Azerbaijan, Belgium, France, Italy, Mexico, Netherlands, Singapore, Spain, South Africa, Switzerland, and Turkey. Finally, Group 3 ("always allowed") is composed of: Germany, Greece, Norway, Sweden, and Taiwan [18].
The last figure shows the functional statistics for testing the null hypothesis that there are no significant differences between the mean group functions. The solid line represents the observed F statistic; the dashed curve indicates the 0.05 point-wise critical value computed with the permutation test, and the dotted line is the 0.05 maximum critical values. This test is based on the null distribution that has been constructed using 1,000 random permutations of the curve labels. Because the observed F statistic does not cross the 0.05 point-wise critical value and the 0.05 maximum critical values, we can conclude that there are no significant differences between the groups in terms of their mean functions. In other words, the legal possibility to abort has not a significant impact on the average trends of the three groups. This means that, regardless of the abortion laws that exist in countries, people are interested in a drug that has abortion as a contraindication and could be used for this purpose in countries where abortion is not legal.

Figure 1. *Smoothed functional Google queries, Functional means, fANOVA test.*





Conclusion

In this study, we presented an interesting application to understand how the legal possibility to abort differently influences the trend of Cytotec' searches in different groups of countries. Regardless of the application presented in this context, the use of functional analysis in medical studies, in recent years, has become very common. FDA, in addition to this possible application, offers numerous advantages such as the possibility of using additional tools such as derivatives, curvature, and dimensionality reduction techniques when there are datasets with few statistical units and many temporal observations, as in this one case. Moreover, recently many scholars have underlined the advantages of this approach with respect to the analysis of historical series because in the FDA we do not need any assumptions and can easily be used even with scattered data.
The idea of analyzing Google trend using FDA was presented for the first in [3] but a different purpose, e.g. clustering soccer players' Google queries using FDA. The possible future developments of this approach are diverse, e.g. forecasting, clustering to identify similar trends in countries or areas, classification, functional regression, and identification of anomalous behavior.
All these possible future developments and their applications in medical statistics can be very interesting for policy-makers and institutions that intend to monitor the phenomenon of drug research on the WWW, but also the pharmaceutical companies that are interested, for marketing purposes and sales, to understand the evolution of people's interest in abortive drugs, and even other medicines.

References

[1] Ramsay JO., Silverman BW., Functional Data Analysis. Springer Series in Statistics, Springer New York, 2007

[2] Ferraty F., Goia A., Vieu P., Nonparametric Functional Methods: New Tools for Chemometric Analysis. Statistical Methods for Biostatistics and Related Fields., 2007, 245–64

[3] Fortuna F., Maturo F., Di Battista T., Clustering functional data streams: Unsupervised classification of soccer top players based on google trends. Quality and Reliability Engineering International, 2018; 34(7): 1448-1460

[4] Cuevas A., A partial overview of the theory of statistics with functional data. Journal of Statistical Planning and Inference, 2014 Apr, 147:1–23

[5] Seifter A., et al., The utility of "Google Trends" for epidemiological research: Lyme disease as an example. Geospatial health, 2010 May 1, 4(2):135-137

[6] Nuti SV., et al., The Use of Google Trends in Health Care Research: A Systematic Review. PLoS ONE, 2014 Oct 22, 9(10): e109583

[7] Hess RF., Women's Stories of Abortion in Southern Gabon, Africa. Journal of Transcultural Nursing, 2007 Jan, 18(1):41–48

[8] Millard C., Brhlikova P., Pollock A., Social networks and health policy: The case of misoprostol and the WHO model essential medicine list. Social Science & Medicine, 2015, 132:190–6

[9] Pulitzer Center. An Abortion Pill in a Murky Market. Retrieved June 01, 2019, from https://pulitzercenter.org/reporting/abortion-pill-murky-market.

[10] Shah IH., Weinberger MB., Expanding access to medical abortion: Perspectives of women and providers in developing countries. International Journal of Gynecology & Obstetrics, 2012, 118 (26): S1–S3

[8] Fernandez MM., et al., Assessing the global availability of misoprostol. International Journal of Gynecology & Obstetrics, 2009 Mar 14, 105(2):180–186

[9] Barbosa RM., Arilha M., The Brazilian Experience with Cytotec. Studies in Family Planning, 1993 Jul, 24(4):236-240

[10] Gipson JD., Hirz AE., Avila JL., Perceptions and Practices of Illegal Abortion among Urban Young Adults in the Philippines: A Qualitative Study. Studies in Family Planning, 2011 Dec, 42(4):261–272

[11] Okonofua F., Shittu O., Shochet T., et al., Acceptability and feasibility of medical abortion with mifepristone and misoprostol in Nigeria. International Journal of Gynecology & Obstetrics, 2014 Jan 7, 125(1):49–52

[12] Atukunda EC., et al., Registration, procurement, distribution, and use of misoprostol in Uganda: an interview-based observational study, The Lancet Global Health, 2013 Oct, 382:10

[13] Coêlho H., et al., Misoprostol and illegal abortion in Fortaleza, Brazil. The Lancet, 1993 May, 341(8855):1261–3

[14] Zamberlin N., Romero M., Ramos S., Latin American women's experiences with medical abortion in settings where abortion is legally restricted. Reproductive Health, 2012 Dec, 9(1): 1-34

[15] Fiol V., et al., The role of medical abortion in the implementation of the law on voluntary termination of pregnancy in Uruguay. International Journal of Gynecology & Obstetrics, 2016 Aug, 134(S1): S12–S15

[16] Di Battista T., Fortuna F., Maturo F., Parametric functional analysis of variance for fish biodiversity assessment. Journal of Environmental Informatics, 2016, 28(2), 101-109

[17] Maturo F., Migliori S., Paolone F., Do Institutional or Foreign Shareholders Influence National Board Diversity? Assessing Board Diversity Through Functional Data Analysis. Studies in Systems, Decision and Control, 2017, 199–217

[18] Wikipedia. Legislazioni sull'aborto. Retrieved June 01, 2019, from https://it.wikipedia.org/wiki/Legislazioni_sull%27aborto

THE EFFECT OF DIFFERENT WEIGHTING SCORE ADJUSTMENT IN LOGISTIC REGRESSION MODELS FOR BINARY CORRELATED RESPONSES.

Marchioni Michele[1], Schips Luigi[2], Di Nicola Marta [1]

[1]*Department of Medical, Oral and Biotechnological Sciences, Laboratory of Biostatistics, University "G. D'Annunzio"; Chieti-Pescara*

[2] *Department of Medical, Oral and Biotechnological Sciences, Urology Unit, University "G. D'Annunzio"; Chieti-Pescara*

Introduction

In the recent years a growing body of literature relied on observational studies to estimate the effect of treatments as well as interventions on health outcomes [1]. However, differently from randomized clinical trials, observational studies are flawed by treatment selection bias since systematic differences in the treatment and control arms exist [1]. In order to reduce the confounding effect of these systematic differences several adjustment methods were introduced including techniques based on propensity score (PS) weighting [2]. The PS is defined as the probability of being assigned to a treatment conditional on baseline covariates. The PS is a balancing score where subjects in the treatment or control arm with the same PS have a similar distribution of the covariates [1]. Alongside with other PS-based methods, including matching, the inverse probability of treatment weighting (IPTW), can be used as adjustment in observational studies, to test the effect of a treatment on binary correlated responses [1]. The success of these methods led to develop and make available commercial packages able to generates weights and form equivalent groups in observational studies [3]. Within the R environment, among all the available packages the WeighIt package allow to generate weights using different algorithms [3]. While several direct or indirect comparison of different PS techniques are available in literature, evidences about comparison of PS and different weighting methods alongside with multivariable adjustment are sparse [4].

Objective

We aim to test differences of different weighting score methods on covariates balance, coefficient and standard error estimation as well as on the model goodness with a direct comparison of those weighting scores with the multivariable adjustment.

Methods

Statistical analyses. The within covariates systematic differences between two groups of patients were evaluated in terms of standardized mean difference (SMD). A SMD ≥0.1 was defined as meaningful, thus variables were considered balanced when SMD <0.1 [5]. Different balance score were estimated with several methods, all the estimation were performed considering the average treatment effect (ATE). First, PS weighting using generalized linear models was used to estimate the IPTW. Second, the covariate balance propensity score weighting (CBPS) was estimated. CBPS models the treatment assignment while optimizing the covariate balance. The estimation with the CBPS aim to improve the poor empirical performance of PS methods [6]. Third, entropy balance (EBAL) was used to estimate weights on a maximum entropy reweighting scheme that calibrates unit weights. Entropy balancing exactly adjusts inequalities in representation with respect of the possibly higher moments of covariates distributions [7]. Fourth, empirical balancing calibration weighting (EBCW) was used to estimated weight without resorting to direct estimation of PS or outcomes regression function [8]. Fifth, PS estimation was performed with a boosted regression (GBM), this method allow to overcome some limitations related to the accurate estimation of PS such as

large numbers of covariates, uncertain functional forms and their association with treatment selection [9]. Sixth, nonparametric covariate balance propensity score weighting (NPCBPS) estimated weights by maximizing the empirical likelihood of the data subject to balance constraints [10]. Seventh, optimization based weighting (OPTWEIGHT) finds the weights of minimum variance that adjust or balance the empirical distribution of the observed covariates [11]. To estimate the effect of exposure to the treatment on the outcome a weighted logistic regression model was fitted. Each weight was applied to different regression models with only the treatment as independent covariate. Coefficients and standard errors were computed. Standard errors and 95% confidence intervals (95%CI) were estimated using the sandwich estimator to compute the robust standard errors assuming a-priori that some of treated patients may have extreme probability compared to the controls [12]. Moreover, multivariable logistic regression adjusting for all the covariates used in the weights computation was also fitted as reference. Forest plots graphically depicted the OR and 95%CI. All the weighted models as well as the multivariable adjusted model were descriptively compared according to Akaike Information Criterion (AIC). The likelihood ratio test compared all the models to the standard multivariable adjusted model. All the analyses were performed to predict two outcomes defined according to the number of event per confounding variables in multivariable adjusted models. Specifically, we identified an outcome with a ratio >10 and one with a ratio ≤10 events/confounding variable [13]. All the test were two sided. All the analyses were performed using the R software environment for statistical computing and graphics (version 3.5.3, R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/).

Case study cohort. The study cohort consisted of a multi-institutional database collecting the surgical experience within 14 Italian centers. All the patients, affected by benign prostatic hyperplasia (BPH), underwent surgery due to moderate-severe bothersome lower tract urinary symptoms (LUTS) between 2011-2016. Based on the surgeon preference two laser surgeries were performed: standard photoselective vaporizazion of the prostate (sPVP) or anatomic photoselective vaporization of the prostate (aPVP), according to well established procedures [14]. Several pre- and intra-operative patients characteristics were collected, namely age (years), prostate volume (mL), prostatic specific antigen (PSA), maximum flow rate (Qmax), international prostate symptoms score (IPSS), American Society of Anesthesiologists (ASA) score, history of catheterization, anticoagulant/antiplatelet use, medical BPH treatment (alpha blocker, 5-alpha reductase inhibitor or no treatment), laser time (min), operative time (min), energy used (kJ) and laser density (kJ/g). Outcomes of interest were early (within 30 days) complications of any nature and delayed (over 90 days complications).

Results

Descriptive analyses. Overall, 522 patients were included. Of all, 325 underwent sPVP and 195 aPVP. Several covariates were unbalanced at the baseline, namely ASA score, Anticoagulant, BPH medical therapy, laser time, energy used and laser density (Figure 1a). After weighting the largest improvement was achieved with the EBAL, EBCW and OPTWEIGHT methods that showed a perfect balance for all the covariates (SMD for all the three methods approximately 0). Conversely, the largest grade of unbalance even after weighting was reached with the CBPS method. Indeed, afte r CBPS-weighting 9 covariates were still unbalanced with the largest SMD for the history of catheter (SMD=0.469) (Figure 1a). Of all, 216 patients (41.4%) experienced early complications. Conversely, only 78 patients had delayed complications (14.9%).

Outcome with <10 events per confounding variable

In all the models the OR for early complications exceeded 1.0 (Figure 1b) with the exception of CBPS-weighted model (OR: 1.0). The smallest standard error and consequently the narrowest 95%CI resulted from the GBM-weighted model. Interestingly, only multivariable models showed p-values <0.05. Moreover, the lowest AIC was showed in the multivariable adjusted model (Figure 1b). Similarly, the likelihood ratio

test showed significantly lower residual deviance for the multivariable adjusted model compared to all the others (all p<0.001).

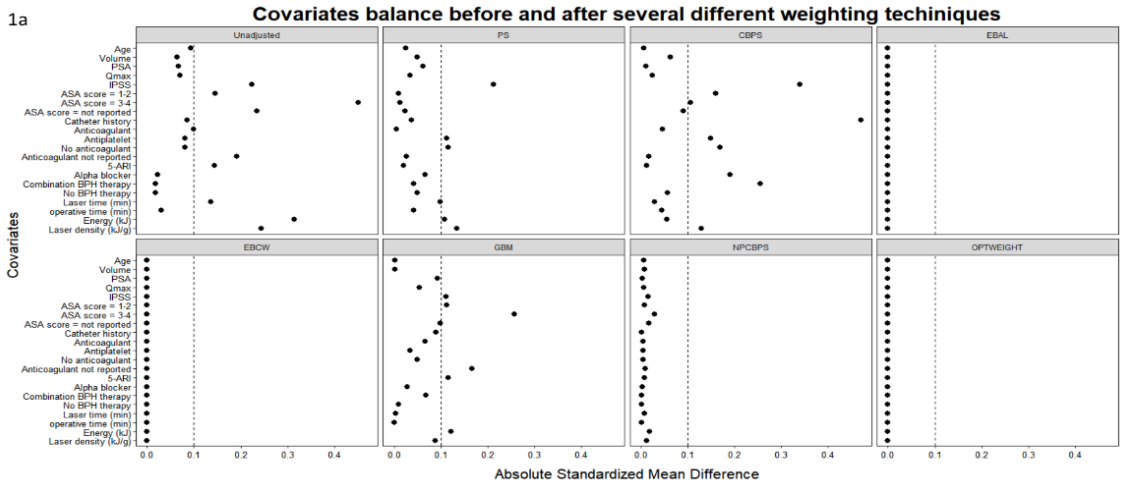
Outcome with ≤10 events per confounding variable

All the OR estimations exceeded 1.0 (Figure 1c). The smallest standard error and consequently the narrowest 95%CI resulted from the GBM-weighted model. In the CBPS- and NPCBPS-weighted models p-value was >0.05. The lowest AIC was shown in the PS-weighted model (Figure 1c). However, the likelihood ratio test showed similar residual deviance for all the models compared to the multivariable adjusted. The only exception were the NPCBPS- and OPTWEIGHT models that showed higher residual deviance compared to the multivariable adjusted one (all p <0.001).

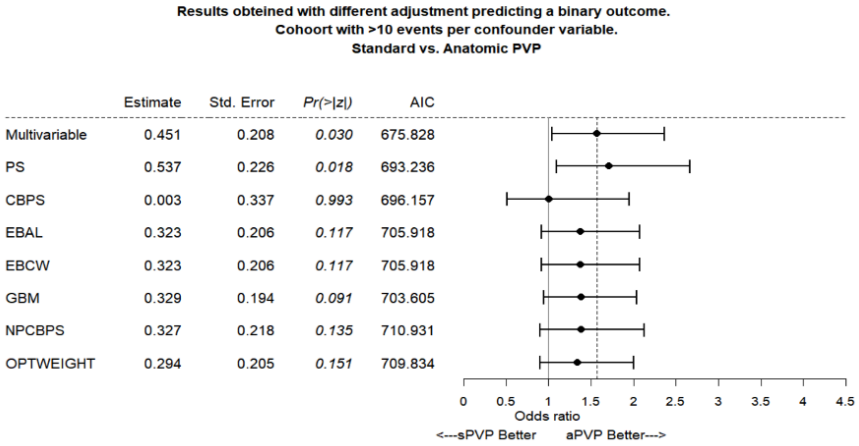
Conclusion

Several weighting methods are available and were implemented in ready-to-use packages. Some of them show the ability to produce an approximately perfect balance. However, such improvement seems to be not efficient to improve the overall quality of the models compared to the standard multivariable adjustment or the more widely used PS. Multivariable adjustment seems to be an effective method to control the selection bias in presence of higher number of events per confounding variables ratio. However, with smaller number of events per confounding variables, PS-weighting showed a better performance compared to the multivariable adjustment.

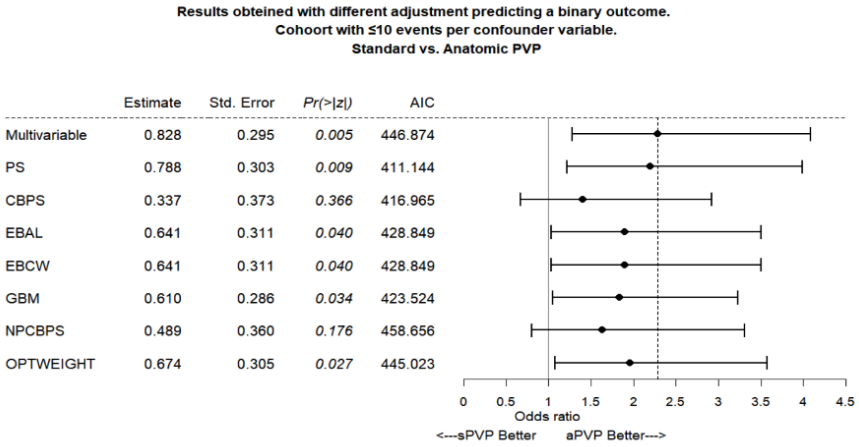
Figure 1. *Covariate balance before and after different weighting techniques (1a), forest plot depicting odds ratios and 95% confidence intervals obtained with different adjustment predicting a binary outcome in a cohort with > 10 events per confounder variable (1b) and ≤10 events per confounder variable (1c).*



1b



1c



References

[1] Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Statistics in Medicine. 2015;34:3661–79.

[2] Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res. 2011;46:399–424.

[3] Greifer N. WeightIt: Weighting for Covariate Balance in Observational Studies. R. 2019. https://cran.r-project.org/web/packages/WeightIt/WeightIt.pdf. Accessed 2 Jun 2019.

[4] Elze MC, Gregson J, Baber U, Williamson E, Sartori S, Mehran R, et al. Comparison of Propensity Score Methods and Covariate Adjustment. Journal of the American College of Cardiology. 2017;69:345–57.

[5] Zhang Z, Kim HJ, Lonjon G, Zhu Y. Balance diagnostics after propensity score matching. Ann Transl Med. 2019;7. doi:10.21037/atm.2018.12.10.

[6] Imai K, Ratkovic M. Covariate balancing propensity score. J R Stat Soc B. 2014;76:243–63.

[7] Hainmueller J. Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. Polit anal. 2012;20:25–46.

[8] Chan KCG, Yam SCP, Zhang Z. Globally efficient non-parametric inference of average treatment effects by empirical balancing calibration weighting. J R Stat Soc Series B Stat Methodol. 2016;78:673–700.

[9] McCaffrey DF, Ridgeway G, Morral AR. Propensity Score Estimation With Boosted Regression for Evaluating Causal Effects in Observational Studies. Psychological Methods. 2004;9:403–25.

[10] Fong C, Hazlett C, Imai K. Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements. Ann Appl Stat. 2018;12:156–77.

[11] Zubizarreta JR. Stable Weights that Balance Covariates for Estimation With Incomplete Outcome Data. Journal of the American Statistical Association. 2015;110:910–22.

[12] Scott A, Wild C. On the robustness of weighted methods for fitting models to case-control data: Methods for Fitting Models to Case-Control Data. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2002;64:207–19.

[13] Cepeda MS. Comparison of Logistic Regression versus Propensity Score When the Number of Events Is Low and There Are Multiple Confounders. American Journal of Epidemiology. 2003;158:280–7.

[14] Cindolo L, Ruggera L, Destefanis P, Dadone C, Ferrari G. Vaporize, anatomically vaporize or enucleate the prostate? The flexible use of the GreenLight laser. Int Urol Nephrol. 2017;49:405–11.

ELEVATED LEVELS OF EEF1A2 PROTEIN EXPRESSION IN TRIPLE NEGATIVE BREAST CANCER RELATE WITH POOR PROGNOSIS

Giudici Fabiola[1-2], Petracci Elisabetta[3], Nanni Oriana[3], Bottin Cristina[4], Pinamonti Maurizio[4], Barbati Giulia[1], Zanconati Fabrizio[4], Scaggiante Bruna[5]

[1] Biostatistics Unit-Department of Medical, Surgical and Health Sciences, University of Trieste, Cattinara Academic Hospitalc, Trieste, Italy

[2]Unit of Biostatistics, Epidemiology and Public Health Department of Cadiac, Thoracic, Vascular Sciences and Public Health University of Padova, Padova, Italy

[3] Unit of Biostatistics and Clinical Trials, Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori, Meldola, Italy

[3] Department of Medical, Surgical and Health Sciences, University of Trieste, Cattinara Academic Hospital, Trieste, Italy

[4] Department of Life Sciences, University of Trieste, Trieste, Italy

Background

Eukaryotic elongation factor 1 alpha 2 (eEF1A2) is a translation factor selectively expressed by heart, skeletal muscle, nervous system and some specialized cells. Its ectopic expression relates with tumorigenesis in several types of human cancer. No data are available about the role of eEF1A2 in Triple Negative Breast Cancers (TNBC). eEF1A2 expression is a covariate with a spike at zero (SAZ), i.e. a continuous covariate with a large amount of observations with value zero. Due to the nonstandard distribution of the covariate, common assumptions for regression techniques are not fulfilled. Moreover, observations with value zero might be substantially different with respect to the ones with a positive value, as they are e.g. lacking a certain characteristic completely. In this case, it might be reasonable to separate the effect estimate of two different types of observations: the zeros and the positive ones. This study investigated the relation between eEF1A2 protein levels and the prognosis of TNBC using two statistical approaches: 1) a standard Cox regression 2) a model based on the inclusion of a binary indicator combined with nonlinear regression modeling using fractional polynomials in situations with one covariate (eEF1A2) with spike at zero (SAZ).

Methods

A total of 84 TNBC diagnosed in the period 2002-2011 were included in the study. eEF1A2 protein level was measured in formalin-fixed paraffin-embedded tissues by immunohistochemistry in a semi-quantitative manner (sum of the percentage of positive cells x staining intensity) on a scale from 0 to 300. In the literature, this score is called H-score and is often adopted to evaluate immunohistochemistry results of tumor samples [1,2]. Regarding dichotomization, different criteria were used. On the one side, we evaluated the effect of the eEF1A2 using cut-point values based on quantiles of the H score distribution such as, median and tertiles. On the other, we used two approaches less tied to data and not based on the H score distribution. The first considered “negative”, patients whose sum of the percentages of not-stained-cells [0] and weakly-stained-cells [1+] was greater than the sum of the percentages of moderately-stained-cells [2+] and strongly-stained-cells [3+]. The second considered “negative”, patients who had no expression of eEF1A2 (100% expression at 0 or 1+). The association of clinical-pathological and molecular factors with the two time-to-event end-points, disease free survival (DFS) and breast cancer specific survival (BCCS) was analyzed separately using cause-specific Cox proportional hazards regression. Hazard ratios (HRs) and their 95%

CIs were reported. Statistically significant variables at 10% level at univariable analysis were selected as candidate prognostic factors for multivariable Cox analysis. Since, in our study eEF1A2 had a large proportion of observations with an H score equals to zero, an additional analysis was performed in presence of a spike at zero (SAZ) covariate, according to the method proposed by Royston and Lorenz, [3-5]. Their method, called "FP-spike", models the continuous non-zero observations of the SAZ covariate parametrically, using the fractional polynomials (FP), and adds to the Cox model a binary indicator variable (*v*), taking value equals to one when the SAZ covariate equals zero, and zero otherwise. The method consists of two stages: the first aims to identify the best FP function when also *v* is included in the model; the second to assess if *v* or the FP component can be eliminated without worsening the model fit. This second stage aims to understand if a kind of "dose-response" effect is present in the data and it is performed with the likelihood ratio test. As usual in Cox models, this method allows to consider the effect of other covariates into the model. Statistical analyses were conducted with R version 3.0.3 and STATA 14.2 (StataCorp, College Station, TX). All p-values (p) were two-tailed and, when not stated differently, differences were considered statistically significant when $p < 0.05$.

Results

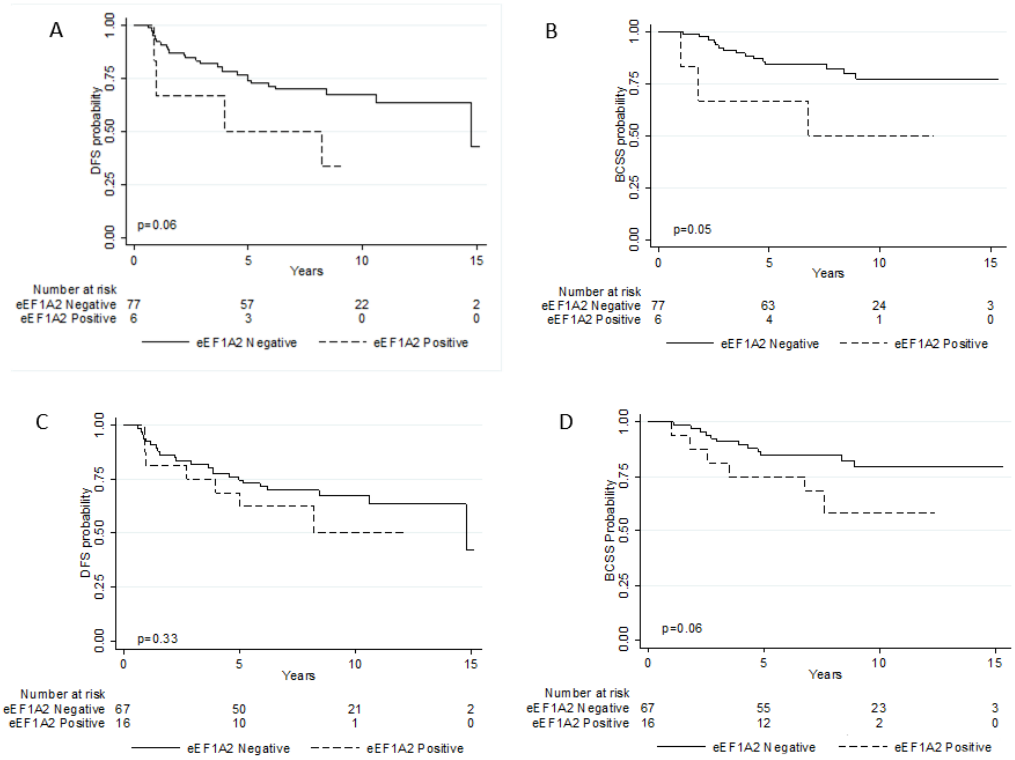
Elevated values of eEF1A2 were associated with older age at diagnosis (p=0.003), and androgen receptors positivity (p=0.002). At univariable Cox analysis, eEF1A2 levels weren't significantly associated with DFS and BCSS (p=0.11 and p=0.08, respectively) whereas adjusting for stage of disease, elevated levels of eEF1A2 protein resulted associated with poor prognosis (HR=1.05, 95% CI: 1.01-1.11, p=0.04 and HR=1.07, 95% CI: 1.01-1.14, p=0.03 for DFS and BCSS, respectively). When dichotomous versions of eEF1A2 were considered, similar results were obtained.
Figure 1 shows the Kaplan-Meier curves for DFS and BCSS for the two dichotomous variables, respectively. In our study, eEF1A2 has a large proportion of observations with an H score of zero (43%): for DFS, a simple linear function (FP1) to model the relationship between the non-zero values of the H score and the hazard was selected by the FP-spike procedure. Moreover, when the indicator variable, *v*, and FP1 component were tested for removal, both resulted important for the model fit (P-values=0.032 and P-value=0.006, respectively). As already shown, adjusting for stage, we observed that increasing values of eEF1A2 were associated with poor prognosis. The FP spike model showed that also women with an H score of zero, have a risk of a recurrence higher compared to women with values different from zero (P-value=0.039). For BCSS, by the FP-spike analysis, the addition of the binary component did not bring any substantial contribution to the model fit. In particular, the best functional relationship between eEF1A2 and the hazard of was described by a linear model confirming that an increase of H-score values was associated with an increase in worse prognosis.

Conclusions

Our data showed a negative prognostic role of eEF1A2 protein in TNBC confirmed by using different strategies. The distribution of the biomarker reported in terms of H score is certainly asymmetric with a large amount of observations with zero-values. Before deciding how to use the biomarker, that is as a "continuous" variable or as a binary/categorical one, different approaches have been considered. It has been known that categorization of continuous variables has pros and cons. The main advantage is certainly the easy interpretation of the study results and the possibility to draw Kaplan-Meier curves. The main disadvantage is that using cut-off values, much information is lost and thus the statistical power to detect a relation between the variable of interest and patient outcome is reduced. A categorical analysis is not always satisfactory, since it implies an arbitrary definition of cut points, a jump in the risk estimate at these cut points which is biologically implausible, and an arbitrary choice of the number of categories. In addition, categorization may introduce residual confounding and results in loss of statistical power.

In our case, the biomarker, when dichotomize-based on the median, was not associated with either DFS and BCSS. Similar results were obtained using tertiles. As mentioned above, other dichotomization criteria were considered, and we found that high values of eEF1A2 were associated with worse prognosis.
Although the use of the biomarker as continuous variable satisfied all the assumptions underlying a Cox regression model, such simplistic approach may not still be the most suitable in the presence of SAZ variables. In addition to these approaches, we dealt the problem of having a “continuous” covariate with a spike at zero (SAZ) using a statistical model which took into account this variable's particular characteristic This model confirms that high values of eEF1A2 expression relate with poor prognosis but it adds an information that would not be evident using a standard model: for DFS the FP-spike model gave a statistically significant adverse effect in women with an H score of zero compared to those with a positive one. The interpretation of this observation deserves further investigations, maybe by evaluating eEF1A1 isoform levels. In fact, it has been demonstrated a reciprocal influence of the two proteins at expression levels in many cancers and the possibility that the overexpression of eEF1A1 protein sustains cancer progression [6,7]
Further studies using a larger independent data set of TNBCs are necessary to confirm the prognostic value of eEF1A2 protein before an implementation in clinical practice. Certainly, the measurement of protein level is very interesting because immunohistochemistry is the main technique available in most pathology labs to evaluate a marker. With this study we encourage the use of appropriate statistical techniques suitable for “particular” covariates, such in this case, a SAZ variables. Even if FP-spike models are not simple to interpret they can contribute to better interpret a role of a biomarker in clinical research.

Figure 1. *Kaplan Meier curves for DFS and BCSS in TNBC cohort*



References

[1] Hirsch FR., Varella-Garcia M., Bunn PA Jr. et al. Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. J Clin Oncol., 2003 Oct 15;21(20):3798-807

[2] John T., Liu G., Tsao MS. Overview of molecular testing in non-small-cell lung cancer: mutational analysis, gene copy number, protein expression and other biomarkers of EGFR for the prediction of response to tyrosine kinase inhibitors.Oncogene, 2009 Aug;28 Suppl 1:S14-23.

[3] Royston P., Sauerbrei W., Becher H. Modelling continuous exposures with a 'spike' at zero: a new procedure based on fractional polynomials. Stat Med. 2010 May 20;29(11):1219-27.

[4] Lorenz E., Jenkner C., Sauerbrei W. et al. Modeling Variables With a Spike at Zero: Examples and Practical Recommendations. Am J Epidemiol. 2017 Apr 15;185(8):650-660.

[5] Lorenz E., Jenkner C., Sauerbrei W. et al. Modeling exposures with a spike at zero:simulation study and pratical application to survival data. Biostatistics & Epidemiology. 2019 Feb 20; 3(1):23-37.

[6] Hassan MK., Kumar D., Naik M. et al.The expression profile and prognostic significance of eukaryotic translation elongation factors in different cancers. PLoS One. 2018 Jan 17;13(1):e0191377

[7] Lin CY., Beattie A., Baradaran B. et al. Contradictory mRNA and protein misexpression of EEF1A1 in ductal breast carcinoma due to cell cycle regulation and cellular stress. Sci Rep. 2018 Sep 17;8(1):13904

STIMA DEI TREND TEMPORALI UTILIZZANDO DATABASE AMMINISTRATIVI: CONSIDERAZIONI METODOLOGICHE ED UN'APPLICAZIONE AI DATI INAIL RELATIVI ALLO SCREENING PER USO DI SOSTANZE STUPEFACENTI IN PROVINCIA DI COMO E VARESE

Raimondi Jessica[1], Cavicchiolo Marco[2], Dashi Oriana[2], Calderini Duccio[3], Aiani Maria Rita[3], Ottaviani Carlo[4], Tosetto Claudio[5], Olandese Raffaela[6], Borchini Rossana[2,7], Ferrario Marco Mario [1,2,7], Veronesi Giovanni[1]

[1] *Centro Ricerche EPIMED Università degli Studi dell'Insubria, Varese;*

[2] *Scuola di specializzazione in Medicina del Lavoro, Università degli Studi dell'Insubria, Varese;*

[3] *PSAL ATS Insubria, Varese-Como;*

[4] *Sovrintendente Sanitario INAIL Regione Lombardia;*

[5] *Ser.D – ASST Sette Laghi, Varese;*

[6] *UOC Dipendenze – ASST Lariana, Como;*

[7] *S.C. Medicina del Lavoro, Preventiva e Tossicologia – ASST Sette Laghi, Varese.*

Introduzione

Il decreto legislativo 81/08 stabilisce l'obbligo annuale da parte del medico competente (MC) di trasmettere tramite piattaforma INAIL schede informative riguardanti i dati aggregati sanitari e di rischio dei lavoratori sottoposti a sorveglianza sanitaria [1]. Tale documento prende il nome di allegato 3B (All-3B). Lo stesso decreto legislativo fa riferimento esplicito alla necessità di effettuare accertamenti periodici, finalizzati ad escludere l'assunzione di sostanze stupefacenti e psicotrope, nei lavoratori addetti a mansioni che possono comportare rischi per la sicurezza propria e di terzi [2-3]. Gli accertamenti per assunzione di sostanze stupefacenti si compongono di due fasi: una prima fase ad opera del MC, il quale effettua i test di screening; i positivi confermati vengono inviati al Ser.D che si occupa di effettuare accertamenti di secondo livello. Nelle Province di Como e Varese, nel periodo 2013-2016 è stato osservato un trend negativo nel numero di lavoratori inviati al Ser.D. L'analisi dei dati provenienti dall'All-3B e relativi al medesimo periodo può fornire importanti indicazioni sui determinanti di tale andamento.

Obiettivi

Lo scopo del presente lavoro è stato quello di stimare i trend temporali nella prevalenza dei lavoratori *i)*. sottoposti a *screening* da parte del MC e *ii)* positivi al test di conferma, utilizzando i dati dell'All-3B relativi alle Unità Produttive (UP) delle Province di Como e Varese, per gli anni 2013-2016.

Metodi

Il database dell'All-3B si presenta frammentato in 12 file per ogni anno e per ogni provincia. Il primo passo è stato quindi quello di ricostruire la relazione tra i diversi database attraverso uno schema Entità-Relazione. Una volta ricreato il database con le variabili necessarie ai fini dell'analisi, si è stimato il tasso di partecipazione delle UP nei diversi anni, come rapporto tra il numero di UP che presentano All-3B, sul totale di UP presenti sul territorio. Il denominatore è stato stimato utilizzando la banca dati INAIL dei premi assicurativi pagati dalle singole aziende (PAT), utilizzando come chiave di link il codice fiscale e la partita IVA. Per la stima dei trend temporali, i dati disponibili aggregati per j-esima UP (j=1,…,J) e per i-esimo anno

(i=1,…,4) sono: il numero medio di lavoratori presenti nell'anno ($N_{i,j}$), il numero di lavoratori screenati ($S_{i,j}$) ed il numero di lavoratori risultati positivi ai test ($P_{i,j}$). Per la stima dei trend abbiamo utilizzato modelli log-lineari a misure ripetute, ovvero: $\log(S_{i,j}) = \alpha + \beta t_i + \log(N_{i,j})$, e $\log(P_{i,j}) = \alpha + \beta t_i + \log(S_{i,j})$ per la prevalenza di lavoratori *screenati* e positivi, rispettivamente. L' *Average Annual Percent Change* (AAPC%) è approssimato da 100*$\hat{\beta}$ [4]. Per la mancanza di identificativo univoco per la stessa UP negli anni, l'analisi è stata ristretta alle sole aziende con una sola UP, ipotizzando che fosse la stessa nel tempo. Per modellare la correlazione negli anni tra le diverse UP, abbiamo considerato una matrice di varianza-covarianza *compound symmetry*. L'AAPC% è stato stimato ipotizzando per $S_{i,j}$ e $P_{i,j}$ sia una distribuzione di Poisson che una distribuzione Binomiale Negativa, confrontando quindi i risultati sia in termini di stima puntuale che di standard error. Infine, come analisi di sensitività, abbiamo ristretto la stima alle UP presenti in tutti e quattro gli anni indagati. Le analisi sono state eseguite con software SAS versione 9.4.

Risultati

Dal confronto effettuato con la fonte dati INAIL-PAT è risultata una partecipazione alla compilazione dell'All-3B per l'anno 2013 del 40% circa, senza distinzione tra aziende delle province di Como e Varese. Tale partecipazione aumenta negli anni, raggiungendo nel 2016 il 45%. Il 50% delle aziende in All-3B aveva dato mancante per il settore merceologico (codice ATECO); grazie a record linkage con la fonte dati INAIL-PAT, questa percentuale è scesa al 12%. Ad inizio periodo, tre settori merceologici erano responsabili del 75% e dell'80% dei lavoratori *screenati* e positivi, rispettivamente: la metalmeccanica, le costruzioni, ed il trasporto e magazzinaggio. Per le analisi dei trend, abbiamo considerato le 7597 UP di questi tre settori, per un totale di 23040 All-3B. Di questi, il 55% non aveva alcun lavoratore sottoposto a screening per droghe. Con il modello di Poisson, la stima di AAPC% nella prevalenza di lavoratori *screenati* è risultata pari a 3.1%±0.7% (IC95% 1.8%-4.4%). L'analoga stima con la distribuzione Binomiale Negativa è 2.4%±0.5% (IC95% 1.5%-3.4%). Anche nelle stime entro settore merceologico, l'AAPC% con Poisson sovrastimava quello ottenuto con la Binomiale Negativa; la sovrastima tendeva ad aumentare per incremento della percentuale di osservazioni con zero lavoratori *screenati*. Tale andamento si confermava restringendo le analisi ai dati completi (UP presenti per tutti e 4 gli anni considerati). La percentuale di lavoratori confermati positivi sul totale dei lavoratori sottoposti a screening (P/S) è inferiore allo 0.5% in tutti i settori considerati. In due settori su tre è risultato un aumento metalmeccanica (+13.3%); trasporto e magazzinaggio (+17.5%), sebbene non in maniera statisticamente significativa.

Conclusioni

Questa analisi di dati provenienti dalla sorveglianza sanitaria dei lavoratori ha confermato che l'utilizzo secondario di banche dati amministrative richiede opportuni accorgimenti metodologici per la stima della partecipazione, l'integrazione dei dati mancanti, e la scelta dell'opportuno modello probabilistico di riferimento. Inoltre vi sono alcune importanti limitazioni, quali l'impossibilità di seguire nel tempo la stessa UP, che obbligano ad una selezione dei dati. L'analisi effettuata ha mostrato un incremento statisticamente significativo nel numero di lavoratori *screenati*, ed un trend in sostanziale aumento anche del numero di lavoratori positivi. Questi risultati, in contro-tendenza con la diminuzione del numero di lavoratori che approdano al Ser.D per test di secondo livello, richiamano la necessità di studi costo-efficacia dell'attuale modalità di *screening* per abuso di sostanze stupefacenti in ambito lavorativo.

Bibliografia

[1] Decreto Ministeriale 9 luglio 2012 Contenuti e modalità di trasmissione delle informazioni relative ai dati aggregati sanitari e di rischio dei lavoratori, ai sensi dell'articolo 40 del decreto legislativo 81/2008 in materia di tutela della salute e della sicurezza nei luoghi di lavoro. G.U. Serie Generale n.173 del 26-07-2012.

[2] Atti del seminario "Alcol, droghe e lavoro" Le nuove regole e i problemi applicativi, Bologna 25 Maggio 2009.

[3] Provvedimento n. 99/CU del 30 ottobre 2007. Intesa ai sensi dell'articolo 8 comma 6, della legge 5 giugno 2003, n. 131, in materia di accertamento di assenza di tossicodipendenza. Gazzetta Ufficiale della Repubblica Italiana 2007, 266.

[4] Kleinbaum et al., Applied Regression Analysis and Other Multivariable Methods, 3rd Edition, Duxbury Press, Pacific Grove, CA (USA) 1998

ITALY AND SWEDEN: A DESCRIPTION ON COGNITIVE TREND AMONG OLDER ADULTS

Ferraro Ottavia Eleonora[1], Rizzuto Debora [2], Grande Giulia [2], Guaita Antonio[3], Villani Simona[1]

[1] *Unit of Biostatistics and Clinical Epidemiology, Department of Public Health, Experimental and Forensic Medicine University of Pavia*

[2] *Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, Stockholm, Sweden*

[3] *Golgi Cenci Foundation, Corso San Martino 10, 20081 Abbiategrasso*

Introduction

Along with the global aging of the population [1], people affected by age-related disorders- such as dementia- will increase dramatically with burdensome consequences at individual and societal levels.
Although different for several cultural, social and economic aspects, Italy and Sweden had, respectively on average, a similar expectancy of life of 83 years[2] and 82.5 years[3] in 2018.
One of the greatest challenges is to prevent the increase of non-communicable diseases, such as dementia, by understanding characteristics and risk factors related to cognitive function in the oldest people.

Aim

In this study we aim at describing trajectories of cognitive decline for older adults' part of two different cohorts in Italy and in Sweden using the group based trajectory models.

Methods

Data were retrieved from two population-based studies: 1) the *INVEcchiamentoCErebrale in Abbiategrasso* (InveCe.Ab) study, settled in the Abbiategrasso, Milan, Italy, in 2010 2) the *Swedish National Study on Aging and Care in Kungsholmen* (SNAC-K) settled in Stockholm, Sweden in 2001.

Outcomes. Cognitive function was measured using the Mini Mental State Examination (MMSE).

Statistical Analysis. Group based trajectory models (GBTM) were used to explore cognitive decline separately by country [4].
This methodology assumes that the population is composed of distinct groups, each with a different trend.
The adequate number of final groups and the best fit for the trajectories were chosen using the higher BIC (Bayesian information criterion) that takes into consideration the total number of assessments used in model estimation across subjects and follow-up times. Each subject was classified in a specific group on the basis of the maximum estimated probability.
The two cohorts were described considering participants with similar follow-up time and similar ages.
People with a diagnosis of dementia at baseline assessment were excluded from the analyses.

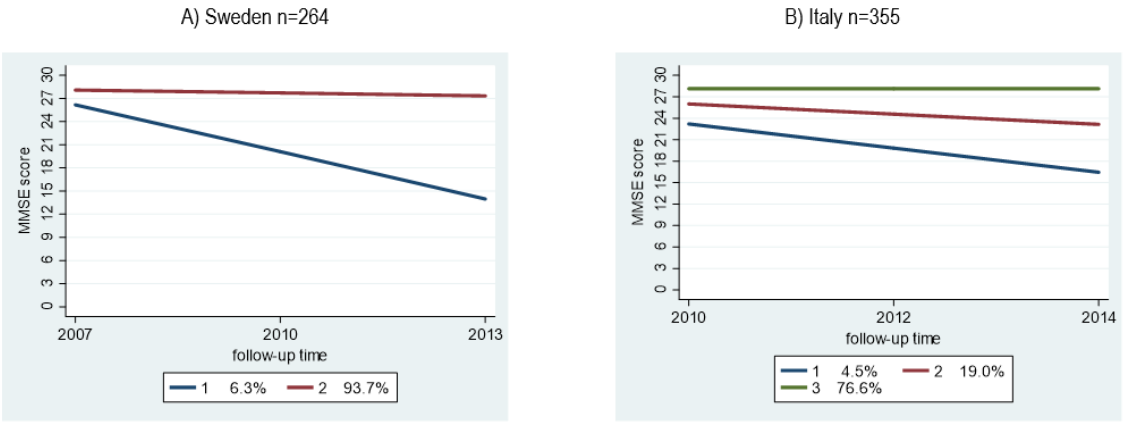
Results

This study included individuals from the two cohorts with similar characteristics as what concern age and follow-up time. Mean age was 76.81 years (sd 0.87 years), 61% of women and 6 years for the follow up (2007-2010-2013) among Swedish. Instead in the Italian cohort the mean age was 74.51 years (sd 0.50 years), 53% of women and 4 years of follow-up (2010-2012-2014).

Sweden. The 264 Swedish people included in this analysis were grouped, as the result of best model from BIC, in two different MMSE trajectories (Figure 1, A): the red line identifies the group of people with the highest and stable MMSE score (n=248). Among them, (n=24) 10% developed dementia over the follow-up time; the blue line (n=10) identifies individuals with the worst MMSE score, with an average annual decline of more than 2 points of MMSE. The latter group is entirely composed by individuals who developed dementia over the follow-up (100%).

Italy. Concerning the Italian cohort, the best model, with the best BIC, was composed by 3 groups with 355 subjects. The green line represents subjects with a stable trend (n=277) above 27 points on MMSE score. In this group less than 1.5% had dementia. The red line is for the middle group where there was a slight decrease in MMSE score less than 1 point in average annual decline in 2014 and 24% were people with dementia. The last group (n=15), in blue, was composed by 53% of subjects with a diagnosis of dementia and an average annual decline of 2 points in MMSE (Figure 1, B).

Figure 1. *Description of cognitive function in Sweden and Italy*



Conclusion

Accordingly to the results of our study the majority of Italian and Swedish older adults had good and stable cognitive trajectories, around 27 MMSE points. More precisely 93.7% in Sweden and 76.6% in Italy belonged to this group. In the groups with the steepest decline on cognitive function, the proportion of subjects with dementia, captured by GBTM, was between 53% and 100% respectively for Italy and Sweden. The GBTM, using only information on cognitive function during time, seems to consistently capture and group together a high proportion of people with dementia in the group with the worst cognitive trajectory. To the best of our knowledge, this is the first time that GBTM has been used to describe cognitive trajectories in two different cohorts. Further studies need to be done to better characterize people belonging to different cognitive trajectories.

Bibliography

[1] WHO, “Global Health and Aging,” 2011.

[2] ISTAT, “Indicatori demografici,” 2019.

[3] P. and E. W. statistics unit Statistics Sweden, “Life expectancy 1751–2018.” [Online]. Available: https://www.scb.se/en/finding-statistics/statistics-by-subject-area/population/population-composition/population-statistics/pong/tables-and-graphs/yearly-statistics--the-whole-country/life-expectancy/. [Accessed: 20-Jun-2019].

[4] D. S. Nagin, *Group-based Modeling Development*. 2005.

A BAYESIAN SAMPLE SIZE ESTIMATION PROCEDURE BASED ON A B-SPLINES SEMIPARAMETRIC ELICITATION METHOD

Danila Azzolina[1], Paola Berchialla[2], Dario Gregori[3], Ileana Baldi[3]

[1]*Department of Translational Medicine, Università del Piemonte Orientale, Novara, Italy*

[2] *Department of Clinical and Biological Sciences, University of Torino, Italy*

[3] *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac Thoracic Vascular Sciences and Public Health, University of Padua, Italy*

Introduction

Sample size estimation is a fundamental element of a clinical trial, and a binomial experiment is the most common situation faced in clinical trial design. The Bayesian sample size estimation is an alternative method to a frequentist procedure, especially for trials conducted on a limited number of patients.
In a Bayesian setting, it is possible to consider the uncertainty in data prediction because the binomial parameter under investigation is not a point estimate but a posterior distribution obtained combining data with available knowledge, translated into a prior distribution.

Objective

When no objective data are available to define a prior, it may be necessary to perform an expert elicitation. This work proposes a generalized version of the Average Length Criterion (ALC) Average Coverage Criterion (ACC), and Worst Outcome Criterion (WOC) for the binomial sample size estimation[1]. The original procedure was defined in a parametric framework[2].

Methods

The proposed method is a flexible approach allowing for the conventional parametric prior but also the semiparametric distributions based on B-splines[3].
The method has been proposed and implemented in a sample size estimation plan using a Generalized ACC, ALC and WOC estimation procedure (GACC, GALC, GWOC) considering different kinds of priors defined in a parametric framework (ACC, ALC, WOC) or using semiparametric B-Spline priors (BSACC, BALC, BSWOC).

Results

The semiparametric sample size is a compromise between the same estimates provided in a beta informative and uninformative setting ensuring also not too different results.
Performing a comparison across methods, GWOC is the more conservative solution. Instead, GALC gives a smaller sample size.

Conclusion

The generalized method leads the definition of flexible Bayesian design using different prior. The approach is useful for a clinical trial conducted on small sample sizes and when no objective evidence is available and may be indispensable to summarize an expert opinion.

References

[1] Adcock C. A Bayesian approach to calculating sample sizes. The statistician. 1988;433–9.

[2] Joseph L, Wolfson DB, Du Berger R. Sample size calculations for binomial proportions via highest posterior density intervals. The Statistician. 1995;143–54.

[3] Bornkamp B, Ickstadt K. A note on B-splines for semiparametric elicitation. The American Statistician. 2009;63(4):373–7.

I MODELLI DI DIFFUSIONE NELLA PREVISIONE E CONTROLLO DI PATOLOGIE INFETTIVE

Di Traglia M., Vestri A.R.

Dipartimento di Sanità Pubblica e malattie Infettive, Università degli Studi di Roma "Sapienza"

Introduzione

La diffusione di malattie infettive nelle popolazioni attuali è facilitata dalla mobilità internazionale degli individui provenienti da qualsiasi luogo del mondo e dalle migrazioni dai luoghi più poveri e disagiati verso quelli più avanzati. Considerando che la probabilità di essere portatori di patologie infettive è maggiore nei paesi poveri rispetto a quelli ricchi e che, questi ultimi, esercitano maggiore attrazione proprio verso i primi, il problema dell'insorgenza e diffusione di malattie infettive è diventato estremamente attuale. Questo meccanismo di attrazione verso i paesi meno sviluppati, da parte dei paesi più avanzati risponde ad un principio naturale noto come "seconda legge della termodinamica". Tale principio afferma che qualunque sistema evolve in modo naturale verso un aumento dell'entropia, ossia, verso l'equidistribuzione. Dal punto di vista della meccanica statistica significa che le molecole di un gas qualsiasi, creano flussi che vanno da porzioni dello spazio dove la pressione è maggiore verso porzioni dello spazio dove essa è minore. Se applichiamo tali concetti al movimento degli esseri umani, per analogia, si vede che sono le stesse leggi che governano i movimenti internazionale di individui comprese le migrazioni. I modelli matematici, introdotti dai fisici, per studiare tale fenomeno sono i modelli di diffusione spazio-tempo. Il modello che meglio sembra rispondere ai requisiti di realtà per poter essere applicato alla diffusione di malattie infettive è quello di R. Feynman. Tale modello tiene conto delle differenze di potenziale tra porzione differenti dello spazio (intese come pressioni differenti tra differenti regioni delle terra, dovute a guerre, povertà, malattie ecc..) attraverso opportune funzioni matematiche delle quali vengono stimati i parametri attraverso i dati statistici disponibili sulle migrazioni. Tali differenze provocano flussi migratori più o meno intensi dai pesi più poveri verso quelli più ricchi. L'applicazione del modello di Feynman allo studio della diffusione di malattie infettive, allo stato attuale, non esiste in letteratura ed il tentativo di questa proposta metodologica è quello di utilizzarlo per sviluppare uno strumento che permetta la previsione e controllo della diffusione spazio-tempo di patologie infettive nelle popolazioni. Il modello di Feyneman è stato applicato ad un problema di diffusione di specie vegetali come reazione ai cambiamenti climatici (Di Traglia et. al. 2011) stimando i parametri attraverso dati satellitari. In questo lavoro, il modello di Feynman viene integrato con quello di Eherenfest per tener conto delle fluttuazioni casuali negli spostamenti da un luogo all'altro degli individui. I modelli matematici nello studio della diffusione di epidemie nelle popolazioni umane sono stati introdotti per la prima volta da D.Bernoulli (1700-1782) per l'epidemia di vaiolo del 1760. Attraverso il modello matematico, Bernoulli dimostrava la necessità della vaccinazione preventiva come argine per la diffusione delle malattie infettive nelle popolazioni. Successivamente sono stati introdotti modelli che tengono conto delle differenti tipologie di infezioni: immunizzanti e non immunizzanti. I primi modelli dividono la popolazione in Suscettibili-Infetti-Rimossi (SIR) mentre i secondi dividono la popolazione in Suscettibili-Infetti-Suscettibili (SIS). Altri approcci utilizzano la tecnica delle reti di comunicazione per calcolare la probabilità di contatto tra infetti e suscettibili (Mata 2015) [1]. I modelli più noti in letteratura, risalenti agli inizi del '900, sono dovuti a Kermack – McKendrick ed a Lotka-Volterra. Tali modelli, con le successive modifiche ed adattamenti, possono classificarsi come "modelli di Campo Medio". I modelli di campo medio sono modelli deterministici i quali descrivono la dinamica delle popolazioni (SIR o SIS) come media (rispetto allo spazio) di variabili casuali dipendenti dal tempo (processi aleatori). Come è noto, la media statistica rispetto a qualsiasi dimensione elimina la variabilità del fenomeno nella dimensione considerata. Nella realtà la componente spaziale della variabilità è fondamentale nello studio delle patologie infettive perché queste si diffondono proprio attraverso la mobilità delle persone infette o degli agenti patogeni o da entrambi da un luogo ad un altro dello spazio

fisico o del territorio. I modelli matematici che descrivono la diffusione di epidemie devono, quindi, essere spazialmente espliciti, proprio perché le epidemie mostrano configurazioni spaziali eterogenee (M. Gatto et al.2013) [2]. L'introduzione della componente spaziale è una delle sfide della modellistica ecologica degli ultimi decenni, (Diekmann e Heesterbeeck 2000) [3]. Negli ultimi anni, per la aumentata disponibilità di dati geo-referenziati e degli strumenti informatici, i modelli di diffusione spazio-tempo hanno ricevuto, infatti, particolare attenzione. Sono stati implementati su supporto informatico alcuni pacchetti attraverso i quali è possibile produrre delle mappe dinamiche atte a descrivere l'evoluzione spazio-tempo di qualsiasi epidemia. Il tool "surveillance", implementato su R permette di simulare e visualizzare differenti tipologie di epidemie: influenza, meningite, coxiellosi, febbre suina ecc.. (S. Mayer et al. 2017) [4].

Obiettivi

Nel presente lavoro si propone l'utilizzo il modello di Feynman per descrivere e prevedere fenomeni di diffusione di malattie infettive Tale modello è stato ampiamente sperimentato nella meccanica statistica e quantistica. Per rendere più aderente alla realtà degli spostamenti di esseri umani, il modello di Feynman viene integrato con un meccanismo, anche questo utilizzato in meccanica statistica, per trattare gli spostamenti che avvengono per effetto del caso: il modello di Ehrenfest. Questo modello si basa sui trasferimenti delle unità tra le differenti aree geografiche e tali spostamenti possono essere dovuti ad un grande e diversificato numero di cause classificabili in macroscopiche e microscopiche. Le cause macroscopiche di trasferimento possono essere considerate le differenti densità di popolazione nei differenti luoghi, le condizioni socio-economiche (pressione demografica), la presenza di carestie o guerre ecc.. Le cause microscopiche sono, invece, in numero elevato in quanto riguardano ogni unità che si sposta e sono differenti per ciascuna di essa unità. In questo lavoro si ipotizza che gli spostamenti siano dovuti solo alle cause microscopiche. Questo implica che non esistono motivi macroscopici ma ogni unità si sposta per motivi personali e quindi, nel complesso, per effetto del caso. La scala territoriale a cui si può applicare il modello di Ehrenfest può essere di ampiezza minima (solo due località) o massima (l'intero pianeta). Il modello di Ehrenfest permette di verificare facilmente che se un sistema si trova in uno stato lontano dall'equilibrio, esso evolverà spontaneamente verso l'equilibrio. In termini epidemiologici, se una patologia infettiva si sviluppa in una data località, la sua diffusione dipende dal numero di suscettibili presente nella località e dalla loro mobilità sul territorio. Se gli spostamenti delle unità avvengono per caso, allora il numero delle unità che si spostano è proporzionale al numero di unità che compongono la popolazione dalla quale partono. Lo scopo del modello di Ehrenfest è di descrivere e prevedere l'evoluzione spazio-tempo delle popolazioni degli infetti e dei suscettibili.

Metodi

Descriviamo il meccanismo attraverso il quale agisce il modello di Feyneman e quello di Ehrenfest, nel caso molto semplice di due sole località (o regioni): A e B. La numerosità delle popolazioni delle due regioni geografiche, indicate brevemente con A e B, sono descritte da un vettore $[N_A, N_B]$, i cui elementi descrivono lo stato del sistema ad un generico tempo t (che per semplicità omettiamo). Lo stato del sistema viene determinato attraverso il modello di Feynman e, successivamente si applica il meccanismo di Eherenfest. Se il sistema si trova, ad un dato istante, nello stato $[N_A, N_B]$ la probabilità che transiti, in seguito a spostamento d'una unità da A a B, allo stato $[N_A-1, N_B+1]$, (ponendo $N = N_A+ N_B$) è uguale al rapporto N_A/N. Viceversa, la probabilità di passare, in seguito a migrazione d'una unità da B ad A, da $[N_A, N_B]$ a $[N_A+1, N_B-1]$ è uguale al rapporto N_B/N. È evidente che: se $N_A > N_B$ è più probabile la transizione da $[N_A , N_B]$ verso $[N_A-1, N_B+1]$ e se $N_A < N_B$ è più probabile la transizione da $[N_A , N_B]$ verso $[N_A+1, N_B-1]$. E' evidente anche che, se si verifica una epidemia in uno dei due luoghi la velocità di transizione da A a B o da B ad A dipende dalla differenza del numero di infetti (e di suscettibili) nelle due regioni. Tale meccanismo è stato proposto in termini di equazioni differenziali alle derivate parziali con l'ipotesi che differenti regioni si comportino in

modo indipendenti rispetto alla diffusione, da E.A Abdel-Rheim; (2009). A parte la difficoltà matematica del modello ci sembra poco realistica l'ipotesi di indipendenza spaziale introdotta per semplificarne gli aspetti tecnici e poco chiara la separazione tra dinamica temporale e variabilità spaziale.

Risultati

Allo stato attuale del lavoro, i risultati riguardano simulazioni numeriche le quali mostrano come il numero di infetti tende ad equi-distribuirsi tra le differenti aree geografiche con il trascorrere del tempo. Questo risultato è in accordo con le previsioni ottenute nella teoria cinetica dei gas (meccanica statistica) per sistemi in equilibrio. Il modello di Feyneman-Ehrenfest, proposto in questo lavoro, tiene conto della dipendenza spazio-temporale nella diffusione delle epidemie e non sostituisce i classici modelli SIR e SIS ma li integra con la componente stocastica e con ulteriori informazioni inerenti i trasferimenti delle infezioni sul territorio, veicolata dalle migrazioni sia dei soggetti infetti che degli agenti patogeni.

Conclusioni

Il modello di feyneman-Herenfest, inserito nel contesto delle equazioni di Kermack –McKendrick (nel senso che esse determinano l'andamento temporale del numero di infetti, di suscettibili e di rimossi (SIR) mentre gli spostamenti territoriali avvengono con il meccanismo di Ehrenfest) tiene conto simultaneamente della variazione temporale e delle variazioni territoriali o spaziali del numero di unità delle popolazioni (SIR o SIS). Il numero di regioni considerate può essere più o meno elevato e senza vincolo di contiguità territoriale. Le cause macroscopiche di migrazioni possono essere introdotte nel modello come funzioni esterne di tipo stocastiche o deterministiche. Il modello necessita ancora di ulteriori simulazioni numeriche per poter essere utilizzato nella realtà medico-epidemiologica.

Bibliografia

[1] A. S. Mata, S. C. Ferreira, (2015) Multiple transitions of the susceptible-infected-susceptible epidemic model on complexnetworks, Phys. Rev. E91, 012816

[2] M. Gatto, L. Mari, A.Rinaldo (2011) ; Modelli spazio-tempo di diffusione , previsione e controllo delle epidemie di colera: dal Sudafrica ad Haiti; Istituto Lombardo (Rend. Scienze) 145, 47-57

[3] O. Diekmann, J. Heesterbeek, (2000) Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis, and Interpretation. Wiley, Hoboken.

[4] S. Meyer, L. Held, M. Höhle. (2017) Spatio-Temporal Analysis of Epidemic Phenomena Using the R Package surveillance; Journal of Statistical Software, Vol. 77, Issue 11. doi:10.18637

[5] E.A. Abdel-Rehim. (2009) From the Ehrenfest model to time-fractional stochastic processes; Journal of Computational and Applied Mathematics 233 197–207

[6] Dan Braha. (2012) Global Civil Unrest: Contagion, Self-Organization, and Prediction; PLOS ONE /www.plosone.org

[7] M. Di Traglia, F. Attorre, F. Francesconi, R. Valenti, M. Vitalea, (2011); Is cellular automata algorithm able to predict the future dynamical shifts of tree species in Italy under climate change scenarios? A methodological approach. Ecological Modelling, journal homepage:www.elsevier.com/locate/ecolmodel.

[8] A. S. Mata, R. S. Ferreira, S.C. Ferreira, (2014). Heterogeneous pair-approximation for the contact Process on complex networks, New J. Phys.16, 053006

[9] F. Maggi, D. Bosco, L. Galetto, S. Palmano, C. Marzachì. (2017) Space-Time Point Pattern Analysis of Flavescence Dorée Epidemic in a Grapevine Field: Disease Progression and Recovery ;Frontier in Plant Science, | https://doi.org/10.3389/fpls.2016.01987

[10] D. Tilman, P. M. Kareiva. (1997) The Role of Space in Population Dynamics and Interspecific Interactions. Spatial Ecology

ACCURATEZZA DEL TEST DI CAMERIERE PER LA STIMA DELLA MAGGIORE ETÀ: UNA REVISIONE SISTEMATICA.

Dajko Marianxhela[1], De Micco Francesco[2], Rosaria Gesuita[1], Andrea Bucci[1], Andrea Faragalli[1], Edlira Skrami[1], Flavia Carle[1], Cameriere Roberto[3], Luigi Ferrante[1].

[1] *Centro di Epidemiologia, Biostatistica e Informatica Medica, Università Politecnica delle Marche, Ancona*

[2] *Dipartimento di Medicina e Scienze della Salute "Vincenzo Tiberio", Università degli Studi del Molise*

[3] *AgEstimation Project, Istituto di Medicina Legale, Università di Macerata, Macerata*

Introduzione

La stima della maggiore età in individui sprovvisti di documenti validi di identità, o dei quali non è comunque nota l'esatta data di nascita, è una questione fondamentale in ambito medico legale che riguarda molti aspetti della società civile che vanno dalla richiesta di asilo politico, alla possibilità di decidere in autonomia, all'imputabilità di un soggetto. In molti paesi nel mondo, e in tutti i paesi europei, la maggiore età equivale ad aver compiuto i 18 anni. Per questo motivo molti lavori di carattere medico legale riguardano l'individuazione di test diagnostici per la stima dei 18 anni. Negli ultimi anni diversi test diagnostici sono stati proposti per la maggiore età. Dal 2008, il metodo riportato in Cameriere et al [1], è uno tra i più utilizzati per eseguire il test per la stima della maggiore età.

Obiettivo

Scopo del presente lavoro è eseguire una revisione sistematica degli studi che hanno utilizzato il test di Cameriere per la maggiore età in popolazioni diverse per area geografica ed etnia.

Metodi

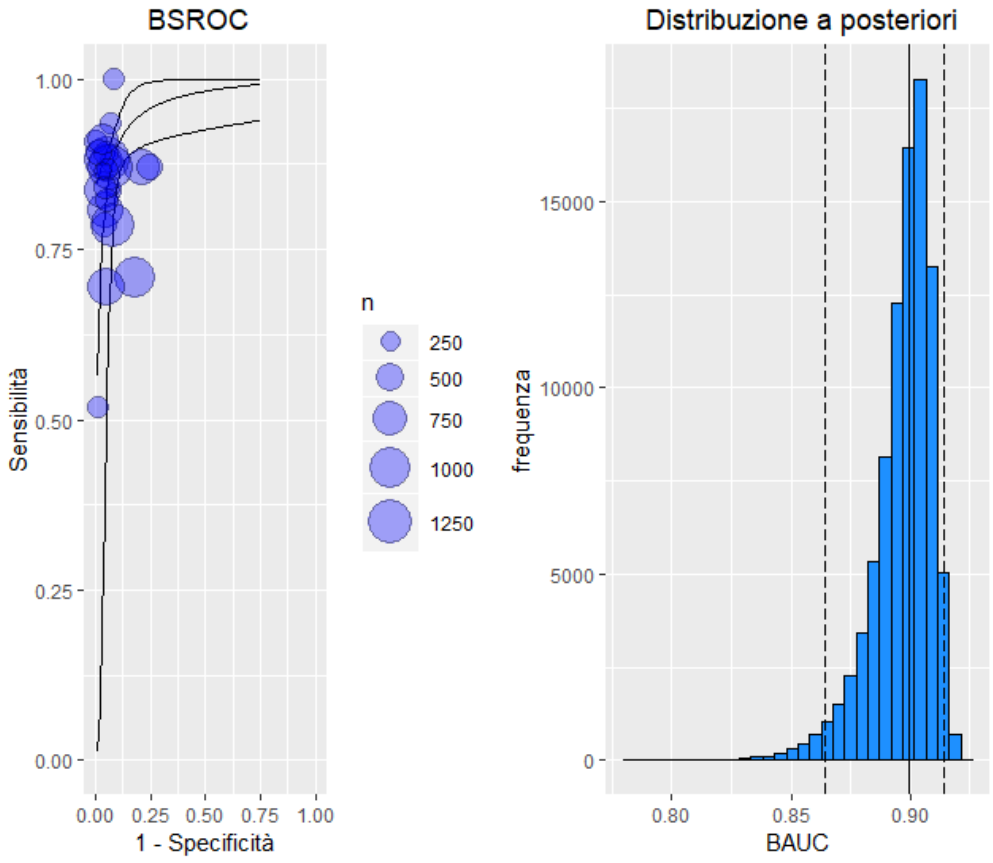
La ricerca è stata effettuata su PubMed, Embase e la Cochrane Library per gli studi pubblicati dal 2008 fino al 31 febbraio 2019. Gli studi sono stati inclusi se riportavano il test basato sull'indice di maturazione I_{3M} come descritto in [1] e contenevano dati sufficienti per costruire una tabella di contingenza 2 × 2, relativa alla classificazione per la maggiore età dei soggetti. In breve, il test si basa sulla misura dell'apertura degli apici, I_{3M}, del terzo molare della mandibola: se $I_{3M} > 0.08$ allora il soggetto viene considerato minorenne, altrimenti maggiorenne. I modelli di meta-analisi applicati a studi relativi ai test diagnostici hanno diversi aspetti peculiari. Uno dei principali aspetti riguarda la correlazione tra le misure d'interesse. Infatti, la sensibilità e la specificità di un test sono generalmente dipendenti e quindi la combinazione delle loro distribuzioni marginali a scopi inferenziali potrebbe essere fuorviante poiché non tiene conto della loro correlazione [2].
Nel presente lavoro è stato utilizzato un modello di tipo Bayesiano che permette di valutare la sensibilità e la specificità aggregate ed anche la loro correlazione, ρ, con i relativi Intervalli di Confidenza al 95% (IC 95%) [3] nella stima della validità globale del test. Il modello statistico è caratterizzato da una mistura di distribuzioni normali con matrici di varianza e covarianza dipendenti da ciascuno studio mediante opportuni pesi w_i che, nel modello Bayesiano, vengono considerati variabili aleatorie con distribuzione gamma inversa e utilizzati per la valutazione dell'eterogeneità degli studi. In particolare, uno studio viene considerato non omogeneo se:

$$P(w_i > 1|Dati) \geq 0.7 \quad i = 1, \dots, k_{studi}$$

Risultati

Ventiquattro studi con un totale di 13900 partecipanti sono stati inclusi in questa meta-analisi. Tra i 24 studi inclusi 11 (45.83%) sono stati svolti in Europa, 3 (12.5%) in Africa, 3 (12.5%) nell'America del Sud, 6(25%) nel continente asiatico e 1 (4.2%) in Australia. La sensibilità e la specificità aggregate globali sono risultate 0.85 (IC 95% 0.80; 0.88) e 0.93 (IC 95% 0.91; 0.95), rispettivamente. La correlazione tra sensibilità e specificità è risultata ρ = -0.10 (IC 95% -0.61; 0.44) L'area sotto la curva BSROC (BAUC), ottenuta con il modello Bayesiano considerato, è risultata pari a 0.90 (IC 95%: 0.86; 0.91). Gli studi presentavano una sostanziale omogeneità. Sei lavori, [4-9], sono risultati non omogenei in relazione alla sensibilità o alla specificità del test. In particolare, il lavoro [6] aveva una sensibilità del 100%, mentre il lavoro [9] aveva una specificità del 100%.

Figura1. *Il grafico di sinistra mostra la curva BSROC: la linea centrale corrisponde alla mediana della distribuzione a posteriori e le curve superiore e inferiore corrispondono ai quantili di ordine 2,5% e 97,5% rispettivamente. Il grafico di destra mostra la distribuzione a posteriori dell'area sotto la curva BSROC.*



Conclusione

I risultati hanno mostrato che il test proposto per la stima della maggiore età è affidabile indipendentemente dalla popolazione in cui viene applicato.

Bibliografia

[1] R. Cameriere, L. Ferrante, D. Angelis, F. Scarpino, F. Galli. "The comparison between measurement of open apices of third molars and Demirjian stages to test chronological age of over 18 year olds in living subjects". International Journal of Legal Medicine, 2008, Volume 122. doi: 10.1007/s00414-008-0279-6.

[2] Irwig L, Macaskill P, Glasziou P, Fahey M (1995). "Meta-Analytic Methods for Diagnostic Test Accuracy." Journal of Clinical Epidemiology, 48, 119–130. doi:10.1016/ 0895-4356(94)00099-c

[3] Verde PE (2018). bamdit: Bayesian Meta-Analysis of Diagnostic Test Data. R package version 3.2.1, URL https://CRAN.R-project.org/package=bamdit.]

[4] Cameriere R, Velandia Palacio LA, Pinares J, Bestetti F, Paba R, Coccia E, Ferrante L. Assessment of second (I2M) and third (I3M) molar indices for establishing 14 and 16 legal ages and validation of the Cameriere's I3M cut-off for 18 years old in Chilean population. Forensic Sci Int. 2018 Apr;285:205.e1-205.e5. doi: 10.1016/j.forsciint.2017.12.043.

[5] Deitos AR, Costa C, Michel-Crosato E, Galić I, Cameriere R, Biazevic MG. Age estimation among Brazilians: Younger or older than 18? J Forensic Leg Med. 2015 Jul;33:111-5. doi: 10.1016/j.jflm. 2015.04.016.

[6] Gulsahi A, De Luca S, Cehreli SB, Tirali RE, Cameriere R. Accuracy of the third molar index for assessing the legal majority of 18 years in Turkish population. Forensic Sci Int. 2016 Sep;266:584. e1-584.e6. doi: 10.1016/j.forsciint.2016.04.034.

[7] AlQahtani S, Kawthar A, AlAraik A, AlShalan A. Third molar cut-off value in assessing the legal age of 18 in Saudi population. Forensic Sci Int. 2017 Mar;272:64-67. doi: 10.1016/j.forsciint.2017. 01.004.

[8] Sharma P, Wadhwan V, Sharma N. Reliability of determining the age of majority: a comparison between measurement of open apices of third molars and Demirjian stages. J Forensic Odontostomatol. 2018 Dec 1;2(36):2-9.

[9] Dardouri AAK,Cameriere R, De Luca S, Vanin S. Third molar maturity index by measurements of open apices in a Libyan sample of living subjects.Forensic Sci Int. 2016 Oct;267:230.e1-230.e6. doi: 10.1016/j.forsciint.2016.07.020.

PREDICTION OF TREATMENT OUTCOME IN CLINICAL TRIALS: A MACHINE LEARNING APPROACH

Berchialla Paola[1], Lanera Corrado[2], Gregori Dario[2], Baldi Ileana[2]

[1] *Depariment of Clinical and Biological Sciences, University of Torino*

[2] *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padova,*

Introduction

A critical early step in a clinical trial design is to define the study population from which the sample will be drawn, i.e. to identify the target population most likely to derive benefit from the experimental treatment. Envisaging a "run-in" period in a study design may accomplish this task. [1]
The "run-in" is a period before randomization, whose aim is detecting/excluding subgroups of patients less likely to respond to the therapy. Incorporating such a pre-randomization period in a study design constitutes a desirable enrichment process of a clinical study [2] but it requires additional patients, increasing times and costs. Being capable to use available a-priori data to inform about the potential patients' outcome, conditionally to the therapy received, could lead to skip such period. This way, in clinical trials, predictions could improve the population segment selection and the consequent enrolment [3].
Machine learning (ML) is an interdisciplinary field aimed at developing models with the maximal predictive accuracy. A distinctive key of ML algorithms is their capability to improve their predictive performance through experience [4]. In a clinical setting, all the amount of different historical, a-priori known information as well as patient everyday-life data [5] can be employed to learn by improving complex tasks such as classification (e.g. response to therapy, cancer clinical-type classification), clustering (e.g. identification of groups of patients with shared characteristics). For these reasons, ML is increasingly applied to clinical studies [6, 7] and it represents a new approach towards conducting medical research and developing ways to predict individual outcomes [7–9].

Objectives

The aim of this paper is to investigate the usage of ML predictive capabilities in clinical trials, during early accrual in the spirit of a population enrichment approach. With this purpose, an ensemble model of ML techniques was trained on a couple of twin randomized clinical trials in order to learn from one trial data and mimic a run-in process.

Methods

We considered the SAIS1 [10], and the PROLOGUE [11] Randomized Clinical Trials (RCTs) for our investigation. Both have studied the effect of Sitagliptin on patients with type 2 diabetes and have collected a common subset of patient measures. For both RCTs, anonymized data are freely available for analysis.
To investigate the ML predictive capabilities to mimic a run-in process, the main strategy was unfold as following: i) a ML algorithm was trained on the treated patients' characteristics (i.e., patients in Sitagliptin arm) of the SAIS1 RCT; ii) the training accuracy was assessed on the same patients through cross-validation; iii) different cut-offs based on the predicted probability of a patient to be a "positive responder were established to select set of patients in the PROLOGUE RCT; iv) the effect of the treatment in the PROLOGUE RCT was evaluated on the above set of patients selected according to their probability of being a responder; (v) finally the impact of this strategy on the variability, heterogeneity of the effect, and the power of the study was assessed.

Investigation design. Common patients' characteristics collected by the two studies were: age (year), gender, Body Mass Index (BMI, kg/cm2), Systolic Blood Pressure (SBP, mmHg), Diastolic Blood Pressure (DBP, mmHg), hypertension (SBP >= 130 mmHg OR DBP >= 80 mmHg), Low-Density Lipoprotein (LDL, mg/dl), High-Density Lipoprotein (HDL, mg/dl), HbA1c (%), Fasting Plasma Glucose (FPG, mmol/l), dyslipidemia (LDL >= 130 mg/dl OR HDL < 35 mg/dl OR triglyceride >= 150 mg/dl OR total cholesterol (= LDL + HDL + (Triglyceride / 5)) >= 200 mg/dl), adiponectin (mg/l).
We set as outcome an improvement of at least -0.5% in HbA1c (dichotomized). As predicting covariates, we used age, gender, BMI, hypertension, dyslipidemia, adiponectin, SBP, DBP, Baseline HbA1c, FPG, and LDL.
Hyper parameter tuning was conduct by a 5-fold cross-validation process for which each fold was balanced to maintain the same ratio of the overall dataset in each training and validation sample. The overall performance was measured by the Area Under the Receiver Operating Characteristics Curve (AUC-ROC).
Missing values of the PROLOGUE dataset were imputed using Multivariate Random Forest Imputations by Chained Equations imputing covariates ordered low to high proportion of missing data.
We use the predicting model estimated on SAIS1 dataset to give a probability to be a responder (i.e. a patient with less or equal of -0/5% of reduction in HbA1c) to each patient into the PROLOGUE dataset.
Using the ten-lowest distinct possible prediction cut-off (i.e., the ten-lowest distinct probability for the trained model applied on the training data itself), we filtered the PROLOGUE dataset into a set of nested subsets.
Each one of the reduced PROLOGUE datasets, as well as the full dataset, were used to fit a "cut-off specific" independent Generalized Linear Model to estimate the effect of Sitagliptin on the variation at 12 months of HbA1c (on the continuous scale). If the model was fitted on a reduced dataset it was also adjusted by the patients' characteristics, i.e., treatment of diabetes mellitus prior to randomization (Pharmacological or not), use of statin, age, gender, SBP at office, baseline HbA1c and Maximum Common Carotid Intimal Medial Thickness.
The effect of the treatment for each subset filtered from the PROLOGUE RCT was finally assessed along with 95% Confidence Interval.
The distributions of the Sitagliptin effect by arm for each cut-off level were reported for both the reduced and excluded populations.

Super Learner. The ML algorithm model was a Super Learner (SL) algorithm. A SL is an ensemble of ML techniques (base learners) combined in a way that it is theoretically proved as being asymptotically as good as the oracle selector, i.e. the best possible weighted combination of the base learners [12]. The base learners used in the combination of the SL were: gradient boosting machine, generalized linear model, polychotomus regression model, random forest, classification trees, Bayesian additive regression trees, support vector machine.

Results

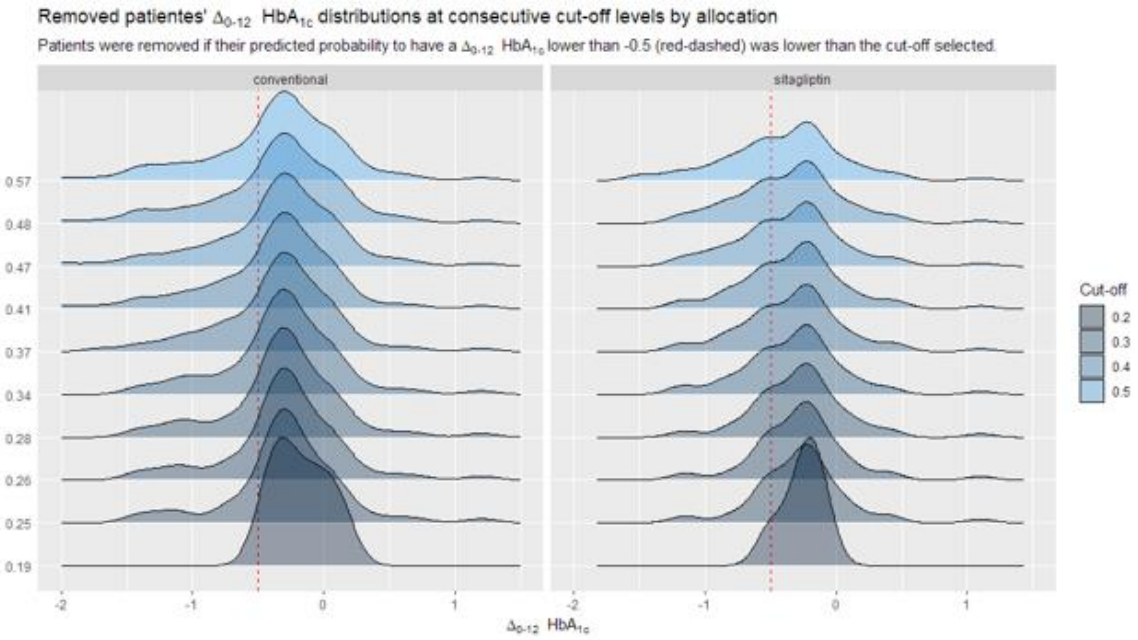
The SAIS1 study involved n = 103 patients with type 2 diabetes, between 20 and 75 years of age, and levels of Glycated Hemoglobin (HbA1c) between 6.9% and 8.4% at the baseline. Patients were allocated on glimepiride (n = 55), and on Sitagliptin (n = 48). Follow-up information were collected once, at 6 months. Levels of HbA1c at the follow-up time were not collected for 5 out of 48 patients in the Sitagliptin arm. Accordingly, we included in our study the 43 SAIS patients with fully collected levels of HbA1c at the follow-up time.
The PROLOGUE study initially involved and analyzed data on n = 385 patients with type 2 diabetes, at least 30 years of age, and levels of HbA1c between 6.2% and 9.4% at the baseline. Patient were allocated on conventional therapy (n = 193), and on Sitagliptin (n = 192). Follow-up information were collected at 12 months and at 24 months.

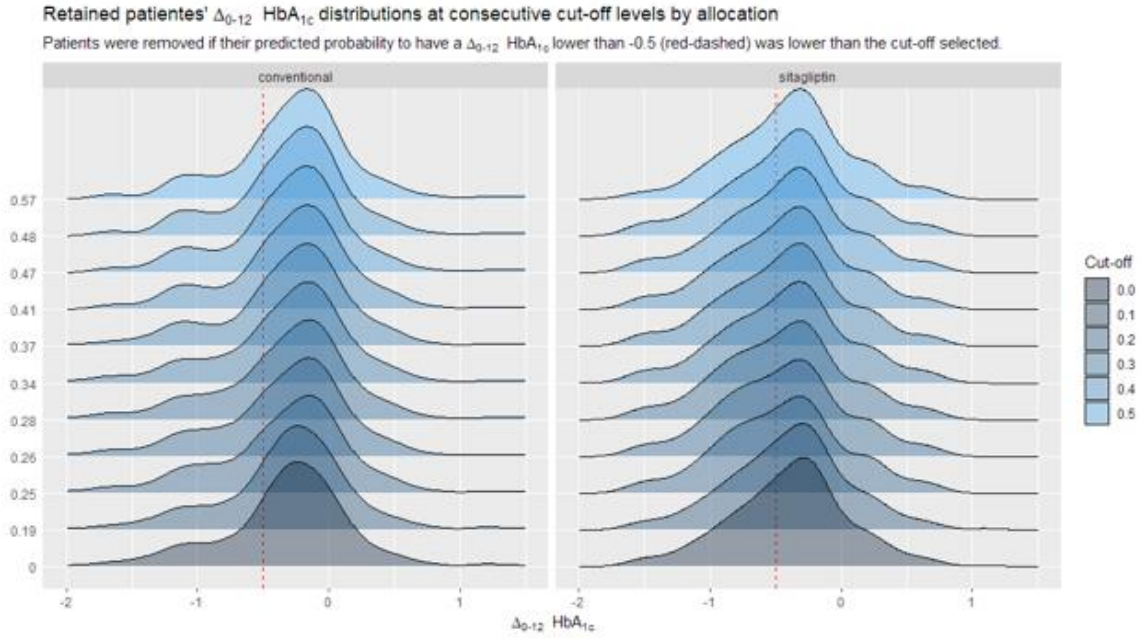
In the top of figure 1 is shown the distribution of HbA1c arm among those patients who were removed from both conventional and Sitagliptin arm because they had a probability to be a responder lower than the pre-specified cut-off.
In the bottom of figure 1 is shown the distribution of HbA1c arm among those patients who were retained in both conventional and Sitagliptin arm at different pre-specified cut-off levels of the probability to be a responder. The treatment effect is assessed at each subset of patients retained and it is not constant across these different patient subpopulations. This effect can be attributed to the heterogeneity of treatment effect and could suggest an interaction between treatment and patient characteristics.

Conclusions

Personalized medicine is deeply connected to machine learning that provides computational methods with the potential to impact current practice [13].
In the present work we used a ML ensemble algorithm applied to RCTs open data to find evidence of patient-specific treatment effect. Investigating variability in patient response is important to assess the possibility of individual response to treatment. In the setting of clinical trials, individual response to treatment can be used to improve patients' enrollment and identify patient sub-populations.
In Senn et al. [14] it has been shown how patient random effects can be estimated in a replicated crossover study. In this study we used two RCTs, similar in terms of inclusion, exclusion criteria and estimands, to train a ML algorithm on one of the trials to predict future responders and optimize treatment selection. The success of this approach depends on the availability of open data clinical trials.

Figure 1. *Top: distribution of HbA1c in the patients excluded on the basis of their predictive probability to be a responder at different cut-offs, in both conventional and Sitagliptin arms; bottom: distribution of HbA1c in the patients selected on the basis of their predictive probability to be a responder at different cut-offs, in both conventional and Sitagliptin arms*





References

[1] Rathod T., Belcher J., Montgomery A.A. et al. Health services changes: is a run-in period necessary before evaluation in randomised clinical trials? Trials. 2014 Jan; 30;15:41.

[2] Berger V.W., Durkalski V.L. Run-In Period. In: Wiley StatsRef: Statistics Reference Online [Internet]. American Cancer Society; 2014 [cited 2019 Feb 19]. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat07112

[3] Fitzpatrick A.M., Jackson D.J., Mauger D.T. et al. Individualized therapy for persistent asthma in young children. J Allergy Clin Immunol. 2016;138(6):1608-1618.e12.

[4] Flach P. Machine Learning by Peter Flach [Internet]. Cambridge Core. 2012 [cited 2019 Feb 19]. Available from: /core/books/machine-learning/621D3E616DF879E494B094CC93ED36A4

[5] Dobkin B.H., Dorsch A. The promise of mHealth: daily activity monitoring and outcome assessments by wearable sensors. Neurorehabil Neural Repair. 2011 Dec;25(9):788–98.

[6] Asadi H, Dowling R, Yan B, Mitchell P. Machine Learning for Outcome Prediction of Acute Ischemic Stroke Post Intra-Arterial Therapy. PLOS ONE. 2014 Feb 10;9(2):e88225.

[7] Oermann E.K., Rubinsteyn A., Ding D. et al. Using a Machine Learning Approach to Predict Outcomes after Radiosurgery for Cerebral Arteriovenous Malformations. Sci Rep. 2016 Feb; 9;6:21161.

[8] Kourou K., Exarchos T.P., Exarchos K.P. et al. Machine learning applications in cancer prognosis and prediction. Computational and Structural Biotechnology Journal. 2015 Jan; 1;13:8–17.

[9] Shouval R., Bondi O., Mishan H. et al. Application of machine learning algorithms for clinical predictive modeling: a data-mining approach in SCT. Bone Marrow Transplant. 2014 Mar;49(3):332–7.

[10] Nomoto H, Miyoshi H, Furumoto T. et al. A Randomized Controlled Trial Comparing the Effects of Sitagliptin and Glimepiride on Endothelial Function and Metabolic Parameters: Sapporo Athero-Incretin Study 1 (SAIS1). PLoS ONE. 2016;11(10): e0164255.

[11] Oyama J.I., Murohara T., Kitakaze M. et al. The Effect of Sitagliptin on Carotid Artery Atherosclerosis in Type 2 Diabetes: The PROLOGUE Randomized Controlled Trial. PLoS Med. 2016 Jun;13(6):e1002051.

[12] Polley E, LeDell E, Kennedy C et al. SuperLearner: Super Learner Prediction [Internet]. 2018 [cited 2019 Feb 22]. Available from: https://CRAN.R-project.org/package=SuperLearner

[13] Frohlich H, Balling R, Beerenwinkel N et al. From hype to reality: data science enabling personalized medicine. BMC Medicine, 2018; 16(1): 150

[14] Senn, S., Rolfe, K., & Julious, S. A. Investigating variability in patient response to treatment – a case study from a replicate cross-over study. Statistical Methods in Medical Research, 2011; 20(6), 657–666.

SINDROME CORONARICA ACUTA: DATI DAL MONDO REALE IN LOMBARDIA

Ronco Raffaella[1,2], Rea Federico[1,2], Merlino Luca[3], Corrao Giovanni[1,2]

[1] *Centro Nazionale di Healthcare Research & Pharmacoepidemiology, Italia*

[2] *Laboratorio di Healthcare Research & Pharmacoepidemiology, Unità di Biostatistica, Epidemiologia e Sanità Pubblica, Dipartimento di Statistica e Metodi Quantitativi, Università degli studi di Milano-Bicocca, Italia*

[3] *Osservatorio epidemiologico, Regione Lombardia, Italia*

Introduzione

La sindrome coronarica acuta (SCA) comprende le diverse manifestazioni dell'ischemia miocardica, ossia dell'assenza (totale o parziale) di afflusso di sangue al tessuto miocardico, con conseguente carenza di ossigeno ed elevato rischio di sviluppare necrosi cardiaca. [1] [2] In Italia si stima che il numero di soggetti colpiti da eventi coronarici in un anno superi i 135.000, e che di questi circa 45.000 siano eventi fatali [3].
La SCA è una sindrome tempo-dipendente: al trascorrere del tempo aumenta la sofferenza cellulare e la probabilità di conseguente necrosi. Per questo motivo, è fondamentale riconoscere i sintomi all'insorgenza e raggiungere il più velocemente possibile la struttura appropriata più vicina.
In letteratura si è evidenziato come una diagnosi tempestiva di SCA, seguita da un opportuno intervento, sia in grado di ridurre il rischio di esiti negativi [4] [5]. Dopo il primo intervento, i pazienti necessitano di un costante monitoraggio, motivo per cui sono solitamente ricoverati in un reparto di degenza o di terapia intensiva coronarica. Anche dopo la dimissione, ai pazienti possono essere prescritte terapie di supporto e di riabilitazione al fine di gestire la cronicità della malattia coronarica, richiedendo un costante monitoraggio da parte del medico di base o ambulatoriale [6].

Obiettivi

Questo studio vuole valutare la qualità della prevenzione secondaria dei pazienti dimessi con diagnosi di SCA In Lombardia ed esaminare l'impatto dell'aderenza a diverse indicazioni terapeutiche rispetto a successivi ricoveri per cause cardiovascolari.

Metodi

Questo studio è basato sui database amministrativi della Regione Lombardia, contenenti informazioni relative alle schede di dimissione ospedaliera, prescrizioni farmaceutiche, prestazioni ambulatoriali, domiciliari e di pronto soccorso ed altre informazioni di natura anagrafica ed amministrativa.
Nello studio sono stati inclusi tutti i soggetti ospedalizzati in regime di urgenza o recatisi in pronto soccorso con diagnosi di SCA nel periodo compreso tra il 01/01/2011 ed il 31/12/2015, definendo la data di dimissione del primo ricovero come data indice. Tra questi, sono poi stati identificati tutti i pazienti con un'età compresa tra i 40 e i 90 anni.
Dallo studio sono poi stati esclusi tutti i pazienti che non risultavano beneficiari del Servizio Sanitario Regionale da almeno 5 anni al momento dell'ingresso in studio o che avevano già sperimentato un ricovero per SCA nello stesso periodo.
Ogni paziente rimanente è stato dunque incluso nella coorte in studio ed ha accumulato anni-persona dalla data indice alla prima tra le date dei seguenti eventi: ricovero con diagnosi di esito cardiovascolare, decesso, emigrazione o fine del periodo di osservazione (30/06/2018).
Per ogni paziente incluso nello studio sono state raccolte le informazioni relative alle caratteristiche misurate al baseline e durante il primo anno di follow-up.

Sono stati raccolti i valori al momento del reclutamento di **(a)** età **(b)** sesso, **(c)** indice di comorbosità (Multisource Comorbidity Score, MCS) [7], e la presenza nei 5 anni precedenti il ricovero indice di: **(a)** ospedalizzazioni per altre patologie (diabete, ipertensione, dislipidemia, depressione, BPCO, ictus/TIA, insufficienza renale cornica [8] [9] [10] [11] [12]), **(b)** prescrizioni di farmaci considerati proxy delle patologie sopraelencate.
Durante il primo anno di follow-up sono stati valutati: **(a)** test di laboratorio (profilo lipidico), **(b)** visite cardiologiche ambulatoriali, **(c)** terapie farmacologiche indicate per questi pazienti (beta-bloccanti, antiaggreganti piastrinici, sostanze che agiscono sul sistema renina-angiosteina, statine).
Per valutare come l'aderenza alle raccomandazioni influenzasse la prognosi dei pazienti, sono stati selezionati i soggetti con un follow-up minimo di un anno.
Disegni user-only sono stati utilizzati per valutare l'impatto dell'aderenza alle singole prescrizioni farmacologiche sull'esito di interesse. Seguendo il principio dell'intention to treat, è stata calcolata l'aderenza terapeutica nel primo anno e sono stati considerati aderenti al trattamento i soggetti che sono risultati coperti dalla terapia per almeno il 75% di tale periodo (in accordo con la metrica della Proportion of Days Coverage, PDC). In un secondo momento, la robustezza dei risultati ottenuti è stata validata replicando le analisi appaiando per High Dimensional Propensity Score (HDPS) [13], età, genere e classe di MCS [7] i soggetti aderenti al trattamento e i soggetti con almeno una prescrizione del farmaco in esame ma non aderenti.
Infine, per valutare l'impatto di test di laboratorio e visite cardiologiche sono stati implementati modelli sulla coorte complessiva. Anche in questo caso i modelli sono poi stati replicati appaiando ad ogni soggetto che ha fatto la procedura di interesse un soggetto con uguale HDPS, genere, età e MCS che non si è sottoposto a tale procedura.
Per valutare quanto la probabilità di un evento cardiovascolare dipenda dalle variabili di interesse, sono stati utilizzati modelli dei rischi proporzionali di Cox, che permettono di caratterizzare la sopravvivenza all'evento stesso attraverso stime di Hazard Ratio (HR) affiancate da intervalli di confidenza (IC) al 95%.

Risultati

Fra il 2011 e il 2015 sono stati registrati 97.353 episodi di SCA fra pronto-soccorsi e ricoveri in regime di urgenza di soggetti con un'età compresa fra i 40 e i 90 anni. Di questi, 5.326 sono stati i decessi durante il primo accesso alla struttura ospedaliera, definito come ricovero indice. Dai 92.027 soggetti sopravvissuti al ricovero indice sono stati selezionati i 87.530 incidenti all'evento oggetto di analisi, ossia senza episodi di SCA nei 5 anni precedenti l'evento.
Fra questi, la probabilità di mortalità intraospedaliera per SCA è emersa essere pari al 5,47, in particolare si distribuisce nel tempo come segue: 5,46 nel 2011, 5,82 nel 2012, 5,15 nel 2013, 5,56 nel 2014 e 5,37 nel 2015.
Infine, sono stati inclusi nello studio i 61.662 soggetti che hanno sperimentato un follow-up minimo di un anno. La coorte così formata è risultata composta per il 67.3% da uomini e per il 32.7% da donne, con un'età media di 67.5 (standard deviation 11.99) fra i due gruppi, distinta in 72.2 nelle donne e 65.2 negli uomini.
47,603 soggetti, circa il 77.2% della coorte, è stato ricoverato con una diagnosi di infarto, mentre il restante di angina.
Valutando gli indicatori di comorbidità nel wash-out è emerso che 3.101 (5,0%) soggetti hanno avuto prescrizioni di antiaritmici, 3.418 (5,5%) di anticoagulanti, 17.880 (29,0%) soggetti hanno avuto prescrizioni o ricoveri indici di BPCO, 4.074 (6,6%) di malattie cerebrovascolari, 10.052 (16,3%) di depressione, 12.621 (21,5%) di diabete, 25.126 (40,7%) di dislipidemia e 559 (0,9%) di insufficienza renale cronica. Inoltre, il 95,0% dei pazienti presentava almeno una delle condizioni che compongono il MCS.
Dei 61.662 soggetti inclusi nello studio e sopravvissuti al primo anno di follow-up, 60.472 (98,1%) hanno almeno una prescrizione delle terapie farmacologiche di interesse; di questi, 53.181 (87,9%) risultano aderenti ad almeno una classe di farmaci nell'anno successivo il ricovero indice.

In particolare, 51.222 (83,1%) hanno avuto almeno una prescrizione di farmaci che agiscono sul sistema renina-angiotensina (ACE inibitori e/o sartani), di questi 28.914 (56,5%) sono risultati aderenti alla terapia. Il modello users-only in cui sono stati valutati i soggetti aderenti alla terapia versus i soggetti non aderenti ha evidenziato un HR di 0,95 (IC: 0,92 – 0,99).
55.106 (89,4%) hanno invece avuto almeno una prescrizione di antiaggreganti piastrinici, di questi 34.209 (62,1%) sono risultati aderenti alla terapia. I soggetti aderenti hanno mostrato una riduzione del 9% (IC: 6-11) nel rischio di ricovero per cause CV.
I soggetti che hanno avuto almeno una prescrizione di betabloccanti sono invece 51.765 (83,9%) di cui 29.399 (56,8%) aderenti alla terapia. Il modello users-only in cui sono stati valutati i soggetti aderenti alla terapia versus i soggetti non aderenti ha evidenziato un HR di 0,94 (0,91 – 0,97).
Infine, l'ultima classe di farmaci considerata è quella delle statine: dei 53.436 (86,7%) soggetti che hanno avuto almeno una prescrizione nel primo anno, 37.466 (70,1%) sono risultati aderenti in tale periodo. Su di essi, un modello users-only ha fornito un HR di 0,88 (0,85 – 0,91).
44,173 soggetti sui 61,662 reclutati, circa il 71.6%, è stato sottoposto ad almeno una visita cardiologica durante il primo anno. Il modello applicato ha fornito un HR di 1.002 (IC: 0.970 – 1.035).
Infine, 49,359 (80.05%) soggetti sono stati sottoposti ad esami di controllo per il profilo lipidico, fornendo un HR pari a 0.934 (0.900 – 0.969).
Tutti i risultati ottenuti sono stati confermati dai modelli ottenuti applicando il metodo di HDPS.

Conclusioni

La probabilità di mortalità intraospedaliera e a breve termine (entro un anno) per SCA sembra essersi ridotta negli ultimi decenni [3], tuttavia questa riduzione della mortalità potrebbe generare un incremento del numero di malati cronici a carico del sistema sanitario regionale. Per questo motivo, è necessario indagare quali fattori possono condizionare la prognosi dei pazienti a lungo termine.
Sebbene si sia osservato che solo il 2% dei soggetti con un episodio di SCA non assumono alcuna terapia farmacologica entro un anno dal ricovero, il 13% di quelli che iniziano una terapia la discontinuano entro tale periodo. Dalle analisi è emerso che un'adeguata aderenza alla terapia farmacologica può ridurre il rischio di esiti cardiovascolari a lungo termine dal 5 al 13% circa.
Risultati protettivi sono emersi anche riguardo l'analisi del profilo lipidico, in particolare si stima che chi viene sottoposto a tale esame rischia il 7% in meno di chi non viene seguito in tal senso.
Sebbene sia consigliato effettuare almeno una visita cardiologica di controllo entro il primo anno e il 30% dei soggetti non segua tale indicazione, la mancata aderenza a tale raccomandazione non ha fornito risultati significativamente importanti.
Si può concludere che i soggetti sottoposti ad un più attento monitoraggio durante il follow-up e maggiormente aderenti alle terapie prescritte, sia in termini di trattamenti farmacologici, sia di esami ambulatoriali, abbiano un rischio ridotto di avere esiti cardiovascolari a lungo termine, il che invita a prestare una particolare attenzione alla prevenzione secondaria nei soggetti colpiti da episodi di SCA.

Bibliografia

[1] Makki N., Brennan T. N. Girotra S., Acute Coronary Syndrome. Journal of Intensive Care Medicine, 2015; 30(4): 186-200.

[2] NHR Biomedical Research Unit, Barts Hearth Centre, London, Acute coronary syndromes. BMJ, 20 Oct. 2015.

[3] Perugini E., Maggioni A. P., Boccanelli A. et al., Epidemiologia delle sindromi coronariche acute in Italia. G Ital Cardiol, 2010; 11(10): 718-729.

[4] Jiang M., Mao J., Pu K. et al, Timing of early angiography in non-ST elevation acute coronary syndrome. J Invasive Cardiology, 2014; 26(2): 47-54.

[5] Deharo P., Bode C., Cohen M. et al, Timing of angiography and outcomes in high-risk patients with non.ST-segment-elevation myocardial infraction managed invasively. Circulation, 2017; 136:1895-1907.

[6] Piironen M., Ukkola O., Huikuri H. et al, Trends in long-term prognosis after acute coronary syndrome. European journal of preventive cardiology, 2017; 24(3): 274-280.

[7] Corrao G., Rea F., Di Martino M. et al, Developing and validating a novel multisource comorbidity score from administrative data: a large population-based cohort study from Italy. BMJ Open, 2017.

[8] Foussas S. G., Acute Coronary Syndrome and Diabetes mellitus. Hellenic Journal of Cardiology, 2016; 57: 375-377.

[9] Foussas S. G., Acute Coronary Syndromes and Renal Disease. Hellenic Journal of Cardiology, 2016; 57, 210-213.

[10] Surani S. R., Diabetes, sleep apnea, obesity and cardiovascular disease: Why not address them together?. World Journal of Diabetes, 15 Jun 2014; 5(3): 381-384.

[11] Woo J. G., Morrison J. A., Stroop D. M. et al, Genetic architecture of lipid traits changes over time and differs by race: Princeton Lipid Follow-up Study. Journal of Lipid Research, 2014; 55: 1515-1524.

[12] Dasgupta K., Quinn R. R., Zarnke K. B. et al, The 2014 Canadian Hypertension Education Program Recommendations for Blood Pressure Measurement, Diagnosis, Assessment of Risk, Prevention, and Treatment of Hypertension. Cardinal Journal of Cardiology, 2014; 30: 485-501.

[13] Schneeweiss S., Rassen J. A., Glynn R. J. et al, High-dimentisonal propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology, 2011.

[14] Mauri F., Epidemiologia, patogenesi e fisiopatologia degli eventi dopo infarto miocardico acuto. AIM, 2008; 6S-11S.

ANALISI DEGLI ERRORI DI I E II TIPO IN ALCUNI TEST DI STATISTICA CIRCOLARE PER IL RICONOSCIMENTO DI TREND CICLICI

Rocchi Ettore[1], Sisti Davide[1], Amatori Stefano[1], Rocchi Marco B.L.[1]

[1]*Dipartimento di Scienze Biomolecolari, Servizio di Biostatistica, Università di Urbino Carlo Bo*

Introduzione

Numerosi test sono impiegati per il riconoscimento e l'analisi di trend ciclici, ma in generale la loro scelta (specie in ambito epidemiologico) non è fondata su una analisi della potenza di tali test nelle diverse condizioni. In questo lavoro vengono presi in considerazione i seguenti approcci: il test di Rayleigh [1], il test di Edwards [2], il test di Roger [3]. In aggiunta viene preso in considerazione il test χ^2 per la goodness-of-fit.
Tutte queste tecniche appartengono alla cosiddetta statistica circolare, ovvero a quella famiglia di tecniche che si applicano a variabili su scala periodica (angoli, ma anche variabili temporali quali giorno dell'anno, giorno della settimana, ora del giorno e così via, tutte trasformabili comunque in variabili angolari). Il test di Rayleigh impiega, come statistica test, la lunghezza del vettore medio, la quale sotto l'ipotesi nulla di distribuzione circolare uniforme (ovvero di isotropia) dovrebbe essere pari a 0. Il test di Edwards impiega come statistica test la distanza tra il centro della circonferenza goniometrica e il baricentro della stessa circonferenza dopo che vi siano stati applicati, in corrispondenza ai dati angolari, dei pesi pari alla radice quadrata delle frequenze degli stessi dati. Il test di Roger, infine, utilizza lo stesso principio del test di Edwards, ma utilizza pesi pari alla frequenza dei dati in luogo della sua radice quadrata.

Obiettivi

L'obiettivo del presente lavoro è quella di fornire una valutazione degli errori di I e II tipo dei quattro test considerati (tutti fondati sull'ipotesi nulla di una distribuzione circolare uniforme) sotto diverse condizioni reali: i. distribuzione di von Mises; ii. mistura di distribuzione di una von Mises e una distribuzione circolare uniforme; iii. mistura di due distribuzioni di von Mises. Tali condizioni vengono studiate al variare dei parametri k (parametro di concentrazione della von Mises), θ (parametro di tendenza centrale della von Mises), p (parametro di proporzione delle misture). In definitiva tutti e tre i casi si possono ridurre al caso generale

$$f(x) = p\frac{1}{2\pi I_0(\mathrm{k}_1)}\exp[\mathrm{k}_1\cos(x-\theta_1)] + (1-p)\frac{1}{2\pi I_0(\mathrm{k}_2)}\exp[\mathrm{k}_2\cos(x-\theta_2)]$$

dove I_0 è la funzione I di Bessel di ordine 0, e dove la funzione $f(x)$ diventa una semplice von Mises se p=1 e la von Mises diventa a sua volta una circolare uniforme se k=0.

Metodi

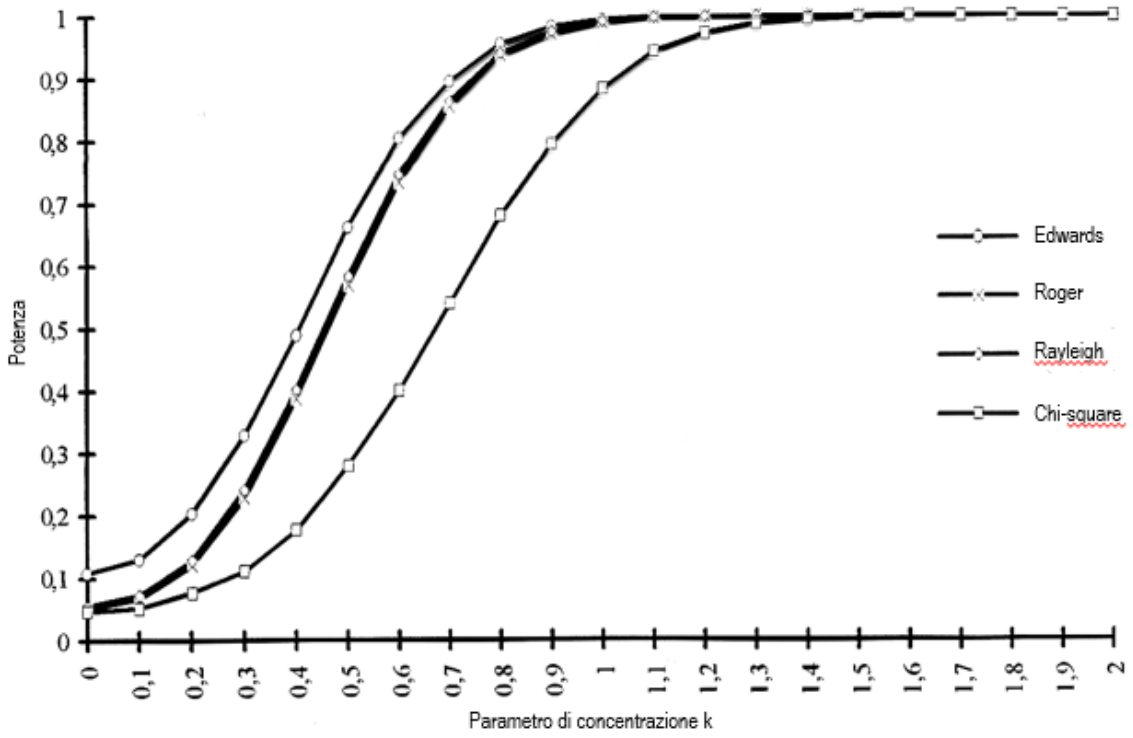
Allo scopo di studiare gli errori di I tipo si è prodotta una generazione di numeri casuali provenienti da una distribuzione circolare uniforme con metodi Montecarlo; sono stati generati 10000 campioni, di ampiezze campionarie da 10 a 500 (partendo da n=10, poi a step di 50 da n=50 a n=500) ed eseguendo i quattro test a livelli di significatività α=0.01, 0.05, 0.10; nel caso del test χ^2 i dati sono stati organizzati in 12 classi di ampiezza $\pi/6$.
Allo scopo di studiare gli errori di II tipo e di stimare le curve di potenza per i test in questione si è prodotta una generazione di numeri casuali mediante metodi Montecarlo (metodo acceptance-rejection di Von Neumann) provenienti da: i. una distribuzione di von Mises al variare del parametro di concentrazione k; ii. una mistura di due distribuzioni di von Mises al variare dei parametri k_1, k_2, θ_1, θ_2, p, tenendo conto che per

k=0 la von Mises diventa una distribuzione circolare uniforme). Per ogni situazione sono stati generati 10000 campioni di ampiezze campionarie da 10 a 500 (partendo da n=10, poi a step di 50 da n=50 a n=500) ed eseguendo poi i quattro test a livelli di significatività α=0.05; nel caso del test χ^2 i dati sono stati organizzati in 12 classi di ampiezza $\pi/6$.

Risultati

I risultati mostrano una evidente inadeguatezza del test di Edwards in termini di errore di I tipo e del test χ^2 in termini di potenza. Al contrario, i test di Rayleigh e di Roger mostrano una rispondenza adeguata alle attese e una sostanziale sovrapponibilità di affidabilità. Tuttavia, nel caso di misture di distribuzioni, appare evidente la perdita di potenza dei test di Edwards, Rayleigh e Roger, tanto più evidente quanto più le mode risultano con densità simili e posizionate a distanza prossima a π. In questo caso il test χ^2 mostra potenza superiore rispetto agli altri test.

__Figura 1__. Il grafico mostra la potenza dei quattro test al variare del parametro di concentrazione k, in 10000 campioni di dimensione n=50 generati con metodo Montecarlo da una distribuzione von Mises.



Conclusioni

Dalle analisi condotte emergono alcune evidenze: i. grossolani limiti del test di Edwards, che andrebbe espunto dai metodi statistici per problemi di identificazione di trend ciclici; ii. sostanziale equivalenza dei test di Rayleigh e di Roger in presenza di distribuzioni unimodali e loro inadeguatezza in caso di bimodalità della distribuzione (specialmente in caso di mode simmetriche attorno all'origine nel grafico polare), nel qual caso diventa vantaggioso l'impiego del test χ^2 (peraltro non si è qui considerato un importante fattore, rappresentato dal numero di classi e dalla loro conseguente ampiezza).

Da queste evidenze emerge una strategia che si può riassumere nei punti seguenti: i. Analisi esplorativa mediante istogramma circolare o, più efficacemente, mediante stima della funzione di densità di probabilità con metodi kernel [4]; ii. test per la goodness-of-fit rispetto a una Von Mises [5] o altra distribuzione circolare; iii. eventuale scomposizione delle misture [6, 7]; esecuzione del test di Rayleigh o di Roger.

Bibliografia

[1] Rayleigh *Lord*, On the problem of random vibrations, and of random flights in one, two or three dimensions. Phil Mag., 1919; 37: 321-47.

[2] Edwards J.H., The recognition and estimation of cyclic trends. Ann Hum Genet., 1961; 25: 83-7.

[3] Roger J.H., A significance test for cyclic trends in incidence data. Biometrika, 1977; 64(1): 152-5.

[4] Fisher N.I., Smoothing a sample of circular data. J Struct Geol., 1989; 11: 775-8.

[5] Lockhart R.A., Stephens M.A., Tests of fit for the Von Mises distribution. Biometrika, 1985; 72: 647-52.

[6] Jones T.A., James W.R., Analysis of bimodal orientation data. Math Geol., 1969; 1: 129-35.

[7] Mardia K.V., Sutton T.W., On the modes of a mixture of two von Mises distributions. Biometrika, 1975; 62: 699-701.

UN APPROCCIO NON PARAMETRICO ALL'ANALISI DELLA VARIANZA A DUE VIE

Borrelli Paola[1], Gigli Berzolari Francesca[1], Pistono Cristiana[2], Mallucci Giulia[3], Monti Maria Cristina[1]

[1] *UO di Biostatistica e Epidemiologia Clinica, Dipartimento di Sanità Pubblica, Medicina Sperimentale e Forense, Università degli Studi di Pavia, Pavia*

[2] *Labortorio di Immunogenetica, Dipartimento di Biologia e Biotecnologie "L. Spallanzani", Università degli Studi di Pavia, Pavia*

[3] *Centro Interdipartimentale Sclerosi Multipla, Fondazione Istituto Neurologico Nazionale "C. Mondino", Pavia*

Introduzione

La validità di un'analisi della varianza passa attraverso la verifica di alcuni assunti: la distribuzione di probabilità della variabile di endpoint nella popolazione obiettivo deve essere normale, le varianze delle popolazioni a confronto devono essere omogenee, il campionamento deve essere casuale e, infine, la variabile di endpoint deve essere misurata almeno in una scala ad intervallo. Quando uno o più di tali assunti non sono rispettati si verifica un aumento degli errori di primo e secondo tipo, con conseguente calo del livello di protezione e della potenza del test di significatività; in tal caso l'approccio non parametrico fornisce risultati più attendibili e la potenza dei test aumenta in relazione all'entità della violazione degli assunti. Tuttavia non esiste un'alternativa non parametrica per condurre un'analisi della varianza a due vie, se non nel modello a blocchi randomizzati completi [1]. Leys et al. propongono una tecnica alternativa che permette di verificare l'effetto dell'interazione, perdendo tuttavia l'informazione sugli effetti principali; inoltre tale tecnica prevede l'utilizzo dell'analisi della varianza parametrica con variabili misurate su una scala ordinale [2].

Obiettivi

Costruire un'alternativa semplice all'analisi della varianza parametrica a due vie quando gli assunti di applicabilità non sono verificati o quando la variabile di endpoint non rispetta i requisiti di una scala parametrica.

Metodi

Un approccio grafico basato sui valori mediani della variabile di endpoint, accompagnato da test non parametrici per la valutazione degli effetti principali e dell'interazione, è in grado di fornire risultati validi e attendibili.
Tale approccio può essere utilizzato nei disegni fattoriali, per misure ripetute e misti. Indipendentemente dal disegno, il metodo consiste nella valutazione degli effetti principali dei due fattori in studio attraverso test non parametrici per popolazioni indipendenti o per popolazioni dipendenti a seconda che il disegno sia fattoriale o per misure ripetute, rispettivamente. La valutazione dell'interazione tra i due fattori può essere effettuata inizialmente con un semplice approccio grafico, seguito, eventualmente, da opportuni test non parametrici che permettono di stabilire se la differenza tra le mediane della variabile di endpoint nei livelli di un fattore differisce significativamente nei livelli dell'altro fattore.
Si supponga di voler valutare l'effetto di due fattori dicotomici A e B sulla variabile di endpoint y misurata in una scala ordinale o ad intervallo (Figura 1, grafico 1).
Per prima cosa si procede alla valutazione degli effetti principali di ogni singolo fattore: l'effetto principale di un fattore (ad esempio il fattore A) è l'effetto che tale fattore ha sulla variabile di endpoint y valutato senza prendere in considerazione l'altro fattore (B); di conseguenza la valutazione di tali effetti può essere effettuata confrontando tra loro i valori mediani della variabile y nei livelli di un fattore (ad esempio A)

indipendentemente dall'altro (B). Per tali confronti si possono utilizzare test non parametrici per popolazioni indipendenti (ad esempio Mann-Withney/Kruskal Wallis test) nei disegni fattoriali o per popolazioni dipendenti (ad esempio test dei segni o di Wilcoxon/Friedman) nei disegni per misure ripetute. Successivamente si procede alla valutazione dell'interazione tra i due fattori: si è in presenza di interazione quando la differenza tra le mediane della variabile di endpoint y nei livelli di un fattore (ad esempio B) differisce significativamente nei livelli dell'altro fattore (A); nel grafico 1 di Figura 1 A1B1, A1B2, A2B1, A2B2 rappresentano i valori mediani della variabile y nei livelli dei fattori A e B oggetto di studio; dalla forma del grafico è ragionevole ipotizzare l'effetto di una interazione tra i due fattori, poiché la differenza tra i valori mediani di y in B1 e B2 del fattore B differisce nei livelli A1 e A2 del fattore A.
Per capire se tale interazione è significativa e quindi attribuibile al caso con una bassa probabilità, si può procedere con due confronti:

1) la mediana A1B1vs la mediana A1B2
2) la mediana A2B1 vs la mediana A2B2

Sono possibili tre scenari:

1. sia in A1, sia in A2 la differenza tra le mediane risulta non significativa, quindi compatibile con il caso; di conseguenza si ha assenza di interazione e si procede quindi alla valutazione dei soli effetti principali;
2. in A1 la differenza tra le mediane risulta non significativa, quindi compatibile con il caso, mentre in A2 risulta statisticamente significativa; di conseguenza si è in presenza di interazione tra i due fattori: la differenza tra le mediane di y in A1 non è costante in A2;
3. sia in A1, sia in A2 la differenza tra le mediane risulta significativa; poiché la differenza in A2 è stimata essere maggiore rispetto a quella osservata in A1, è ragionevole ipotizzare la presenza di interazione tra i fattori.

Risultati

Un esempio di applicazione del metodo.
La sclerosi multipla (SM) è una patologia multifattoriale e tra i meccanismi di suscettibilità vi è lo stress ossidativo. Si è voluto investigare *ex vivo* l'influenza dello stress ossidativo sulle cellule mononucleate del sangue periferico, confrontando pazienti affetti da SM (casi) con un gruppo di controllo (controlli). A tale scopo le cellule sono state trattate con uno stress ossidativo (H_2O_2) [3] [4]; è stata quindi valutata la loro attività mitocondriale attraverso un dosaggio colorimetrico (MTT) dopo15 minuti e dopo tre ore. Sono stati analizzati 53 pazienti affetti da SM e 45 controlli. Per l'analisi è stato utilizzato un approccio non parametrico, studiando la relazione tra i livelli della variabile di endpoint "MTT" e i fattori gruppo (caso/controllo) e tempo (basale, 15 minuti, 3 ore); è corretta l'applicazione di un modello misto di analisi della varianza (il gruppo è la componente fattoriale e il tempo è la componente ripetuta).
Seguendo le linee guida del metodo sopra descritto, si è proceduto all'analisi degli effetti principali, utilizzando il test di Mann-Whitney per il confronto tra casi e controlli e il test di Friedman per il confronto tra i tempi (Figura 1, grafico 2). Si sono ottenuti i seguenti risultati:

- l'effetto principale del gruppo non è risultato significativo ($mediana_{casi}$=177,00 IQR:138,17-212,00 $mediana_{controlli}$=187,00 IQR:132,00-232,67 U=1171,00 p=0,880).
- L'effetto principale del tempo è risultato statisticamente significativo ($mediana_{basale}$=182,00 IQR: 145,75-218,75 $mediana_{15min}$=171,50 IQR: 127,00-214,25 $mediana_{3h}$=186,75 IQR:134,00-236.25 chi^2=16,88 p<0,001). L'analisi a posteriori, effettuata utilizzando il test dei segni con

correzione per confronti multipli, ha evidenziato una riduzione statisticamente significativa dei valori mediani tra il livello "basale" e il livello "dopo 15 minuti" ($p<0,001$) e tra il livello "dopo15 minuti" e il livello "dopo 3 ore" ($p=0,005$); nessuna differenza significativa è risultata nel confronto tra il livello "basale" e il livello "dopo 3 ore" ($p=0,077$).

- Per quel che riguarda l'effetto dell'interazione, l'approccio grafico (grafico 2 di Figura 1) suggerisce la presenza di una possibile interazione tra il fattore gruppo e il fattore tempo poiché la differenza tra le mediane tra casi e controlli non si mantiene costante nel tempo.

Pertanto si è proceduto a verificare la significatività statistica delle differenze tra casi e controlli nei livelli del tempo; a tale scopo è stato utilizzato il test di Mann-Whitney, ottenendo i seguenti risultati:

- ✓ in basale: $U=1138,00$ $p=0,700$
- ✓ dopo 15 minuti: $U=1181,00$ $p=0,936$
- ✓ dopo 3 ore: $U=1185$ $p=0,959$

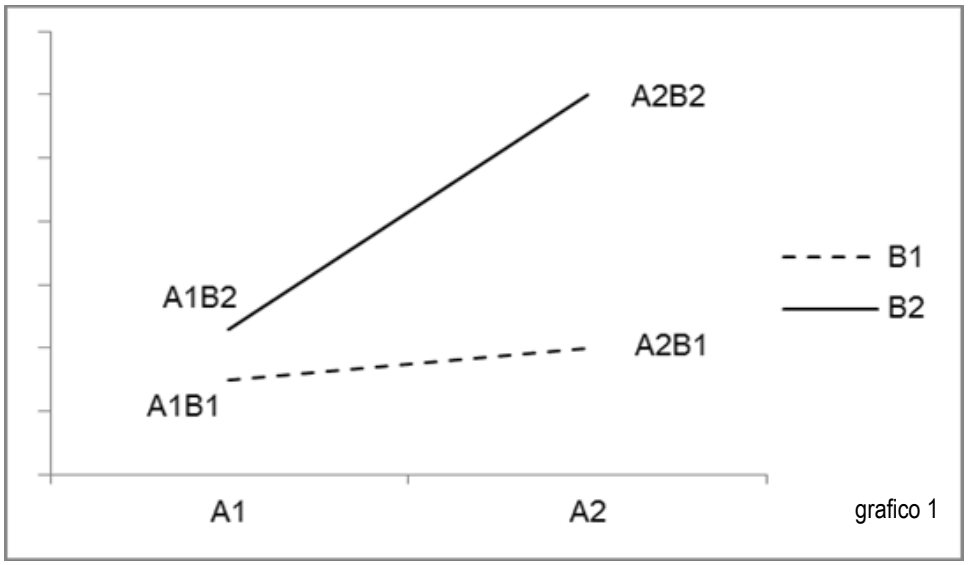
Dai risultati si evince che, non essendo risultata significativa alcuna differenza, è possibile escludere la presenza di interazione tra il fattore tempo e il fattore gruppo; pertanto si procede alla valutazione dei soli effetti principali dei due fattori in studio:

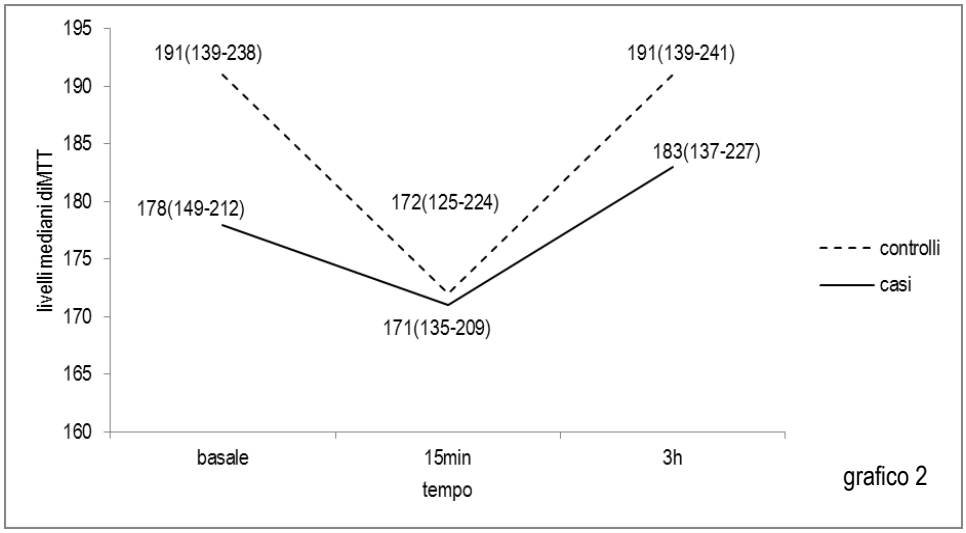
- indipendentemente dal fattore tempo, la risposta nel tempo allo stress ossidativo dei casi è mediamente uguale a quella dei controlli,
- indipendentemente dal fattore gruppo, si osserva una variabilità nel tempo statisticamente significativa: i livelli mediani di MTT decrescono dopo 15 minuti di stress ossidativo, passando da 182,00 a 171,50; dopo 3 ore si osserva un interessante ripristino dei livelli, passando da 171,50 a 186,75.

Conclusioni

Si ritiene che il metodo illustrato possa essere una valida alternativa all'analoga tecnica parametrica, poiché semplice da realizzare, attendibile nei risultati (in modo particolare con casistiche non troppo ridotte), applicabile a diversi disegni di analisi della varianza e alla valutazione di effetti che graficamente presentano andamenti differenti.
In relazione all'applicazione sui dati SM, il metodo permette di concludere che le cellule mononucleate del sangue periferico mostrano una rilevante variazione della loro funzione mitocondriale, misurata attraverso un dosaggio colorimetrico, sia nei soggetti affetti da sclerosi multipla, sia nei controlli, considerando diversi tempi di esposizione. In particolare, dopo 15 minuti dallo stress ossidativo l'attività mitocondriale decresce e viene ripristinata dopo tre ore dall'esposizione allo stress. Il fatto che non sia risultata alcuna differenza significativa tra casi e controlli suggerisce che le cellule mononucleate del sangue periferico di entrambi i gruppi inizialmente (dopo 15 minuti) riducono la loro attività mitocondriale, ma sono in grado di ripristinare i loro livelli fisiologici in un periodo di tempo più lungo (3 ore).

Figura 1





Bibliografia

[1] Leys C., Schuman S., A nonparametric method to analyze interactions: the adjusted rank transform test. Journal of Experimental Social Psychology, 2010; 46(4):684-688

[2] Sokal R.R.,Rohlf F. J., Biometry,1981; second edition: 445-450

[3] Li M, Zhao L, Liu J, et al, Hydrogen peroxide induces G2 cell cycle arrest and inhibits cell proliferation in osteoblasts. Anat Rec (Hoboken),2009; 292:1107–13

[4] Nogueira-Pedro A, Cesário T., Dias C. et al, Hydrogen peroxide (H2O2) induces leukemic but not normal hematopoietic cell death in a dose-dependent manner. Cancer Cell Int, (2013); 13:123

DAILY PATTERNS OF CARBOHYDRATE INTAKE IN ADULTS BY NON-PARAMETRIC MULTI-LEVEL LATENT CLASS ANALYSIS: AN APPLICATION TO A REPRESENTATIVE SAMPLE OF BRITISH ADULTS FROM THE UK NATIONAL DIET AND NUTRITION SURVEY (2008/09 - 2015/16)

Chaochen Wang[1,2], Suzana Almoosawi[3], Luigi Palla[1]

[1] *Department of Medical Statistics, London School of Hygiene & Tropical Medicine, London, UK;*

[2] *Department of Public Health, Aichi Medical University, Nagakute, Aichi, JAPAN;*

[3] *Brain, Performance and Nutrition Research Centre, Northumbria University, Newcastle, UK*

Introduction

Timing of eating is believed to affect a wide variety of physiological processes as well as health outcomes. Recent evidence on adults suggests that there are three types of eaters (grazers, early eaters and late eaters) according to the timing of energy intake. However, to date challenges remain on how diurnal eating patterns could be derived and their association with health outcomes analyzed.

Objectives

This study aims at combining time and quantity of carbohydrate (CH) intake in the definition of eating patterns in UK adults. And to take into account in the derivation of such patterns the multi-level nature of the data, i.e.the repeated measures of daily diet diaries for each individual in a national nutrition survey, representative of the British population.

Methods

Data from the National Diet and Nutrition Survey (NDNS) Rolling Programme (2008/09-15/16) included 6155 aged 19 or older, a representative cross-sectional sample of the adult British population taken on a rolling basis for 8 years between 2008 and 2016. Time of the day was categorized into 7 slots: 6-9 am, 9-12 noon, 12-2 pm, 2-5 pm, 5-8 pm, 8-10 pm and 10 pm-6 am. Responses for carbohydrate intake within each time slot were categorized into: (i) no energy intake, (ii) CH contributed $<$ 50% or (iii) CH contributed $\geq$ 50% of total energy at that time slot. Non-parametric multilevel latent class analysis (MLCA) models were applied to identify latent classes of CH consumption, accounting for the repeated measurement of intake on 3-4 days nested within individuals: the multilevel procedure would identify patterns of carbohydrate intake across the time-slots within a day (day pattern) and reclassify survey participants according to their probability to follow each day pattern. Survey-designed multivariable regression models were then used to assess the associations of CH eating patterns with hypertension and obesity.

Results

Three CH eating day patterns (named as low percentage, high percentage and regular meal day) emerged from 24483 observation days; based on these three types of days, CH eaters were defined which could be broadly labelled as low (28.1%), moderate (28.8%) and high (43.1%) CH eaters. On average, low-CH eaters consumed the highest amount of total energy intake (7985.8 kJ, p-value $<$ 0.001), and they had higher percentages of energy contributed by fat and alcohol, especially after 8 pm. Moderate-CH eaters consumed the lowest amount of total energy (7341.8 kJ) while they had the tendency of eating meals later in the day. High-CH eaters consumed most of their carbohydrates and energy within the time slots of 6-9 am, 12-2 pm and 5-8 pm. More details with average values of nutrients (also by time-slot) for different types of CH eaters

are presented in Table 1. In relation to health outcomes, low-CH eaters had higher BMI, larger WC and higher prevalence of diabetes and hypertension compared to moderate and high CH eaters.

Table 1. *Weighted means and percentages (95% CI) of the nutrients intake according to individual level carbohydrate eating classes.*

(NDNS RP 2008/09-15/16, sample size = 6155)

	Low CH eaters (n = 1730)	**Moderate CH eaters (n = 1772)**	**High CH eaters (n = 2653)**	**P value**
Total energy intake (kJ)	7985.8 (7823.3, 8146.295)	7341.8 (7825.3, 8146.3)	7677.8 (7555.8, 7799,8)	< 0.001
Carbohydrate intake (g)	203.8 (199.8, 207.8)	218.3 (212.9, 223.7)	233.4 (229.6, 237.2)	< 0.001
6 am – 9 am	23.0 (21.8, 24.3)	11.2 (10.0, 12.3)	37.9 (36.8, 39.2)	
Fiber (g)	1.4 (1.3, 1.5)	0.6 (0.5, 0.7)	2.0 (1.9, 2.2)	
Sugar (g)	10.2 (9.6, 10.9)	5.3 (4.8, 5.8)	19.7 (19.0, 20.4)	
Non-milk extrinsic sugar (g)	4.7 (4.3, 5.1)	3.2 (2.9, 3.6)	11.1 (10.7, 11.6)	
Starch (g)	12.8 (12.0, 13.5)	5.9 (5.1, 6.6)	18.3 (17.6, 19.1)	
9 am – 12 noon	25.1 (23.9, 26.3)	33.0 (31.4, 34.6)	26.2 (25.1, 27.2)	
Fiber (g)	1.5 (1.4, 1.6)	1.6 (1.5, 1.7)	1.3 (1.2, 1.3)	
Sugar (g)	11.6 (10.9, 12.3)	15.7 (14.8, 16.6)	14.2 (13.6, 14.8)	
Non-milk extrinsic sugar (g)	5.7 (5.2, 6.2)	9.6 (8.9, 10.2)	8.1 (7.7, 8.5)	
Starch (g)	13.5 (12.8, 14.3)	17.3 (16.4, 18.3)	11.9 (11.3, 12.6)	
12 noon – 2 pm	42.6 (40.9, 44.3)	38.7 (37.0, 40.4)	51.6 (50.2, 52.9)	
Fiber (g)	3.1 (2.9, 3.2)	2.3 (2.2, 2.5)	3.6 (3.5, 3.7)	
Sugar (g)	14.7 (14.0, 15.4)	14.9 (14.0, 15.7)	19.4 (18.7, 20.0)	
Non-milk extrinsic sugar (g)	7.3 (6.7, 7.8)	9.1 (8.4, 9.8)	10.3 (9.8, 10.8)	
2 pm – 5 pm	25.0 (23.6, 26.4)	33.6 (31.6, 35.6)	24.7 (23.6, 25.7)	
Fiber (g)	1.6 (1.5, 1.7)	1.9 (1.7, 2.0)	1.3 (1.2, 1.4)	
Sugar (g)	11.9 (11.3, 12.7)	14.5 (13.5, 15.5)	13.4 (12.8, 13.9)	
Non-milk extrinsic sugar (g)	6.9 (6.4, 7.5)	9.9 (9.0, 8.6)	8.6 (8.2, 9.1)	
Starch (g)	13.1 (12.1, 13.9)	19.1 (17.7, 20.4)	11.3 (10.6, 11.9)	
5 pm – 8 pm	55.9 (54.1, 57.9)	54.6 (52.1, 57.0)	62.9 (61.3, 64.4)	
Fiber (g)	4.4 (4.2, 4.5)	3.7 (3.5, 3.9)	4.9 (4.7, 5.0)	

Sugar (g)	18.7 (17.9, 19.5)	18.6 (17.6, 19.5)	21.8 (20.9, 22.5)	
Non-milk extrinsic sugar (g)	10.2 (9.6, 10.8)	11.8 (10.9, 12.6)	12.1 (11.4, 12.7)	
Starch (g)	37.3 (35.8, 38.8)	35.9 (34.1, 37.9)	41.1 (39.9, 42.2)	
8 pm – 10 pm	23.3 (21.9, 24.6)	29.7 (27.6, 31.7)	23.7 (22.5, 24.9)	
Fiber (g)	1.4 (1.3, 1.6)	1.6 (1.5, 1.8)	1.3 (1.5, 1.8)	
Sugar (g)	10.9 (10.3, 11.5)	13.2 (12.2, 14.2)	12.4 (11.8, 13.0)	
Non-milk extrinsic sugar (g)	7.3 (6.8, 7.8)	9.4 (8.5, 10.4)	8.3 (7.8, 8.8)	
Starch (g)	12.3 (11.4, 13.3)	16.4 (15.0, 17.8)	11.3 (10.5, 12.1)	
10 pm – 6 am	8.8 (7.7, 9.8)	17.6 (15.2, 19.9)	6.4 (5.8, 7.1)	
Fiber (g)	0.34 (0.29, 0.39)	0.74 (0.63, 0.85)	0.24 (0.21, 0.27)	
Sugar (g)	5.3 (4.6, 6.1)	10.0 (8.6, 11.5)	4.1 (3.7, 4.5)	
Non-milk extrinsic sugar (g)	3.9 (3.3, 4.6)	7.7 (6.4, 8.9)	2.9 (2.6, 3.3)	
Starch (g)	3.5 (2.9, 3.9)	7.5 (6.3, 8.8)	2.3 (1.9, 2.7)	
Carbohydrate percent (%)	40.6 (40.2, 41.0)	47.3 (46.8, 47.8)	48.3 (47.9, 48.6)	< 0.001
Total fiber (g)	13.7 (13.4, 14.0)	12.5 (12.1, 12.9)	14.7g (14.4, 14.9)	< 0.001
Protein intake (g)	79.9 (77.9, 81.8)	69.3 (67.6, 71.0)	73.7 (72.5, 74.8)	< 0.001
Protein percent (%)	17.2 (16.9, 17.5)	16.3 (16.0, 16.6)	16.5 (16.3, 16.6)	< 0.001
Fat intake (g)	74.7 (73.1, 76.4)	63.8 (62.1, 65.5)	65.7 (64.4, 67.0)	< 0.001
Fat percent (%)	35.4 (34.9, 35.8)	32.5 (32.1, 32.9)	32.0 (31.7, 32.3)	< 0.001
P values were obtained from chi-square test for categorical variables, and one-way ANOVA comparing the means in multiple groups for continuous variables; CH, carbohydrate. Non-milk extrinsic sugar is defined as: additionally added free sugar, such as table sugar, honey, glucose, fructose and glucose syrups, sugars added to food and sugars in fruit juices				

Discussion

Whether diurnal variation in their intake of energy, fat and alcohol may have contributed to this higher prevalence or may be the result of reverse causation will need to be addressed by longitudinal studies. Some partial statistical and epidemiological limitations of this study include: ignoring the order of the observation days within individuals (exchangeability of repeated measures) and the probabilistic nature of the MLCA classification in subsequent regression analyses; the cross-sectional nature of the study design; lack of information about shift work and the difficulty in estimating underreporting of survey participants. Strengths include the ability of the data-driven procedure to find subgroups within a complex structure in a representative large survey sample of British adults, with a latent class method recently proposed and available in M-plus.

Conclusions

The high-CH eaters profile seemed to be the healthiest. Low-CH eaters were associated with higher prevalence of obesity which may have resulted from health/weight concerns, leading to greater consumption of fat or alcohol as replacements for CH. To ascertain the direction of causality in the association of CH patterns with blood pressure and obesity, prospective longitudinal studies are warranted.

References

[1] Leech, R.M.; Worsley, A.; Timperio, A. et al., Diurnal eating patterns: a latent class analysis approach. Int J Behav Nutr Phys Act, 2017. 14(1): p. 3.

[2] Mansukhani R., Palla L. Investigating eating time patterns in UK adults from the 2008-2012 NDNS. Proc Nut Soc, 2018, 77 (OCE1)

MEASURING MULTIMORBIDITY INEQUALITY ACROSS ITALY THROUGH THE MULTISOURCE COMORBIDITY SCORE: A NATIONWIDE STUDY

Rea Federico[1,2], Cantarutti Anna[1,2], Merlino Luca[1,3], Zanier Loris[4], Di Lenarda Andrea[5], De Palma Rossana[6], Belotti Laura[6], Francesconi Paolo[7], Profili Francesco[7], Forni Silvia[7], Carle Flavia[1,8], Bucci Andrea[1,8], Davoli Marina[9], Di Martino Mirko[9], Lallo Adele[9], Lepore Vito[10], Martini Nello[11], Dondi Letizia[11], Scondotto Salvatore[1,12], Pollina Addario Sebastiano[1,12], Fantaci Giovanna[1,12], Garau Donatella[13], Sechi Giuseppe[14], Lispi Lucia[15], Carbone Simona[15], Corrao Giovanni[1,2] and on behalf of the working group 'Monitoring and assessing diagnostic therapeutic paths' of the Italian Health Ministry.

[1] *National Centre for Healthcare Research and Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy*

[2] *Laboratory of Healthcare Research and Pharmacoepidemiology, Unit of Biostatistics, Epidemiology and Public Health, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy*

[3] *Head, Epidemiologic Observatory, Regional Welfare Service, Lombardy Region, Milan, Italy*

[4] *Epidemiological Service, Azienda Regionale di Coordinamento per la Salute (ARCS), Udine, Italy*

[5] *Cardiovascular Center, University Hospital and Health Services, Trieste, Italy*

[6] *Authority for Healthcare and Welfare, Emilia Romagna Regional Health Service, Bologna, Italy*

[7] *Regional Health Agency of Tuscany, Florence, Italy*

[8] *Center of Epidemiology and Biostatistics, Polytechnic University of Marche, Ancona, Italy*

[9] *Department of Epidemiology, Lazio Regional Health Service, Roma, Italy*

[10] *Regional Healthcare Agency of Puglia Region, Bari, Italy*

[11] *ReS (Research & Health) Foundation, Rome, Italy*

[12] *Epidemiologic Observatory, Sicily Regional Health Service, Palermo, Italy*

[13] *Regional Councillorship of Health, "Regione Autonoma della Sardegna" Cagliari, Italy*

[14] *Department of Clinical Governance, Sardinia Region Health Service, Cagliari, Italy*

[15] *Department of Health Planning, Italian Health Ministry, Rome, Italy*

Introduction

Measuring inequality in the prevalence of chronic diseases is requested for proper resources allocation, particularly in systems based on National Health Service (NHS).
A simple Multisource Comorbidity Score (MCS) combining population-based administrative data on healthcare utilization from the NHS has been recently developed and validated in Italy [1]. As MCS showed an impressive stability of between-region performance in predicting clinical outcomes among beneficiaries of the NHS, other than for risk adjustment in clinical and epidemiological studies, it can represent a useful tool for health policy planning.

Aim

A very large real-world investigation covering 63% of the Italian population aged 50-85 years was conducted under the auspices of the Italian Health Ministry with the aim of measuring between-region differences in the distribution of the multimorbidity through the MCS

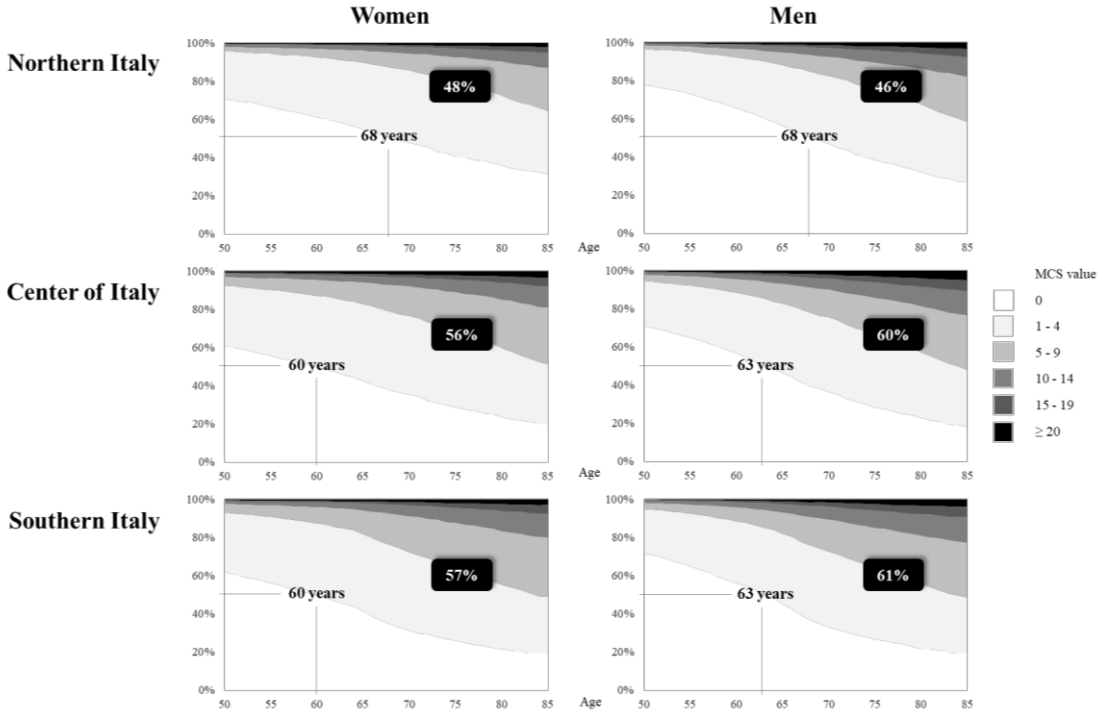
Methods

This study is based on computerized healthcare utilization databases from ten of the twenty Italian regions, which voluntarily adhered to the initiative. Three regions are located in the northern (Lombardy, Friuli-Venezia-Giulia and Emilia-Romagna), three in central (Marche, Tuscany and Lazio) and four in the southern Italy (Puglia, Basilicata, Sicily and Sardinia). Overall, these data covered more 38 million beneficiaries of the Italian NHS, just under two out of three of the entire Italian population.
Beneficiaries of the NHS who in the index year (i.e., 2017) were aged from 50 years to 85 years, and were resident in one of the ten participant Italian regions from at least two years (i.e., were recorded as beneficiaries of the regional NHS before the year 2015), formed the study population.
For each cohort member, MCS was calculated according to their previous contacts with the NHS. Individuals were classified into 1 of the 6 categories denoting a progressive worsening comorbidity status (i.e., with values 0, 1-4, 5-9, 10-14, 15-19 and ≥ 20).
With the aim of comparing the MCS distribution across women and men belonging to geographic aggregates (i.e., between north, centre and south Italy), we proceeded as follows. The percentage of beneficiaries belonging to the MCS categories was calculated for each individual age, and the corresponding distributions stratified for gender and geographic aggregates were represented. Each distribution, which was denoted as age-specific MCS distribution, consisted in a rectangle (with the 35 ages as base and 1 as height) within which the area was portioned into sub-areas with different shades of grey which gradually become darker according to progressively increasing values of the MCS.
Three approaches were used for synthesizing and comparing the age-specific MCS distributions. One, the age above which half of the beneficiaries suffered from at least one comorbidity was calculated and denoted as MCS median age. Two, the portion of the rectangle area covered by beneficiaries suffering from at least one comorbidity was also calculated and denoted as percentage of comorbid population. Three, the ridit transformation of the MCS was finally performed [2]. For each individual age, we ridit transformed the MCS so to compare the comorbidity profile for NHS beneficiaries from a given geographic aggregate with respect to NHS beneficiaries of the entire Italian population. A ridit score with value lower or higher of 0.5 is obtained respectively according to whether better or worse comorbidity profile characterized the population from a given geographical aggregate with respect to the entire Italian population.

Results

Figure 1 shows that MCS median age ranged from 60 years (center and south) to 68 years (north) in women and from 63 years (center and south) to 68 years (north) in men. The percentage of comorbid population was lower than 50% for northern population, while it was around 60% for central and southern populations. Ridit analysis displayed that, compared with NHS beneficiaries from central and southern Italy, those from northern Italy always showed better comorbidity profile for both women and men. Men, and especially women, aged 65 years or older from southern Italy showed the worse comorbidity profile.

Figure 1. MCS median age and percentage of comorbid population according to gender and geographic aggregate. Italy, 2017



Conclusions

MCS may be easily derived from data usually used for health system management, and has relevant value for conducting population-based investigations aimed to compare multimorbidity prevalence and evaluating health policy planning.

References

[1] Corrao G, Rea F, Di Martino M, et al., Developing and validating a novel multisource comorbidity score from administrative data: a large population-based cohort study from Italy. BMJ Open., 2017 Dec 26;7(12): e019503

[2] Donaldson GW. Ridit scores for analysis and interpretation of ordinal pain data. Eur J Pain., 1998;2(3):221-7

IDENTIFICATION OF GENE EXPRESSION IN THE BRAIN CAUSALLY LINKED TO MULTIPLE SCLEROSIS: A TWO-SAMPLE MENDELIAN RANDOMIZATION APPROACH

Nova Andrea[1], Fazia Teresa[1], Gentilini Davide [1], Adaikalavan Ramasamy[2], Berzuini Carlo[3], Bernardinelli Luisa[1]

[1] *Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy.*

[2] *Department of Medical and Molecular Genetics, King's College London, Guy's Hospital, London, UK*

[3] *Centre for Biostatistics, University of Manchester, Manchester, United Kingdom*

Introduzione

MS is a multifactorial disease characterized by chronic inflammation and demyelination in the Central Nervous System (CNS). Although its etiology remains unknown, MS is believed to result from an intricate interplay between inflammatory effector cells and CNS resident cells [1]. A large number of MS-associated genes have been found in the recent genome-wide association studies (GWASs), fine-mapping and meta-analysis [2-3]. Multiple NF-κB polymorphisms have been associated with increased susceptibility to MS [4]. Variants have been identified within the NF-κB signaling pathway in subjects with autoimmune diseases through genome-wide association studies. In MS patients, variants near genes involved in NF-κB signaling have been found and are functional, potentially due to location in regulatory elements. One such variant, proximal to NFKB1, increased expression by twenty-fold [5] NF-κB signaling plays a central role in the activation of both peripheral inflammatory cells and the CNS resident glial cells that ultimately mediate inflammatory demyelination. Evidence suggests that NF-κB activation in inflammatory cells, including T cells, microglia/macrophages and astrocytes, enhances inflammation and promotes the development of MS and Experimental Autoimmune Encephalomyelitis (EAE). Conversely, data indicate that NF-κB activation in oligodendrocytes and neurons protects these cells against inflammation in MS and EAE. NF-κB exerts different roles in different cell types in MS and EAE, which are paradoxical and even conflicting [6]. Evidently, NF-κB plays a key role in MS development, through many different signaling pathways. Each signaling pathway has many influencers, some of which act in many of the NF-κB signaling pathways, and others that are limited to only a few NF-κB signaling pathways. Whether direct inhibition of NF-κB is the best mechanism, it still remains to be elucidated. Being MS a multifactorial disease not dependent on only one gene, or inciting insult, it is possible that there are a variety of aberrant genes present in MS patients that decrease the functioning of any of the many kinases and molecules implicated in the NF-κB signaling pathway [5].

Obiettivi

Most of the GWAS-highlighted loci that appear to operate via gene expression regulation, rather than protein coding, spurred the researchers to perform expression quantitative trait locus (eQTL) analyses, a key tool for the functional understanding of GWASs results, that highlighted genes whose expression in specific tissues is regulated by loci associated with the disease [7]. The aim of this study is to asses if gene expression in brain tissue could be causally related to MS, focusing on genes involved in the NF-kB pathway. The assessment of possible causal links between gene expression and disease is based on a class of statistical methods known as Mendelian Randomization (MR), since these methods will help to eliminate the risk of estimate causal effects biased by confounding and/or reverse causation [8].

Metodi

MR uses measured variation in DNA sequence, Z, to assess the possible causal effect of a biological phenotype, X, on a response variable, Y, without requiring any experimental intervention on X. The variable Z is called instrumental variable (IV) and it is an instrument for the estimation of the effect of interest. We applied the method to a restricted selected a priori set of genes identified as plausible MS risk genes on the basis of previous, independent studies. In particular genes were selected on the basis of canonical and non-canonical NF-kB pathway, analysing receptors (i.e TNFRS1A), protein kinases (i.e MAP3K7), inhibitor IKBα, NF-kB transcription complex subunits (i.e p50), transcription products and inductors (i.e TNF, ICAM1), for a total of 21 genes.
This study is based on a two-samples MR, where the genetic associations with outcome and exposure were estimated respectively in two independent and non-overlapping samples: Sardinian multiplex families and data provided by the UK Brain Expression Consortium (UKBEC) [9-10]. Sardinia data provided the summary statistics of association (regression coefficients and standard error) between the IVs and MS, obtained in a previous study analysing Immunochip genotyping data from 15 multiplex Sardinian families (75 affected and 254 unaffected members), plus 94 unrelated healthy subjects from the same population [9]. UKBEC data provided Immunochip genotyping and gene expression measurements from a collection of autoptic tissue samples of ten brain regions (cerebellar cortex, frontal cortex, hippocampus, medulla specifically inferior olivary nucleus, occipital cortex specifically primary visual cortex, putamen, substantia nigra, temporal cortex, thalamus and intralobular white matter) in a sample of 134 neuropathological confirmed control individuals of European descent. UKBEC data was used to obtain the required instrument-exposure association summary statistics, identifying for each gene the genetic variant having a significant ($p<1x10-3$) effect in the linear regression of the transcript-level expression profile on the genetic variant to be used a IVs for MR analysis. Heterogeneous instruments were removed using Radial MR with a significance level $\alpha=0.25$ to avoid potential pleiotropy and invalid instruments. The MR analyses were performed using the MendelianRandomization R package, that reported the results of MR using the Inverse-Variance Weighted estimator (IVW). Sensitivity analysis were conducted using different significant levels when selecting IVs to include more instruments ($p<1.25x10-3$ and $p<1.5x10-3$), a lower significance level to remove heterogeneous instruments ($\alpha=0.10$), estimate from Weighted Median Estimator (WME) and a control for potential pleiotropy with MR-Egger regression. Multiple testing was corrected with Bonferroni method and False Discovery Rate (FDR).

Risultati

Results show that gene expression in brain seems to play a role in MS pathogenesis, both in a protective or a risky way. Results confirm p52 increased expression as a risk factor in cerebellum (OR=1.44 [CI 0.25-0.48]; p-value<0.0001) and in frontal cortex (OR=1.32 CI [0.19-0.37]; p-value< 0.0001) while increased expression IKBα, the inhibitor of p52, as a protective factor in different brain regions. Many of the products of NF-kB pathway seem to have a different role depending on brain tissue. For example, TNFRSF1A receptor increased expression is causally protective in frontal and temporal cortex while it is a risk factor in occipital cortex, indicating different possible responses and different balances in each brain tissue. RelB, which is a transcription subunit of non-canonical NF-kB pathway, increased expression is causally related to an increased risk of MS in temporal cortex and putamen, suggesting how both canonical and non-canonical NF-kB pathways could be potential risk factors in brain tissues. Upstream protein kinases as IBKβ, CHUK, NEMO, TAK1 and receptors TNFRSF1A and LTBR have both a protective and risk role, probably due to involvement in different pathways.

Conlcusioni

Results suggest that NF-kB canonical pathway seems to have a role in MS occurrence since IKBα inhibitor increased expression has a protective role in different brain regions while transcription subunit p50 have a risk role, a result already highlighted in literature [11]. Non-canonical NF-kB pathway seems to be involved too, since increased transcription subunit RelB and transcription of BAFF and ICAM1 have a risk role too. ICAM1, which is involved in leukocytes recruitment through the blood-brain barrier, have a risk role in thalamus and substantia nigra and it would be interesting to investigate if these regions are susceptible to blood-brain barrier disruption. CCL2 and CXCL10 increased expression results as a risk factor and it confirms previous studies [12-13-14]. Results are different and even opposite depending on the tissue, suggesting more investigation at tissue level and precise therapeutic targeting. Finally, we have a possible clue that gene expression in brain could play a role in pathogenesis of MS independently of peripherical autoimmune stimulation, as suggested by the “inside-out” hypothesis, where MS is initiated by primary degeneration of oligodendrocytes and myelin, resulting in the release of antigenic myelin components that elicits an autoimmune reaction against myelin components [6].

Bibliografia

[1] Ghasemi N., Razavi S., Nikzad E., Multiple Sclerosis: Pathogenesis, Symptoms, Diagnoses and Cell-Based Therapy. Cell J., 2017 Apr-Jun; 19(1): 1–10.

[2] Consortium IMSG, Beecham AH., Patsopoulos NA., Xifara DK., Davis MF, Kemppinen A., et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013;45:1353–60.

[3] Patsopoulos NA., Barcellos LF., Hintzen RQ., et al. Fine-Mapping the Genetic Association of the Major Histocompatibility Complex in Multiple Sclerosis: HLA and Non-HLA Effects. PLoS Genet. 2013;9:e1003926.

[4] Srinivasan M., Lahiri DB., Significance of NF-κB as a pivotal therapeutic target in the neurodegenerative pathologies of Alzheimer’s disease and multiple sclerosis. Expert Opin Ther Targets. 2015 Apr; 19(4): 471–487.

[5] Leibowitz SM., Yan J., NF-κB Pathways in the Pathogenesis of Multiple Sclerosis and the Therapeutic Implications. Front Mol Neurosci. 2016; 9: 84.

[6] Yuan Y., Stone S., Lin., Role of nuclear factor κB in multiple sclerosis and experimental autoimmune encephalomyelitis. Neural Regen Res. 2018 Sep; 13(9): 1507–1515.

[7] Edwards SL., Beesley J., French JD., Dunning AM. Beyond GWASs: Illuminating the Dark Road from Association to Function. Am J Hum Genet. 2013;93:779–97.

[8] Davey Smith G., Hemani G., Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23:R89–98.

[9] Fazia T., Pastorino R., Foco L., Han L., Abney M., Beecham A., et al. Investigating multiple sclerosis genetic susceptibility on the founder population of east-central Sardinia via association and linkage analysis of immune-related loci. Mult Scler J. 2017;24:1815–24.

[10] Ramasamy A., Trabzuni D., Guelfi S., Varghese V., Smith C., Walker R., et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. Nat Neurosci. 2014;17:1418–28.

[11] Housley W., Fernandez SD., Vera K.., Genetic variants associated with autoimmunity drive NFκB signaling and responses to inflammatory stimuli. Sci Transl Med. 2015 Jun 10; 7(291): 291ra93.

[12] Kim RY., Hoffman AS., Itoh N. ed al., Astrocyte CCL2 sustains immune cell infiltration in chronic experimental autoimmune encephalomyelitis. J Neuroimmunol. 2014 Sep 15;274(1-2):53-61.

[13] Ponath G., Park C., Pitt D., The Role of Astrocytes in Multiple Sclerosis. Front Immunol. 2018; 9: 217.

[14] Skripuletz T., Hackstette D., Bauer K., Astrocytes regulate myelin clearance through recruitment of microglia during cuprizone-induced demyelination. Brain. 2013 Jan;136(Pt 1):147-67.

RELATIONSHIP BETWEEN TOTAL PROTEIN INTAKE AND KIDNEY FUNCTION IN ADULTS FROM THE COOPERATIVE HEALTH RESEARCH IN SOUTH TYROL (CHRIS) STUDY

Vukovic Vladimir[1*], Gögele Martin[1], Rainer Johannes[1], Garcia-Larsen Vanessa[2], Domingues Francisco[1], Pramstaller P Peter[1], Pattaro Cristian[1]

[1] *Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Bolzano, Italy*

[2] *Center for Human Nutrition, Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America*

Introduction

There is increasing public health concern about the consequences that modern diet, which is rich in processed and red meat, may have on the risk of chronic non-communicable diseases. Chronic kidney disease (CKD) is among the most widespread common chronic diseases in industrialized countries, where it represents a major public health issue. CKD affects nearly 10% of the world's population, with globally increasing prevalence and limited treatment options [1]. It represents a high burden for the health system and is associated with high disease-attributable mortality and multi-systemic complications [2]. Elevated protein intake may damage the kidney, leading to increased glomerular pressure and glomerular hyperfiltration: the consequent damage to the glomerular structure may lead to CKD or accelerate its progression [3]. However, cross-sectional and longitudinal analyses of association between protein intake and kidney function provide different and often contrasting perspectives [4]. While data from the general population are scarce, they may help elucidate the protein intake - kidney function relationship in a real-world scenario.

Objectives

Aim of our study was to investigate the relationship between total daily protein intake (TDPI) and kidney function and the CKD in a general population sample.

Methods

The analysis was carried out in the Cooperative Health Research In South Tyrol (CHRIS) study, an ongoing population-based study being conducted in Val Venosta (South Tyrol, Italy) [5]. Participants completed a Food Frequency Questionnaire (FFQ) developed by the Global Allergy and Asthma European Network (GA2LEN) study [6], which included local foods of the South Tyrolean region. The frequency of food consumption was transformed into grams per day for each food item and converted into nutrient estimates such as the TDPI, using the UK's Food Composition Table [7]. Kidney function was quantified by the glomerular filtration rate estimated from serum creatinine (eGFR) using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation [8]. CKD was defined as eGFR<60 ml/min/1.73m2. The relationship between eGFR and TDPI was assessed by fitting linear mixed regression models adjusting for age, sex, and other potential confounders, while logistic regression was used to asses the association between CKD and TDPI.

Results

Complete data for analyses was available for 5883 out of 10,518 participants (CHRIS study data release no. 3.0, June 7th, 2018), of whom 55% were females. The median age was 46 years old (interquartile range, IQR = 33, 58). The mean eGFR was 93.27 ml/min/1.73m2 (standard deviation, SD = 15.89). One-hundred thirty-three participants (2.26%, 95% confidence interval, CI = 1.90, 2.67) were affected by CKD. The median

calculated TDPI was 83.17 g (IQR = 66.58, 103.52). The effect of TDPI on eGFR levels was of –0.003 ml/min/1.73 m2 per one unit increase of TDPI (standard error, SE = 0.005; p-value = 0.48). Results did not change when additionally accounting for the body mass index (BMI; effect = -0.003, SE = 0.005, p-value = 0.53). The odds ratio (OR) of CKD per one unit increase of TDPI was 1.00 (95% CI = 0.99, 1.01; p-value = 0.77) and remained unchanged when accounting for BMI (OR = 1.00, 95% CI = 0.99, 1.01; p-value = 0.79).

Conclusions

Our results did not support an association between TDPI and kidney function or CKD. This might be explained by CKD being typically silent in its earliest stages, for which reason, people might not modify their diet early enough to prevent progression of the kidney damage. Another aspect is the temporal allocation of our measurements: while eGFR chronically declines over time, along with people ageing, we could only assess dietary habits relatively to 12 months prior to the participation, which may prevent from observing a long-term effect of TDPI on kidney function. The main strength of our study is the assessment of kidney function and dietary habits in a large sample from the general population. Among the limitations, total protein intake was estimated based on an FFQ and not objectively quantified. In addition, eGFR was approximated by serum creatinine and CKD was estimated based on a single eGFR measurement, whilst at least two measures at a distance of at least 3 months should be considered for a proper diagnosis. In conclusion, longitudinal studies that can allocate dietary habits earlier in time as compared to the CKD onset are warranted. Additionally, for chronic conditions that remain silent for a long time, it would be important to include the information about clinical diagnosis or individuals' awareness in the models, for a better interpretation of possible cause-effect relationship.

Bibliography

[1] Eckardt K-U, Coresh J, Devuyst O, et al. Evolving importance of kidney disease: from subspecialty to global health burden. Lancet. 2013;382:158–69

[2] Jha V, Garcia-Garcia G, Iseki K, et al. Chronic kidney disease: global dimension and perspectives. Lancet. 2013;382(9888):260-72.

[3] Ko GJ, Obi Y, Tortorici AR, et al. Dietary protein intake and chronic kidney disease. Curr Opin Clin Nutr Metab Care. 2017;20:77–85.

[4] Cirillo M, Lombardi C, Chiricone D, et al. Protein intake and kidney function in the middle-age population: contrast between cross-sectional and longitudinal data. Nephrol Dial Transplant. Oxford University Press;2014;29:1733–40

[5] Pattaro C, Gögele M, Mascalzoni D, et al. The Cooperative Health Research in South Tyrol (CHRIS) study: rationale, objectives, and preliminary results. J Transl Med. 2015;13:348

[6] Garcia-Larsen V, Luczynska M, Kowalski ML, et al. Use of a common food frequency questionnaire (FFQ) to assess dietary patterns and their relation to allergy and asthma in Europe: pilot study of the GA2LEN FFQ. Eur J Clin Nutr. 2011;65:750–6

[1] Food Standard Agency, Public Health England. McCance and Widdowson's The Composition of Foods. Cambridge: Royal Society of Chemistry; 2014

[8] Pattaro C, Riegler P, Stifter G, et al. Estimating the Glomerular Filtration Rate in the General Population Using Different Equations: Effects on Classification and Association. Nephron Clin Pract. 2013;123:102–11

DIETARY NON-ENZYMATIC ANTIOXIDANT CAPACITY AND THE RISK OF CARDIOVASCULAR DISEASES IN A LARGE COHORT OF SWEDISH WOMEN

Essi Hantikainen[1], Marie Löf[2,3], Alessandra Grotta[4,5], Ylva Trolle Lagerros[6,7], Mauro Serafini[8], Rino Bellocco[1,5], Elisabete Weiderpass[5]

[1] *Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy.*

[2] *Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden.*

[3] *Department of Medical and Health Sciences, Linköping University, Linköping, Sweden.*

[4] *Department of Public Health Sciences, Centre for Health Equity Studies (CHESS), Stockholm University, Stockholm, Sweden*

[5] *Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.*

[6] *Clinical Epidemiology Unit, Department of Medicine, Karolinska University Hospital Huddinge, Stockholm, Sweden.*

[7] *Obesity Center, Academic Specialist Center, Stockholm Health Services, Stockholm, Sweden.*

[8] *Functional Food and Metabolic Stress Prevention Laboratory, Faculty of BioSciences and Technology for Food, Agriculture and Environment, University of Teramo, Teramo, Italy.*

Introduction

Because of controversial findings from randomized clinical trials, antioxidant supplements, such as vitamin E or C, are not recommended for the prevention of cardiovascular diseases [1]. Antioxidants are known to interact synergistically with each other [2] and therefore vitamin supplementation might not yield the same beneficial effects as would be achieved through the consumption of whole foods rich in antioxidants.

Aim

To further clarify the role of dietary antioxidants on the risk of CVDs, we aimed to prospectively investigate the association between dietary Non Enzymatic Antioxidant Capacity (NEAC), which measures the antioxidant potential of the whole diet, an the risk of several cardiovascular outcomes, i.e. myocardial infarction (MI), heart failure (HF) and stroke.

Methods

In the Swedish Women's Lifestyle and Health Cohort 45,882 women aged 30 to 49 years and free from CVDs and cancer were followed from 1991 until 2012 through linkages to National Health Registers. We identified 657 incident cases of MI, 881 incident cases of stroke and 245 incident cases of heart failure. Total dietary NEAC was assessed by combining information from a validated food frequency questionnaire collected at baseline with NEAC values from an antioxidant food database [3]. We categorized the distribution of NEAC into quantiles and used multivariable adjusted Cox proportional hazards regression models to estimate hazard ratios with 95% confidence intervals. Dose-response relationships were further investigated using restricted cubic splines [4].

Results

We identified 657 incident cases of MI, 881 incident cases of stroke and 245 incident cases of heart failure. A higher baseline dietary NEAC was inversely associated with a lower risk of MI (p for trend <0.001) [5] and

heart failure (p for trend: 0.02) and based on the spline regression the relationship seemed to be linear (p values for non-linearity > 0.05). No association was found with stroke (p for trend >0.05).

Conclusion

Our findings suggest that a higher baseline dietary NEAC might be associated with lower risk of MI and HF but not stroke in young and middle aged women.

Bibliography

[1] Ye Y, Li J, Yuan Z, Effect of antioxidant vitamin supplementation on cardiovascular outcomes: a meta-analysis of randomized controlled trials. Plos One. 2013;8(2):e56803.

[2] Serafini M, Del Rio D, Understanding the association between dietary antioxidants, redox status and disease: is the Total Antioxidant Capacity the right tool? Redox Rep. 2004;9(3):145-52.

[3] Carlsen MH, Halvorsen BL, Holte K et al., The total antioxidant content of more than 3100 foods, beverages, spices, herbs and supplements used worldwide. Nutr J. 2010;9.

[4] Harrell F, Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Springer; 2015.

[5] Hantikainen E, Lof M, Grotta A, Trolle Lagerros Y, Serafini M, Bellocco R, et al. Dietary non enzymatic antioxidant capacity and the risk of myocardial infarction in the Swedish women's lifestyle and health cohort. Eur J Epidemiol. 2018;33(2):213-21.

EFFECT OF DEPRESSION AND LOW-GRADE SYSTEMIC INFLAMMATION ON MORTALITY AND HOSPITALIZATION RISK: DATA FROM THE MOLI-SANI STUDY

Gialluisi Alessandro[1*], Costanzo Simona[1], Di Castelnuovo Augusto[2], Bonaccio Marialaura[1], Bracone Francesca[1], De Curtis Amalia[1], Cerletti Chiara[1], Donati Maria Benedetta[1], De Gaetano Giovanni[1], Iacoviello Licia[1,3], on behalf of the Moli-sani Study Investigators[4]

[1] *Department of Epidemiology and Prevention, IRCCS NEUROMED, Pozzilli, Italy*

[2] *Mediterranea Cardiocentro, Napoli, Italy*

[3] *Research Center in Epidemiology and Preventive Medicine (EPIMED), Department of Medicine and Surgery, University of Insubria, Varese, Italy*

[4] *A complete list of the Moli-sani Project Investigators is reported online at https://www.moli-sani.org/?page_id=181*

Introduction

Depression and low-grade systemic inflammation are associated risk factors for mortality and hospitalizations, although the nature of this relationship is underinvestigated.

Aims

To clarify the effect of depression and low-grade systemic inflammation on all-cause mortality and first hospitalization risk.

Methods

In an Italian population cohort, the Moli-sani study (N=24,325; age≥35 years; 51.5% women) [1], depressive symptoms were assessed to classify subjects as non- , moderately-, and severely-depressed [2]. Death events and first hospital admissions were recorded through linkage with regional mortality and hospital discharge registries. Multivariable Cox regressions on all-cause mortality/hospitalizations were performed using depression forms as exposure, adjusting for age, sex, education level, prevalent chronic conditions (cardiovascular disease, cancer and diabetes) and main lifestyle factors (smoking, diet and physical activity) [3]. The effect of inflammation was quantified through i) a mediation analysis on INFLA-score, a global composite index based on four circulating inflammatory biomarkers [3], and ii) alternative Cox regressions including INFLA-score as covariate.

Results

Over 12,742 subjects with complete information (440 death events; median follow-up 8.24 years), severely-depressed subjects showed about twice the risk of all-cause mortality, compared to non-depressed (HR=1.95; $CI_{95\%}$=1.11-3.44; p=0.021). Similarly, analysis of hospitalizations (N=12,753; 4,640 events; 7.28 years) revealed an increased risk for moderately- (HR=1.20; 1.13-1.28; $p<10^{-4}$) and severely-depressed participants (HR=1.46; 1.19-1.80; $p=3x10^{-4}$), respectively. These estimates were not substantially attenuated by including INFLA-score as additional covariate in the models. INFLA-score showed a significant additive (but not interactive) effect on both mortality (HR=1.018; $CI_{95\%}$=1.001-1.034; p=0.033) and hospitalizations (1.008; 1.003-1.013; p=0.001). Moreover, it mediated a small but significant proportion of the effect of depression on mortality (7.5%; p=0.021), but not on hospitalizations (1.2%; p=0.066).

Conclusion

Depression increases the risk of both mortality and hospitalizations. This effect is only partly explained by low-grade systemic inflammation, which exerts an independent additive influence on both clinical events.

References

[1] Bonaccio M, Di Castelnuovo A, Costanzo S, et al. Mediterranean diet and mortality in the elderly: a prospective cohort study and a meta-analysis. Br J Nutr., 2018 Oct; 120(8):841-854.

[2] Kroenke K, Strine TW, Spitzer RL, et al. The PHQ-8 as a measure of current depression in the general population. J Affect Disord., 2009 Apr; 114(1-3):163-73.

[3] Gialluisi A, Bonaccio M, Di Castelnuovo A, et al. Lifestyle and biological factors influence the relationship between mental health and low-grade inflammation. Brain Behav Immun., 2019 May; doi:10.1016/j.bbi.2019.04.041.

LO STUDIO VIADANA III: INQUINAMENTO ATMOSFERICO E SALUTE IN ETÀ PEDIATRICA

Silvia Panunzi[1], Pierpaolo Marchetti[1], Chiara Badaloni[2], Kees de Hoogh[3,4], Linda Guarda[5], Francesca Locatelli[1], Vanda Pironi[5], Caterina Silocchi[6], Massimo Stafoggia[2], Paolo Ricci[5], Alessandro Marcon[1]

[1] *Sezione di Epidemiologia e Statistica Medica, Dipartimento di Diagnostica e Sanità Pubblica, Università di Verona, Verona*

[2] *Dipartimento di Epidemiologia, Servizio Sanitario Regione Lazio, Roma*

[3] *Swiss Tropical and Public Health Institute, Basel, Svizzera*

[4] *Department of Public Health, University of Basel, Basel, Svizzera*

[5] *UOC Osservatorio Epidemiologico, ATS Val Padana, Mantova*

[6] *UOS Salute e Ambiente, ATS Val Padana, Mantova*

Introduzione

Lo studio Viadana III nasce come approfondimento di precedenti studi effettuati nel distretto sanitario di Viadana, in provincia di Mantova, sede del più grande polo industriale italiano del pannello truciolare e del legno. Il distretto ospita due grandi aziende a sud, nei comuni di Viadana e Pomponesco, e altre fabbriche più piccole di lavorazione del legno, distribuite al centro e al sud del territorio.
A partire dal 2005, l'Università di Verona e l'ATS Val Padana (ex-Asl di Mantova), in collaborazione con ARPA Mantova, hanno condotto due indagini epidemiologiche sulla popolazione pediatrica residente nel distretto di Viadana (http://biometria.univr.it/viadanastudy), con l'obiettivo di valutare l'impatto dell'inquinamento atmosferico di origine industriale sulla salute pubblica.
Gli studi precedenti hanno prodotto evidenze che si è ritenuto utile approfondire, quali la presenza di sintomi respiratori e irritativi e di eccessi di ricoveri per cause respiratorie nei bambini che vivevano in prossimità delle fabbriche del distretto e un più alto livello di biomarcatori di danno genotossico nelle cellule della mucosa orale dei bambini residenti in zone con concentrazioni maggiori di formaldeide e biossido di azoto (NO_2) [1-4].

Obiettivi

Viadana III si propone di integrare le indagini precedenti valutando se i bambini e ragazzi per i quali si sono riscontrati esiti sanitari e biologici potenzialmente correlati alle attività industriali siano tuttora a rischio di eventi avversi e se un analogo rischio risulti documentabile su una nuova popolazione pediatrica identificata nel distretto. A tal fine sarà valutata l'esistenza di associazioni tra esposizione a inquinanti atmosferici ed indicatori sanitari ricavati dai dati correnti, utilizzando il disegno di coorte prospettico.
Il presente abstract ha l'obiettivo di fornire una descrizione generale dello studio Viadana III, attualmente in corso, e riportare alcune statistiche descrittive sui dati di esposizione raccolti fino ad oggi.

Metodi

Popolazioni in studio. Saranno analizzate le due coorti sopra menzionate: la 'coorte storica' di 3756 ragazzi nati negli anni 1992-2003, che avevano 3-14 anni nell'anno 2006, in cui si è condotto il primo studio e una 'nuova coorte' di 4206 bambini nati tra gli anni 2004-2012 e residenti nel distretto di Viadana al 31 dicembre 2012.

Fonti dei dati sanitari e confondenti. Le informazioni demografiche per i bambini della coorte storica saranno ricavate dal questionario compilato dai genitori nel 2006, nell'ambito dello studio Viadana I e comprendono sesso, età, nazionalità, livello di istruzione dei genitori (come *proxy* di stato socio-economico), abitudine al fumo dei genitori, esposizione al traffico veicolare. Per la nuova coorte le informazioni a disposizione saranno sesso, età e nazionalità, desumibili dall'anagrafe sanitaria.
Gli outcomes in esame saranno in particolare l'occorrenza di ricoveri ospedalieri, accessi al pronto soccorso, visite mediche specialistiche e uso dei farmaci, soprattutto per patologie respiratorie, durante il periodo di follow-up (2013-2017).
I dati dei consumi sanitari sono forniti e verificati dall'UOC Osservatorio Epidemiologico, ATS Val Padana, che utilizza questi flussi sanitari al fine di ricostruire, con cadenza annuale, la storia dei consumi sanitari di tutta la popolazione assistita della provincia di Mantova.
Fonti di dati ambientali. La principale fonte di dati sulle concentrazioni degli inquinanti nel distretto saranno le mappe di NO_2 e formaldeide, elaborate nel contesto di Viadana II mediante ordinary kriging [4]. Queste mappe sono state costruite interpolando le concentrazioni annuali dei due inquinanti, stimate mediante campagne di monitoraggio condotte nel 2010 [4] utilizzando 62 campionatori passivi. I campionatori erano stati collocati nel territorio del distretto in modo da massimizzare la rappresentatività delle esposizioni sulla popolazione target e l'accuratezza delle misure effettuate nelle aree di ricaduta attesa delle emissioni, approssimate da buffer circolari di 3 km.
Saranno inoltre utilizzati i dati ambientali ottenuti in altri studi, con lo scopo di stimare l'esposizione anche ad inquinanti precedentemente non disponibili ma anche per ottenere, per l'NO_2, un confronto con i nostri dati. In questa prima fase dell'indagine sono state prodotte stime di esposizione di carbonio elementare, NO_2 e $PM_{2.5}$ (anno di riferimento 2010) sulla base del progetto ELAPSE (Effects of Low-Level Air Pollution: A Study in Europe) [5,6] per l'intera area europea e stime di esposizione di $PM_{2.5}$ (2013) e PM_{10} (2012) prodotte a livello nazionale in EPISAT (Dati satellitari ed uso del territorio per la stima delle esposizioni a livello nazionale) [7,8].Entrambi gli studi hanno prodotto modelli di esposizione Land Use Regression utilizzando i dati delle stazioni di monitoraggio della qualità dell'aria operanti nel relativo contesto geografico, uniti ai dati cartografici (per esempio uso del suolo, densità di popolazione, strade e traffico) e da satellite disponibili.
Attribuzione delle stime individuali di esposizione. Gli indirizzi di residenza dei bambini delle due coorti al baseline (2012) sono stati georeferenziati, ottenendo coordinate geografiche di qualità accurata per circa il 95% dei partecipanti (3482 soggetti della coorte storica e 4043 della nuova coorte). Le stime di esposizione agli inquinanti in corrispondenza delle abitazioni sono state ottenute utilizzando per formaldeide e NO_2 i modelli variogramma (ordinary kriging) precedentemente sviluppati [4]. Per gli inquinanti disponibili in ELAPSE le mappe sono state utilizzate le mappe europee con risoluzione 100×100 m fornite da Kees de Hoogh [6]. Per gli inquinanti disponibili in EPISAT le stime sono state calcolate specificamente per questa indagine da Massimo Stafoggia e Chiara Badaloni, ottimizzando le stime ottenute su celle con risoluzione 1x1 km^2 (stage 3) sugli indirizzi geocodificati (stage 4), a partire da concentrazioni su base giornaliera a livello italiano [6,7].
È stata inoltre calcolata la distanza delle abitazioni dei soggetti dalle fonti emissive industriali. I bambini e ragazzi sono stati classificati in tre gruppi: "non esposti" (distanza > 2km da tutte le fabbriche del distretto), gruppo a "bassa esposizione" (≤ 2km dalle fabbriche più piccole), gruppo ad "alta esposizione" (≤ 2km dalle grandi aziende del pannello truciolare).

Risultati

La tabella 1 riporta la distribuzione delle stime di esposizione agli inquinanti atmosferici al baseline (2012) in funzione della distanza dalle fabbriche. Visti gli scopi descrittivi della tabella, i dati delle due coorti sono stati combinati. L'associazione tra esposizioni e distanza dalle fabbriche è stata testata tramite test ANOVA.

Da questi risultati si può notare la tendenza di tutti gli inquinanti stimati ad aumentare avvicinandosi alle fonti di emissioni, in particolare le aziende più grandi, con l'eccezione delle polveri sottili ottenute nell'ambito di EPISAT.
Si osserva inoltre che le stime di esposizione a NO_2 ottenute utilizzando i campionatori passivi risultano circa 10 µg/m^3 inferiori rispetto a quelle ottenute dalle mappe europee di ELAPSE, verosimilmente per differenze nei dati di input e nei metodi di stima.

Tabella 1. *Distribuzione delle stime di esposizione al baseline in funzione della distanza dalle fabbriche.*

Distanza dalle fabbriche	**> 2km dalle fabbriche (N=2687)**	**≤ 2km dalle piccole aziende (N=2402)**	**≤ 2km dalle grandi aziende (N=2436)**	p-value
NO_2 µg/m^3				< 0.001
Media (SD)	15.3 (1.6)	15.9 (2.0)	17.9 (1.8)	
Mediana (5°-95° percentile)	15.4 (12.7- 18.0)	16.4 (12.9- 18.6)	18.3 (14.7- 20.4)	
Formaldeide µg/m^3				< 0.001
Media (SD)	2.4 (0.2)	2.5 (0.1)	2.7 (0.2)	
Mediana (5°-95° percentile)	2.5 (2.1- 2.7)	2.5 (2.3- 2.7)	2.7 (2.5- 3.0)	
NO_2 (Elapse) µg/m^3				< 0.001
Media (SD)	26.1 (2.6)	26.6 (3.0)	28.3 (2.2)	
Mediana (5°-95° percentile)	26.2 (21.7- 29.6)	26.5 (21.9- 31.2)	28.9 (23.9- 31.2)	
$PM_{2.5}$ (Elapse) µg/m^3				< 0.001
Media (SD)	24.3 (1.4)	25.1 (1.5)	26.6 (0.6)	
Mediana (5°-95° percentile)	24.0 (22.2- 26.6)	25.5 (22.5- 26.9)	26.8 (25.5- 27.2)	
Carbonio elementare (Elapse) 10^{-5}m^{-1}				< 0.001
Media (SD)	1.9 (0.1)	1.9 (0.1)	2.0 (0.1)	
Mediana (5°-95° percentile)	1.9 (1.7- 2.1)	1.9 (1.7- 2.1)	2.0 (1.9- 2.1)	
PM_{10} (Episat) * µg/m^3				< 0.001
Media (SD)	37.0 (3.6)	35.0 (3.4)	37.1 (3.3)	
Mediana (5°-95° percentile)	36.9 (31.6- 43.2)	35.1 (29.4- 40.3)	36.9 (30.5- 42.5)	
$PM_{2.5}$ (Episat) * µg/m^3				0.001
Media (SD)	25.3 (2.2)	25.2 (1.6)	25.4 (1.4)	
Mediana (5°-95° percentile)	25.3 (21.7- 28.9)	25.2 (22.2- 27.6)	25.4 (23.5- 27.7)	

* al momento disponibili solo per la nuova coorte di bambini (n=4043).

Conclusioni

In questa prima fase dell'indagine abbiamo prodotto stime di esposizione agli inquinanti atmosferici in corrispondenza dell'abitazione dei partecipanti. Le stime di esposizione ottenute mostrano, in generale, una buona concordanza con le distanze dalle fabbriche, con l'eccezione delle polveri sottili disponibili in EPISAT. Queste differenze potrebbero essere in parte dovute alla distribuzione più "regionale" delle polveri sottili, meno correlata alle fonti di emissione e più influenzata dalle condizioni metereologiche rispetto per esempio all'NO_2. Nelle successive fasi dell'indagine si approfondiranno aspetti quali la correlazione tra i diversi inquinanti e la presenza di trend temporali nell'andamento dell'inquinamento atmosferico durante il periodo dello studio.
Sarà anche valutata la stabilità nella distribuzione spaziale degli inquinanti misurati nel 2010 mediante campionatori passivi, anche utilizzando i dati di una nuova campagna di monitoraggio condotta nel 2018 da ARPA Mantova nel comune di Viadana.

Bibliografia

[1] de Marco, Roberto, et al. "Proximity to chipboard industries increases the risk of respiratory and irritation symptoms in children: the Viadana study." Science of the total environment408.3 (2010): 511-517.

[2] Rava, Marta, et al. "Proximity to wood factories and hospitalizations for respiratory diseases in children." Science of the Total Environment 410 (2011): 80-86.

[3] Marchetti, Pierpaolo, et al. "Children living near chipboard and wood industries are at an increased risk of hospitalization for respiratory diseases: a prospective study." International journal of hygiene and environmental health 217.1 (2014): 95-101.

[4] Marcon, Alessandro, et al. "Outdoor formaldehyde and NO2 exposures and markers of genotoxicity in children living near chipboard industries." Environmental health perspectives 122.6 (2014): 639-645.

[5] Vienneau, Danielle, et al. "Western European land use regression incorporating satellite-and ground-based measurements of NO2 and PM10." Environmental science & technology 47.23 (2013): 13555-13564.

[6] De Hoogh, Kees, et al. "Spatial PM2. 5, NO2, O3 and BC models for Western Europe–Evaluation of spatiotemporal stability." Environment international 120 (2018): 81-92.

[7] Stafoggia, Massimo, et al. "Estimation of daily PM10 concentrations in Italy (2006–2012) using finely resolved satellite data, land use variables and meteorology." Environment international 99 (2017): 234-244.

[8] Badaloni, Chiara, et al. "Big data in epidemiologia ambientale. Dati satellitari e uso del territorio per la stima delle esposizioni a livello nazionale." Epidemiol Prev 42.1 (2018): 46-59.

LONG TERM INCIDENCE OF SYMPTOMATIC CARDIAC EVENTS IN 2,809 BREAST CANCER PATIENTS TREATED WITH ADJUVANT TRASTUZUMAB IN REAL-WORLD PRACTICE.

Franchi Matteo[1,2], Di Cosimo Serena[3], Trama Annalisa[4], Merlo Ivan[1,2], Minicozzi Pamela[5], Tarantini Luigi[6], Merlino Luca[7], Apolone Giovanni[8], Corrao Giovanni[1,2]

[1] *National Centre for Healthcare Research and Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy*

[2] *Laboratory of Healthcare Research and Pharmacoepidemiology, Unit of Biostatistics, Epidemiology and Public Health, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy*

[3] *Biomarker Unit, Department of Applied Research and Technological Development, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*

[4] *Evaluative Epidemiology Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*

[5] *Analytical Epidemiology and Health Impact Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*

[6] *Cardiology Department, Ospedale San Martino, S.C. Cardiologia, Ospedale San Martino, Azienda ULSS 1, Belluno*

[7] *Head, Epidemiologic Observatory, Regional Welfare Service, Lombardy Region, Milan, Italy*

[8] *Scientific Directorate, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*

Introduction

Clinical trials, which describe well the acute toxicities in selected patients [1-2], may have underestimated toxicities in patients with advanced age and co-morbidities, which are common in real-world practice. More importantly, clinical trials do not provide extended follow-up. On the other hand, few observational studies have been conducted in order to assess the long-term safety of adjuvant trastuzumab in clinical practice [3-5].

Aims

The main objective of the study was to assess acute and long-term symptomatic cardiac events (SCEs) after adjuvant trastuzumab-based chemotherapy (T) in a large and unselected breast cancer patient's population. The secondary objective was to compare the risk of SCEs in women treated with T as compare to those treated with chemotherapy only.

Methods

Using healthcare administrative database, i.e. clinical discharge records and drug prescriptions of the Lombardy region (Italy), we selected patients newly diagnosed with early breast cancer between 01/2008 and 12/2011 and monitored until 12/2016. Patients treated with T were 1:2 matched with patients treated with chemotherapy only for age, date of treatment start, and cardiovascular risk factors (including diabetes, hypertension, cardiac arrhythmia and renal diseases). SCEs included heart failure and cardiomyopathy, based on ICD9-CM codes. The cumulative risk of SCEs was estimated using the Kaplan-Meier method, and independent predictors were assessed by means of a Cox regression model.

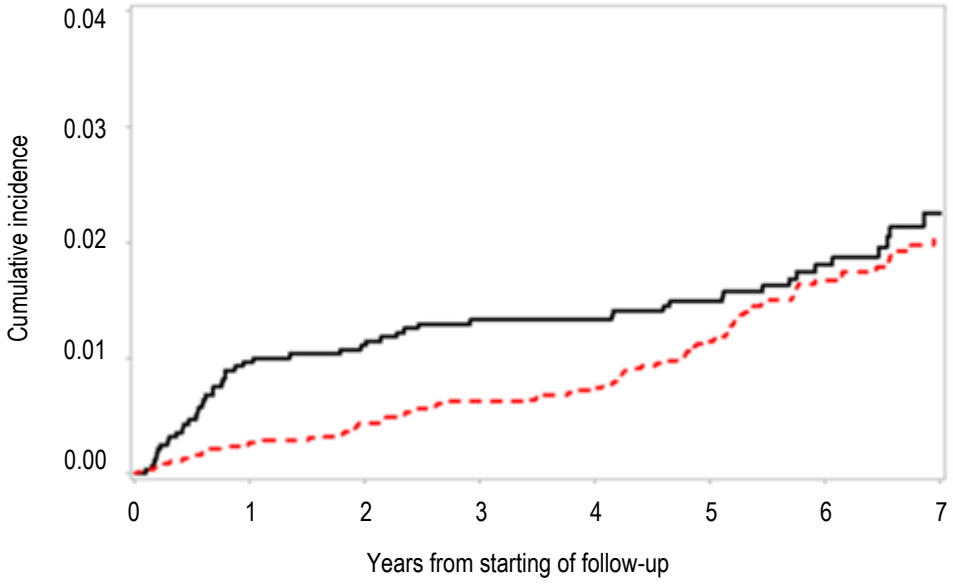
Results

Of a cohort of 34,218 pts with incident BC, 2,809 patients treated with T were matched to 5,618 pts treated with chemotherapy only. Among T-users, 52 (1.8%) experienced one SCE. No cardiac death occurred. The cumulative risk of SCEs was 2.25% and increased up to 1 year after starting T (0.96%), remitting after treatment. One fourth of SCEs occurred within the first 6 months. The cumulative risk in non T-users was 2.0% (0.16% during the first year). The hazard ratio [HR] was 9.96 (95% CI 3-78-26.2) in the first year, and 1.41 (95%CI 0.99-2.02) during the entire follow-up. In the unmatched cohort, HR was 8.77 (95% CI 5.25-14.64) in patients aged ≥70 years, and 2.32 (95% CI 1.68-3.22) in patients with at least one pre-existing cardiovascular risk factor.

Conclusions

In real-world practice, SCEs during or after T are infrequent, early and self-limiting. Based on the timeline of SCEs, it might be unnecessary to monitor cardiac toxicity beyond the period of T treatment. Besides, to reduce the excess of cardiac risk of the first year, strategies including shortening T exposure or increasing the number of check-ups in asymptomatic patients should be accompanied by the development of biomarker able to identify patients at risk before or immediately after T initiation.

Figure 1. *Cumulative incidence of hospitalization for cardiovascular events (heart failure or cardiomyopathy) in women treated with trastuzumab-based chemotherapy (black solid line) or chemotherapy alone (red dotted line).*



References

[1] Slamon D, Eiermann W, Robert N, et al. Adjuvant trastuzumab in HER2-positive breast cancer. N Engl J Med. 2011 Oct 6;365(14):1273-83

[2] Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med. 2005 Oct 20;353(16):1659-72

[3] Bonifazi M, Franchi M, Rossi M, et al. Long term survival of HER2-positive early breast cancer treated with trastuzumab-based adjuvant regimen: a large cohort study from clinical practice. Breast. 2014 Oct;23(5):573-8

[4] Chen J, Long JB, Hurria A, et al. Incidence of heart failure or cardiomyopathy after adjuvant trastuzumab therapy for breast cancer. J Am Coll Cardiol. 2012 Dec 18;60(24):2504-12

[5] Bowles EJ, Wellman R, Feigelson HS, et al; Pharmacovigilance Study Team. Risk of heart failure in breast cancer patients after anthracycline and trastuzumab treatment: a retrospective cohort study. J Natl Cancer Inst. 2012 Sep 5;104(17):1293-305.

DEVELOPMENT OF A FOOD KNOWLEDGE QUESTIONNAIRE FOR TANZANIAN WOMEN OF CHILDBEARING AGE

Maria Vittoria Conti[1*], Marco Gnesi[2*], Maria Cristina Monti[2], Rachele De Giuseppe[1], Angela G. Mkindi[3], Naelijwa Mshanga[3], Hellas Cena[1,4]

*same contribution

[1] *Laboratory of Dietetics and Clinical Nutrition, Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia (Italy)*

[2] *Section of Biostatistics and Clinical Epidemiology, Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy*

[3] *Nelson Mandela African Institution of science and Technology, Arusha, Tanzania*

[4] *Clinical Nutrition and Dietetics Service, Unit of Internal Medicine and Endocrinology, ICS Maugeri IRCCS, Pavia, Italy*

Introduction

The incidence of malnutrition in developing countries, such as Sub Saharan Africa, is still high, especially in people with increased nutritional needs, such as children and women in childbearing age [1,2]. It has been reported that women are food systems "pillars", because of their role as agricultural food producers and because they are the ones in charge with the entire process of food preparation, food preservation, and processing [3].
Women's food knowledge (FK), defined as the "*understanding of basic nutrition concepts"* [4,5], is positively associated with child nutritional outcomes [4]. Since women are the primary caregivers, they impact children's nutrition indirectly through their own nutritional status as well as directly throughout childcare [3]. Thus, enhancing FK will likely lead to greater awareness and autonomy in food choices, and will help to develop public health policies aimed at tackling malnutrition at different levels.
It is acknowledged that food choices and eating behaviors contribute to malnutrition [4]. In addition, evidences show that FK is one of the key aspects among those influencing dietary intake and eating habits [5]. Enhancing FK throughout education could play a pivotal role in a socio-political and agricultural strategy for counteracting malnutrition [6]. Specific and validated questionnaires could be useful to identify areas of weakness in people's understanding of healthy eating and provide data on the relationship between FK and dietary behaviour. The first FK questionnaire (FKQ) was drafted by Parmenter et al. [7] in a UK population of adults and then adapted and validated by other researchers in other specific populations. From the literature only one FKQ was validated in Sub Saharan area, specifically for Uganda population, by Richard Bukenya et al. [6]. The Uganda questionnaire is composed from different sections, that include: topics on food groups; selecting food; nutrition and disease relationship; food fortification in Uganda. Richard Bukenya et al. carry on a questionnaire validity in three studies: one about Assessment of Content Validity, the second study aims to assess the face validity and the last one study aims to construct validity and Test-Retest Reliability. Started from the Uganda FKQ the authors create a target specific population questionnaire addressed to childbearing age Tanzanian women.

Aim

The aim of this study was to describe the framework for the development and refinement of a questionnaire assessing food knowledge (FKQ), designed for Tanzanian women of childbearing age.

Methods

A FKQ was specifically designed for Tanzanian childbearing-age women, starting from a similar questionnaire previously proposed in Ugandan adults [6]. The FKQ was developed in English and then translated to Swahili by a forward- and back-translation process [8], and intended for survey administration by trained interviewers. A part from the initial section on sociodemographic characteristics (Section A), all the items have binary answers (right or wrong). Specifically, the FQK consists of 139 close-ended questions, divided into 4 sections: A) related to personal information about the sample; B) what people think about nutrition expert advices regarding healthy food habits; C) knowledge about macro-/micro-nutrients content in different food items; D) knowledge about the impact of nutrition on health: E) food safety and food contamination.
As a "proof of concept", the questionnaire was administered to a pilot sample of Tanzanian women aged 14-55 years, with no diagnosed disease, living in the region of Arusha (three urban and peri-urban wards). The interviews were conducted by two local well skilled and previously trained operators from the Nelson Mandela University (Tanzania). Each woman involved in the survey was informed about the study and asked to signed a consent. In order to provide a preliminary assessment of the items, individually and as part of latent constructs, exploratory statistical methods were applied. First of all, the items were recoded to correct/incorrect answers (the "not sure" option was coded together with incorrect options). All items were described by means of their frequency distributions, in order to identify those showing either no variability or an under-represented category (<10%), thus indicating a scarce discriminating power, or a high occurrence of missing answers (>10%) [9]. The remaining items of each section were then included in a Multiple Correspondence Analysis (Burt's method, representation in principal coordinates), in order to inspect reciprocal relations between items in each section; in addition, the optimal quantifications for each category of each item were computed. Globally, these analytical steps give a hint regarding the relation between items and latent constructs. All the analyses were performed in Stata 15. With regard to optimal quantifications, they were obtained by means of a user-written command.

Results

The FQK was administered to a pilot sample of 81 women of childbearing-age, selected in six different wards of the region of Arusha. Forty and 41 women were living in urban areas and in peri-urban areas, respectively; the whole sample was almost equally distributed in four age classes: 14-24 yrs (30%), 25-34 yrs (21%), 35-44 yrs (27%) and 45-55 yrs (22%); 71.6% of the women had at least one child (median 2, IQR:0-3 children). The education level for the majority of the sample (76.5%) was "primary or secondary school", 19.8% of the women reported a high school or higher degree education level and 3.7% did not have any education training. Through the A section of the FKQ the authors collected also some social and economic information such as the money management, detecting that 69.1% of the women declared to manage themselves money to buy food at the household level. Moreover, in the same section the possibility to follow a diet has been investigated and a 33.3% of the sample declared to follow a diet elaborated by a nutrition expert. The preliminary assessment of the items (identification of items with no variability, underrepresented categories or high missingness) led to discard several items (15 out of 19 for Section B, 39 out of 74 for Section C, 9 out of 16 for Section D, all items in Section E). Then, items were included in a series of Multiple Correspondence Analyses, which allowed to disclose a few items that behaved unexpectedly (e.g. showing a counter-relation with the other items), and eventually remove them. Optimal quantifications were applied on the final list of items; Guttman's Eta was sufficient (>0.2) for Section D and very close to the threshold for Section B, indicating good homogeneity of these two scales. In the transformation plots, all items appeared to be discriminating, although not all items were parallel.

Conclusion

According to the exploratory analyses presented above, several changes were introduced with the goal of refining FKQ. First of all, the inclusion of those questions that proved to be scarcely discriminant or had high missing rates was carefully evaluated, and some questions were dropped. Analogously, the behavior of each item within its construct, as showed by Multiple Correspondence Analysis, was inspected, and some questions were reworded or dropped. In addition, subsequently the scoring system for all the questions has been harmonized, adopting as possible answers “Agree”/ ”Disagree”/ ”Not Sure” and rewording questions accordingly (if necessary). Adopting such answer options was useful in reducing the feeling of “being judged” in respondents [10] besides, using the same scoring throughout the questionnaire led to a more homogeneous final construct. This pilot study allowed to collect 81 FKQ of childbearing age women living in the urban and peri-urban area in the district of Arusha and it could represent the start point to create a FKQ validable, population-specific.

References

[1] Development Initiatives, 2018. 2018 Global Nutrition Report: Shining a light to spur action on nutrition. Bristol, UK: Development Initiatives.

[2] Fadare O, Amare M, Mavrotas G et al. Mother’s nutrition-related knowledge and child nutrition outcomes: Empirical evidence from Nigeria. PLOS ONE, 2019 Feb; 14(4): e0215110.

[3] Conti MV, Campanaro A, Coccetti P et al. The potential role of Neglected and Underutilized Plant Species (NUS) in improving women’s empowerment and nutrition in Sub-Saharan area. Nutrition Reviews.

[4] Debela BL, Demmler KM, Rischke R et al. Maternal nutrition knowledge and child nutritional outcomes in urban Kenya. Appetite. 2017 Sep 1;116:518-526.

[5] Barbosa LB, Vasconcelos SM, Correia LO et al. Nutrition knowledge assessment studies in adults: a systematic review. Cien Saude Colet. 2016 Feb;21(2):449-62.

[6] Bukenya R, Ahmed A, Andrade JM et al. Validity and Reliability of General Nutrition Knowledge Questionnaire for Adults in Uganda. Nutrients. 2017, 9(2), 172.

[7] K Parmenter and J Wardle. Development of a general nutrition knowledge questionnaire for

Adults. European Journal of Clinical Nutrition. 1999, 53, 298±308

[8] https://www.who.int/substance_abuse/research_tools/translation/en/

[9] Kline P.Handbook of Psychological Testing.Abingdon: Routledge. 2013

[10] Trakman GL, Forsyth A, Russell H, et al. Developing and validating a nutrition knowledge questionnaire: key methods and considerations. Public Health Nutrition, 2017 Jul 24; 20(15):2670-2679

USE OF HEALTH INFORMATION FLOWS TO ESTIMATE THE PREVALENCE OF ALZHEIMER'S DISEASE AND OTHER DEMENTIAS IN SICILY

Maniscalco L.[1], Fantaci G.[2], Pollina Addario S.[2]., Scondotto S.[2], Monastero R.[1], Matranga D.[3]

[1] *Dipartimento di Biomedicina, Neuroscienze e Diagnostica Avanzata, Università degli Studi di Palermo, Italia; laura.maniscalco04@unipa.it*

[2] *Dipartimento per le Attività Sanitarie e Osservatorio Epidemiologico, Assessorato Regionale Salute, Palermo, Italia*

[3] *Dipartimento di Promozione della Salute, Materno-Infantile, Medicina Interna e Specialistica di eccellenza "G. D'Alessandro", Università degli Studi di Palermo, Italia; domenica.matranga@unipa.it*

Introduction

Dementia is one of the most frequent neurodegenerative diseases of the human brain worldwide. Dementias are characterized by the rapid progression of cognitive deficits, behavioral disorders and functional damage with the loss of autonomy and self-sufficiency [1,2]. In Italy, the number of patients with dementia is estimated at 1.2 million people, of whom about half are affected by Alzheimer's disease (AD), and the progressive increase in the elderly population will lead to an increase of the prevalence of elderly suffering from dementia [3]. Four medications are approved for the treatment of dementia by AIFA. Three of them belong to acetylcholinesterase inhibitors as donepezil, rivastigmine and galantamine whereas the fourth is memantine [4,5]. To date, no information is available about the prevalence of AD and other dementias in Sicily based on administrative flows.

Aim

To estimate the prevalence of individuals with AD and other dementias at 31/12/2017, aged 50 years or older residing in Sicily and assisted by the National Health Service (NHS).

Methods

The analyzed data were retrieved from an integrated archive of three administrative databases: the Hospital Discharge Records database, the Exemption Codes database, the database of outpatient dispensations of NHS-reimbursable drugs. The Residents' Registry was used to calculate the population residing in Sicily and the Register of causes of death was used (RENCAM archive) to assess the life status of each individual. Individuals aged 50 years or more, residing in Sicily that either i) have carried out at least one hospitalization for AD and other dementias (ICD-9 codes: 290.X excluded (290.8 e 290.9), 294.1, 294.10, 294.11, 331.0, 331.1X, 331.2, 331.82, 046.1) as main or secondary diagnosis (from the 2nd to the 5th diagnosis), or ii) have received at least one prescription of acetylcholinesterase inhibitors and/or memantine (ATC codes N06DA02, N06DA03 , N06DA04 , N06DX01) or have been found in the Exemption Codes database (exemption codes: 011.290.X, 011.291.1, 011.294.0, 029.331.0). The span period for data extraction was between January 2013 and December 2017.
Exclusion regarded patients (1) deceased as assessed in the RENCAM, (2) not residing or not beneficiary of the SSR, 3) aged less than 50 at diagnosis.
Statistical methods. The annual prevalence was computed by each age group and gender at 31 December 2017. The number of subjects suffering from AD and other dementias, residing in Sicily and assisted by the NHS, was divided by the population residing in Sicily and assisted by the NHS at the same date, including deceased and emigrated people. All analyses were performed using the SAS (version 9.4). A P-value <0.05 were significant.

Table 1. Prevalence for AD and other dementias by gender and age in Sicily, 31 December 2017

Age group	Male	Female	Total
50-54	0.05	0.04	0.04
55-59	0.12	0.10	0.11
60-64	0.28	0.23	0.25
65-69	0.61	0.58	0.59
70-74	1.13	1.30	1.22
75-79	2.34	2.96	2.68
80-84	3.68	5.06	4.49
85+	4.79	5.98	5.57
Total	1.05	1.55	1.32

Results

The number of selected individuals affected by dementia amounted to 47,009. After removing the deceased and non-residents and/or unassisted subjects we obtained the number of residents and assisted in the region of Sicily (n = 27,826). After selecting subjects with an age> 50 at the time of diagnosis, the number of individuals with AD and other dementias, residing in Sicily was estimated 27,577 at 31/12/2017. The 38% of subjects were extracted from the Hospital Discharge Records, the 41% from the database of outpatient dispensations of NHS-reimbursable drugs and only 6% from the Exemption Codes. The remaining subjects were extracted by the intersection among flows. Looking at the distribution of prevalent cases by age, it could be observed that there is a growing trend for both males and females by increasing age. Prevalence is higher for males compared to females from 50-54 years to 70-74 years. Afterwards, there is a reversal of this trend and females show higher prevalence than males (Table 1).

Conclusions

This study represents the first step of a more ambitious project, which is aimed to estimate also the incidence and the compliance to treatment in a real clinical setting. The progressive population aging has among its consequences the increased prevalence of dementias and other chronic diseases, with serious consequences on families, communities and health care systems. Estimating the prevalence is mandatory in order to achieve an accurate understanding of the overall dementia disease burden.

Bibliography

[1] Prince, M., Wilmo, A., Guerchet, M., Ali, G.-C. et al. World Alzheimer Report 2015, AA.VV. The Global Impact of Dementia An analysis of prevalence, incidence, cost and trends. London: Alzheimer's Disease International, 2015.

[2] GBD 2016 Dementia Collaborators. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol, 2019; 18: 88–106

[3] ISTAT, Anziani: le condizioni di salute in Italia e nell'Unione Europea (settembre 2017) - www.istat.it/it/archivio/203820

[4] Linee di indirizzo Nazionali sull'uso dei Sistemi Informativi per caratterizzare il fenomeno delle demenze. http://www.salute.gov.it/imgs/C_17_pagineAree_4893_listaFile_itemName_1_file.pdf Consultato il 7 Giugno 2019.

[5] Definizione di algoritmi per la identificazione dei pazienti con malattia di Alzheimer e altre forme di demenza. ATS della città metropolitana di Milano. https://www.ats-milano.it/Portale/Portals/0/AtsMilano_Documenti/A002-MS005%20Definizione%20Algoritmi%20Alzheimer%20e%20altre%20forme%20di%20demenza%20rev03_38fca72f-5e35-4499-a63c-e8c3676b14ae.pdf

TOTAL ANATOMICAL RECONSTRUCTION DURING PROSTATECTOMY: RISK OF POST-OPERATIVE INCONTINENCE PREDICTION

Stura Ilaria[1], Manfredi Matteo[2,3], Checcucci Enrico[2,3], Garrou Diletta[2,3], Amparore Daniele[2,3], Fiori Cristian[2,3], Porpiglia Franco[2,3] AND Migliaretti Giuseppe[1]

[1] *Dip. Scienze della Sanità Pubblica e Pediatriche, Università degli Studi di Torino, Italia*

[2] *Dip. Oncologia, Divisione di Urologia, Regione Gonzole 10, 10043, Orbassano, Italy*

[2] *Dip. Scienze Cliniche e Biologiche, Università degli Studi di Torino, Regione Gonzole 10, 10043, Orbassano, Italy*

Introduction

Prostate cancer is one of the most common solid organ malignancy in men western countries [1] and its incidence is expected to rise as the population ages and longevity increases. Radical prostatectomy (RP) is the gold standard among radical treatments in eligible patients. In particular, we focused on Total Anatomical Reconstruction (TAR) [2] during the robot assisted RP (RARP), which goal is to achieve complete cancer control while preserving functional outcomes in terms of urinary and sexual faculties.

Aim

The aim of this communication is to present a nomogram that evaluates which factors influence the postoperative continence recovery after three months.

Methods

The enrolment phase began in June 2013 and ended in May 2017; the follow-up period was formally closed in June 2018. The study was approved by the local ethics committee of San Luigi Gonzaga Hospital in Orbassano, Italy. Patients were prospectively included in the study with the following inclusion criteria: (1) localized PCa (clinical stages cT1-3, cN0, cM0); (2) indication for radical prostatectomy; (3) preoperative multiparametric prostate magnetic resonance imaging (mp-MRI). The exclusion criteria were: (1) contraindications for undergoing RARP; (2) neo-adjuvant hormone therapy; (3) anterior tumors with suspected extracapsular extension on mp-MRI.
Collected data included preoperative, intraoperative, postoperative and pathological variables. In particular, the nomogram was created considering a logistic regression model performed on pre- and intra-operative factors on the postoperative urinary continence recovery. All the analysis were performed with SAS ® Statistics Software.

Results

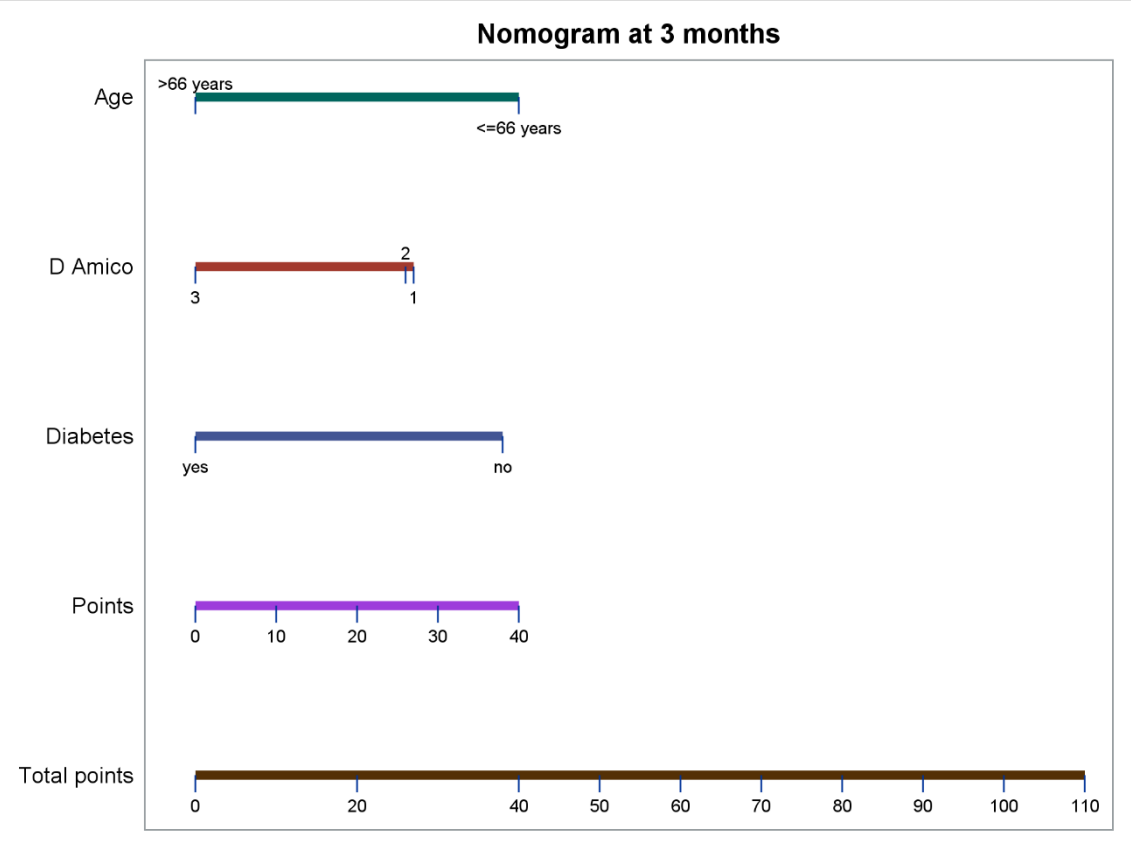
One-thousand-eight patients were enrolled in our study.
Logistic regression model shown that age, diabetes and D'Amico classification were the only significant pre-operative variables that are related to the continence recovery. A nomogram for the continence probability after three months were then generated (Fig. 1).

Conclusions

The strengths of the study are the large sample size considered in a limited period of time, the use of our standardized surgical technique, that allows to obtain successful results on post-operative continence from the beginning of the study, and the creation of a nomogram which clearly link pre-operative variables with

continence recovery. However, the nomogram requires external validation in a multicentre study to verify its applicability in a general context and to consider it for routine clinical implementation.
In the next future, an intra-operative use of bioengineering materials could improve the urinary continence recovery and better the presented nomogram.

Figure 1. *Nomogram for continence recovery after 3 months*



Bibliography

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, e A. Jemal, «Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries», *CA. Cancer J. Clin.*, vol. 68, n. 6, pagg. 394–424, 2018.

[2] F. Porpiglia *et al.*, «Total Anatomical Reconstruction During Robot-assisted Radical Prostatectomy: Implications on Early Recovery of Urinary Continence», *Eur. Urol.*, vol. 69, n. 3, pagg. 485–495, mar. 2016.

CONSUMO DI STATINE E SOPRAVVIVENZA IN UNA COORTE DI PAZIENTI CON NEOPLASIA DELLA MAMMELLA

Comoretto Rosanna Irene[1], Folino Franco[1], Zanier Loris[2], Gregori Dario[1], Canova Cristina[1]

[1] *Department of Cardiac, Thoracic and Vascular Sciences and Public Health, University of Padua, Padua, Italy*

[2] *Epidemiological Service, Health Directorate, Friuli Venezia-Giulia Region, Udine, Italy*

Introduzione

Le statine sono i farmaci più comunemente utilizzati per diminuire i livelli sierici di colesterolo, dimostrandosi in grado di ridurre la morbilità e la mortalità per patologie cardiovascolari [1–3]. Alcuni studi epidemiologici hanno dimostrato come i soggetti con patologia tumorale possono presentare livelli anomali di colesterolo [4] e che ci sia un'associazione positiva tra livelli elevati di colesterolo sierico e il rischio di sviluppare alcuni tipi di tumore [5–7]. Recenti revisioni sistematiche hanno inoltre suggerito che l'utilizzo di statine è associato ad un aumento dei tassi di sopravvivenza per diverse patologie tumorali,tra cui in particolare per donne con carcinoma mammario [8–12]. I risultati sono tuttavia in parte contrastanti a causa di diversi criteri di (i) selezione del campione, (ii) definizione dell'esposizione, (iii) definizione del tempo di follow-up, (iv) selezione delle covariate e potenziali fattori confondenti. Inoltre, l'informazione sulla stadiazione tumorale, la cui mancata considerazione può portare a stime distorte degli effetti delle statine sulla sopravvivenza, non risulta spesso disponibile. Infine, è necessario tenere in considerazione l'utilizzo concomitante di altre terapie (tra cui chemio- e radio-terapia) come fattori modificatori d'effetto, mentre l'eventuale confondimento da indicazione dovrebbe essere valutato attraverso la rilevazione della presenza di patologia cardiovascolare.

Obiettivi

Determinare se il trattamento con statine è associato alla riduzione della mortalità (*overall* e causa-specifica) in soggetti con prima diagnosi di tumore della mammella. Si vuole inoltre determinare se (i) alcune specifiche statine siano maggiormente efficaci nella prevenzione e nel trattamento delle neoplasie, (ii) la stadiazione del tumore si associ ad una differente efficacia terapeutica delle statine, (iii) l'associazione possa essere modificata dall'utilizzo concomitante di altre terapia (tra i quali farmaci antitumorali, chemio/radioterapia).

Metodi

Dal registro tumori della regione Friuli-Venezia Giulia possono essere identificati i casi incidenti di tumore della mammella nel periodo 1995-2013. Da tale registro è inoltre possibile avere a disposizione informazioni dettagliate su localizzazione, istologia, stadiazione e morfologia/topografia della patologia tumorale. Attraverso record-linkage utilizzando una chiave paziente univoca anonimizzata, per questi soggetti sono state raccolte le informazioni relative a: ricoveri ospedalieri (precedenti e successivi la diagnosi di tumore), farmaci assunti (sia dai registri della farmaceutica territoriale, sia dai registri della distribuzione intra-ospedaliera), somministrazione di chemioterapici ed esecuzione di radioterapia, prestazioni ambulatoriali ed esami di laboratorio (con i relativi risultati), mortalità (con le relative cause). L'esposizione a statine viene definita come variabile tempo-dipendente e nell'analisi verranno distinti i soggetti incidenti e prevalenti per uso di statine (rispettivamente dopo la diagnosi di tumore, oppure *any time*). Inoltre, la presenza di patologia diabetica e/o cardiovascolare (deducibili da registri di patologia implementati in regione e dai flussi della specialistica ambulatoriale e degli esami di laboratorio) verrà considerata come covariata e/o modificatore d'effetto. Modelli di regressione di Cox verranno utilizzati per stimare gli hazard ratio (HR) con i relativi

intervalli di confidenza al 95% per l'associazione tra l'utilizzo di statine e la mortalità overall e causa specifica.

Risultati

Dal registro tumori sono state identificate 22,090 donne con prima diagnosi di tumore della mammella, di cui circa il 40% risultano decedute al 31/12/2017. Da una preliminare analisi, 4,727 soggetti della coorte risultano essere utilizzatori incidenti di statine (prima prescrizione dopo la diagnosi di tumore).

Conclusioni

Rispetto a quanto già riportato in letteratura, l'utilizzo del registro tumori per l'identificazione della coorte in studio, oltre che la possibilità di record-linkage con registri di patologie cardio-metaboliche e la cartella oncologica con informazione sui trattamenti chemio-radioterapici consentirà la conduzione di analisi in sottogruppi identificati dalla stadiazione tumorale, il possibile effetto di modificazione de ll'utilizzo concomitante di altre terapia, oltre che la possibilità di valutare eventuali fattori confondenti che non sono completamente tracciabili e/o identificabili dai flussi amministrativi.

Bibliografia

[1] Colhoun HM, Betteridge DJ, Durrington PN, Hitman GA, Neil HAW, Livingstone SJ, et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomised placebo-controlled trial. Lancet. 2004;364:685–96.

[2]Collins R, Armitage J, Parish S, Sleigh P, Peto R, Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol-lowering with simvastatin in 5963 people with diabetes: a randomised placebo-controlled trial. Lancet. 2003;361:2005–16.

[3] Cholesterol Treatment Trialists' (CTT) Collaborators, Mihaylova B, Emberson J, Blackwell L, Keech A, Simes J, et al. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. Lancet. 2012;380:581–90.

[4] Cantafora A, Blotta I. Neutral lipids production, transport, utilization. Anticancer Res. 1996;16:1441–9.

[5] Shafique K, McLoone P, Qureshi K, Leung H, Hart C, Morrison DS. Cholesterol and the risk of grade-specific prostate cancer incidence: evidence from two large prospective cohort studies with up to 37 years' follow up. BMC Cancer. 2012;12:25.

[6] Pelton K, Freeman MR, Solomon KR. Cholesterol and Prostate Cancer. Curr Opin Pharmacol. 2012;12:751–9.

[7] Allott EH, Howard LE, Cooperberg MR, Kane CJ, Aronson WJ, Terris MK, et al. Serum lipid profile and risk of prostate cancer recurrence: Results from the SEARCH database. Cancer Epidemiol Biomarkers Prev. 2014;23:2349–56.

[8] Wu BU, Chang J, Jeon CY, Pandol SJ, Huang B, Ngor EW, et al. Impact of statin use on survival in patients undergoing resection for early-stage pancreatic cancer. Am J Gastroenterol. 2015;110:1233–9.

[9] Mansourian M, Haghjooy-Javanmard S, Eshraghi A, Vaseghi G, Hayatshahi A, Thomas J. Statins Use and Risk of Breast Cancer Recurrence and Death: A Systematic Review and Meta-Analysis of Observational Studies. Journal of Pharmacy & Pharmaceutical Sciences. 2016;19:72.

[10] Manthravadi S, Shrestha A, Madhusudhana S. Impact of statin use on cancer recurrence and mortality in breast cancer: A systematic review and meta-analysis. Int J Cancer. 2016;139:1281–8.

[11] Mei Z, Liang M, Li L, Zhang Y, Wang Q, Yang W. Effects of statins on cancer mortality and progression: A systematic review and meta-analysis of 95 cohorts including 1,111,407 individuals: Statins on cancer mortality and progression. International Journal of Cancer. 2017;140:1068–81.

[12] Liu B, Yi Z, Guan X, Zeng Y-X, Ma F. The relationship between statins and breast cancer prognosis varies by statin type and exposure time: a meta-analysis. Breast Cancer Res Treat. 2017;164:1–11.

VULNERABILITY AND SOCIAL CONNECTEDNESS IN YOUNG MIGRANTS: AN ANALYSIS BASED ON BAYESIAN NETWORKS

Charrier Lorena[1], Berchialla Paola[2], Borraccino Alberto[1], Lemma Patrizia[1], Sciannameo Veronica[3], Dalmasso Paola[1]

[1] *Dipartimento di Scienze della Sanità Pubblica e Pediatriche, Università di Torino*

[2] *Dipartimento di Scienze Cliniche e Biologiche, Università di Torino*

[3]*Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac Thoracic Vascular Sciences and Public Health, University of Padova*

Background

Most young people in developed countries grow up healthy and safe, receiving the assistance they need from families and from the communities during the transition from adolescence to adulthood. However, some of them are called to face barriers that put them at risk of developing dissatisfaction with life and behaviours that could have a negative impact on health outcomes [1]. Those young people can be identified as vulnerable. Among the known risk factors, belonging to an ethnic minority group is one of the most representative condition of vulnerability and it has also been associated to many health problems. More recently, it has been shown that having low socio-economic status is also frequently associated with the above-mentioned vulnerable condition [2,3]. Conversely, the social connectedness is recognized to be a potential protective factor among vulnerable youths [4].

Aim

The aim of this work is to study whether specific types of connectedness (family, school, or peers) could act as a potential protective factor associated to wellbeing in vulnerable migrant youths.

Methods

Data were collected as part of the 2014 Health Behaviour in School-Aged Children study in Italy. A national representative sample of students aged 11, 13, and 15 years were recruited: 40,224 hosting students and 7,175 immigrants. Based on mother's country of birth, immigrants were categorized as non-vulnerable youths (from western countries), and more vulnerable adolescents (from non-western countries) [5]. Three sets of multidimensional scales of support were used to measure connectedness to family, peers and school, and perceived wellbeing was described as the occurrence of health complaints and the level of life satisfaction [6,7]. Bayesian Networks (BNs) were used to make inference on the joint conditional distribution with the aim to get the posterior conditional probability of specific outcomes after observing evidence, i.e. conditioning on the values of some nodes (ethnic background, social support and socio-economic status).The PC stable algorithm [8] with a 100- fold bootstrap was to learn the structure of the BN. PC stable is a constrained-based learning algorithm, which uses mutual information tests to estimate whether conditional independence between variables hold.The 100 BNs learned were then averaged to obtain a more robust model. The averaged BN was created using the arcs present in at least 90% of the BNs. Finally, a Maximum a Posteriori estimation computed the set of parameters of the conditional probability distribution at each node. The learning algorithms did not use prior knowledge nor constraints.

Results

Adolescent immigrants from Eastern Europe and from non-Western/non-European countries are significantly more likely to belong to a low SES category when compared to native or immigrants from Western countries. Similarly, these EE+non-W/non-EU immigrants show a significantly lower life satisfaction, higher health complaints and lower family, school and peer support as compared to native/Western immigrants (all $p<0.001$). Family support turned out to reduce the risk of a low life satisfaction of about 67% (RR = 0.33, with 99% probability of a RR < 1) and the risk of high health complaints of about 59% (RR = 0.41, with 99% probability of a RR < 1). Comparing adolescents who are vulnerable for both ethnic background and SES status with adolescents who are vulnerable for ethnic background only, it turned out that family support has still a protective effect on both low life satisfaction (RR = 0.41, with 99% probability of a RR < 1) and high health complaints (RR = 0.43, with 99% probability of a RR < 1. When comparing vulnerable adolescents with social support with adolescents who are vulnerable for SES only or are not vulnerable at all, we still find a protective effect of social support for both low life satisfaction and high risk complaints.

Conclusion

Our results showed that different sources of support can be acknowledged as protective factors on individual psychosocial stress and perceived wellbeing. We found that family and school support have the potential to mitigate the negative outcomes on wellbeing adolescents who are vulnerable due to their socioeconomic status (low SES) and ethnic background (immigrants from Eastern Europe and from non-Western/non-European countries).

References

[1] Fernandes-Alcantara, A.L. (2014). Vulnerable youth: Background and policies. Washington, DC: Congressional Research Service.

[2] Mood, C, Jonsson, JO, Brolin Låftman S Immigrant Integration and Youth Mental Health in Four European Countries. European Sociological Review, 2016,32(6): 716-729

[3] Belhadj Kouider E, Koglin U, Petermann F. Emotional and behavioral problems in migrant children and adolescents in Europe: a systematic review. Eur Child Adolesc Psychiatry. 2014 Jun;23(6) 373-391.

[4] Ewell Foster C, Horwitz A, Thomas A, et al. Connectedness to family, school, peers, and community in socially vulnerable adolescents, Children and Youth Services Review, 10.1016/j.childyouth.2017.08.011, 81, (321-331), (2017)

[5] Dalmasso P, Borraccino A, Lazzeri G, et al. Being a Young Migrant in Italy: The Effect of Perceived Social Support in Adolescence. J Immigr Minor Health. 2018 Oct;20(5) 1044-1052

[6] Currie, C, Health Behaviour in School-aged Children Currie, 2010, Health Behaviour in School-aged Children

[7] Ravens-Sieberer, U, Kokonyei, G.,Thomas, C. (2004). School and health. in Currie, C. et al. (eds.), Young People's Health in Context. Health Behaviour in School-aged Children Study: International Report from the 2001/2002 Survey. WHO Regional Office for Europe, Copenhagen

[8] Colombo D, Maathuis MH Order-independent constraint-based causal structure learning. J. Mach. Learn. Res., 2014,15:3741–3782

REQUIEM PER IL "ONE SAMPLE ONE STAGE DESIGN (A'Hern's design)".

Zolin Anna[1], Tresoldi Claudio[1], Antonelli Paolo[2], Cesana Bruno Mario[3]

1 Dipartimento di Scienze Cliniche e di Comunità, Università degli Studi di Milano

2 già professore a contratto Università di Brescia

3 già professore associato di Statistica Medica Università di Brescia, professore a contratto: Dipartimento di Scienze Cliniche e di Comunità, Unità di Statistica Medica, Biometria e Bioinformatica "Giulio A. Maccacaro", Facoltà di Medicina e Chirurgia, Università di Milano.

Introduzione

Nel contesto dei disegni di Fase II miranti alla valutazione della potenziale efficacia di un farmaco, da definire successivamente nell'ambito degli studi comparativi della Fase III, un disegno ampiamente utilizzato è quello a uno stadio. Infatti, una ricerca su ClinicalTrial.gov [1,2] effettuata con "Phase II" come unico termine ha prodotto 63.109 Studi e 13,884 (22.0%) unendo anche "one stage".
Questo disegno prevede il calcolo della numerosità dell'unico campione fissando i valori di probabilità di efficacia non interessante (π_0) e interessante (π_1) per l'ulteriore sviluppo del farmaco, del livello di significatività (α=0.05, usualmente o anche α=0.10), della potenza (1-β=0.80, usualmente) e del test per saggiare il confronto di una proporzione verso un valore atteso (test binomiale esatto o test z approssimato). In accordo a questo disegno, si effettua una sola analisi statistica finale quando tutti i risultati saranno disponibili.
Storicamente, il disegno di Gehan[3] è il primo modello per tale fase, dimensionato con un numero di pazienti tale che l'ottenere nessun successo terapeutico costituisse un risultato statisticamente significativo ($P \leq 0.05$ o $P \leq 0.10$, calcolato in accordo a una distribuzione binomiale) che permetteva di rifiutare l'ipotesi nulla relativa a una probabilità attesa di efficacia molto piccola (0.10, ad esempio). Viceversa, nel caso di un risultato statisticamente significativo che rifiutava la probabilità attesa di efficacia molto piccola, si suggeriva di aumentare la dimensione campionaria per ottenere una adeguata precisione delle stime.
Ricordiamo per una rassegna dei disegni di Fase II, i disegni Bayesiani a più stadi proposti da Herson [4] e la "One-Sample Multiple Testing Procedure" proposta da Fleming[5] riprendendo un lavoro di Schultz *et al*.[6] che avevano formalizzato le proprietà operative dei test sequenziali chiusi proposti da Aroian[7].
La procedura di Fleming [5] è basata su due diversi valori di probabilità: π_0, che se fosse vera sicuramente non comporterebbe un'ulteriore sviluppo del farmaco e il più piccolo possibile (π_A) che se fosse vera determinerebbe l'ulteriore valutazione dell'efficacia terapeutica del farmaco. Formalmente, le ipotesi unilaterali da saggiare sono: $H_0: \pi \leq \pi_0$ verso $H_A: \pi \geq \pi_A$ e il calcolo della dimensione campionaria per il disegno a uno stadio è basato sull'approssimazione Gaussiana. Tuttavia, nel prosieguo il lavoro [5] sottolinea la necessità di pianificare un disegno a due o tre stadi unitamente alla metodologia statistica per il calcolo della dimensione campionaria basata sulla distribuzione binomiale e ai valori critici di successo per accedere agli stadi successivi o per interrompere lo studio senza poter rifiutare l'ipotesi nulla. Infatti, non è eticamente accettabile esporre pazienti a un trattamento scarsamente efficace (interruzione per futilità) e non trattare altri pazienti con un farmaco promettente se tale risultato fosse acquisito il più precocemente possibile (interruzione per efficacia).
Ricordiamo doverosamente anche il disegno a due stadi ("optimum" e "minimax") proposto da Simon [8], esteso successivamente a un approccio Bayesiano [9]. Si rimanda a Mariani e Marubini [10] e al Capitolo 7 "Overview of Phase II Clinical Trials" di Green [11] per una rassegna aggiornata al 2010-2011 dell'evoluzione metodologica dei disegni della Fase II.

Focalizziamo il nostro interesse sul disegno a un singolo stadio noto come disegno di A'Hern [12] che è analizzato dopo aver trattato tutti i pazienti richiesti, sulle problematiche relative al calcolo della dimensione campionaria basata sul test binomiale e le criticità connesse con la scelta dell'uso del test binomiale e della procedura di stima intervallare a esso connessa.
Si può presumere che l'aver illustrato un disegno a uno stadio nella prima parte dell'articolo di Fleming [7] e le problematiche statistiche insite nel calcolo della dimensione campionaria e dei relativi cut-off per l'arresto per efficacia o per futilità dei disegni a due o più stadi abbia fatto ritenere che il disegno a uno stadio potesse costituire una valida alternativa, soprattutto se il calcolo della dimensione campionaria fosse basato sul test binomiale esatto.

Metodi

Calcolo della dimensione campionaria per il test binomiale unilaterale. Il calcolo della dimensione campionaria è basato sul verificarsi congiunto di due condizioni: rifiutare H_0 sotto H0 con probabilità minore o uguale ad α e rifiutare H_0 sotto H_A con probabilità maggiore o uguale a $(1-\beta)$.

Indicando con K il cut-off, si ha:

$$1)-\text{per}\ \ \pi_0<\pi_A:\left(\sum_{k=K}^{N}\binom{N}{k}\pi_0^k\left(1-\pi_0\right)^{N-k}\right)\le\alpha\ \ \text{et}\ \ \left(\sum_{k=K}^{N}\binom{N}{k}\pi_A^k\left(1-\pi_A\right)^{N-k}\right)\ge 1-\beta$$

$$2)-\text{per}\ \ \pi_0>\pi_A:\left(\sum_{k=0}^{K}\binom{N}{k}\pi_0^k\left(1-\pi_0\right)^{N-k}\right)\le\alpha\ \ \text{et}\ \ \left(\sum_{k=0}^{N}\binom{N}{k}\pi_A^k\left(1-\pi_A\right)^{N-k}\right)\ge 1-\beta$$

Quanto sopra considera la situazione della probabilità di successo (1) o di insuccesso (2) con la probabilità dell'ipotesi alternativa (H_0) rispettivamente maggiore e minore. Stante che il valore della funzione cumulativa della distribuzione Binomiale con parametri N and π:

$$P=\left(X\le k\right)=\sum_{i=0}^{k}\binom{N}{k}\pi^i\left(1-\pi\right)^{N-i}$$

è uguale a quello della funzione cumulativa della distribuzione Beta ("regularized incomplete beta function") data da $\text{cdf}_{\text{BETA}}(1-\pi,\alpha_1,\beta_1)$ con parametri α_1=N-k e β_1=k+1, le suddette espressioni possono facilmente essere sostituite da quelle equivalenti della distribuzione Beta. Inoltre, un'alternativa formulazione della suddetta uguaglianza è data da: $1-\text{cdf}_{\text{BETA}}(\pi,\alpha_2,\beta_2)$, con parametri α_2=k+1 e β_2=N–k.

Calcolo della dimensione campionaria per il test binomiale bilaterale. Armitage *et al.*[13] riportano: "for a two-sided test we could add the probabilities of deviations at least as large as that observed, in the other direction". Così per esempio, se il numero di successi (x) è maggiore o uguale a quello atteso sotto H_0 ($N\pi_0$), si deve sommare il valore di probabilità calcolato da x a N (incluso) alla coda destra della distribuzione a quello calcolato alla coda sinistra da 0 a x'=$2N\pi_0$–x; infine, x' deve essere incluso nei calcoli se è intero o deve essere incluso l'intero ottenuto arrotondando x' per difetto.

Formalmente, con x successi su N prove, il livello di significatività del test bilaterale è calcolato:

$$\text{per } x \geq N\pi_0 : P(X \geq x) + P(X \leq x') = \sum_{k=x}^{N} \binom{N}{k} \pi_0^k (1-\pi_0)^{N-k} + \sum_{k=0}^{x'} \binom{N}{k} \pi_0^k (1-\pi_0)^{N-k}$$

$$\text{per } x < N\pi_0 : P(X \leq x) + P(X \geq x') = \sum_{k=0}^{x} \binom{N}{k} \pi_0^k (1-\pi_0)^{n-k} + \sum_{k=x'}^{N} \binom{N}{k} \pi_0^k (1-\pi_0)^{N-k}$$

Una procedura alternativa somma a entrambe le code i valori di probabilità minori di quello effettivamente calcolato (SPVM, Small P-Values Method); tralasciamo l'approccio di raddoppiare il minore dei valori calcolati alle due code poco raccomandabile stante l'asimmetria della distribuzione Binomiale.

La potenza è calcolata:

$$1) - \text{per } \pi_0 < \pi_A : \left(\sum_{k=K}^{N} \binom{N}{k} \pi_A^k (1-\pi_A)^{N-k} \right) + \left(\sum_{0=k}^{K^*} \binom{N}{k} \pi_A^k (1-\pi_A)^{N-k} \right) = 1-\beta$$

$$2) - \text{per } \pi_0 > \pi_A : \left(\sum_{k=0}^{K} \binom{N}{k} \pi_A^k (1-\pi_A)^{N-k} \right) + \left(\sum_{k=K^*}^{N} \binom{N}{k} \pi_A^k (1-\pi_A)^{N-k} \right) = 1-\beta$$

Dove $K^*=2N\pi_0 - x$ per il metodo della somma delle due aree e corrisponde, per SPVM, all'ultimo numero di successi (partendo da zero o da N) che ha una probabilità minore di quella ottenuta col numero di successi campionario.
L'interesse del ricercatore è quello di rifiutare l'ipotesi nulla e di concludere per il corretto verso della differenza tra la probabilità prevista e il valore atteso nel caso di un solo gruppo e tra quella del trattamento sperimentale e quella dello standard nel caso di due gruppi, evitando di commettere l'errore di Tipo III o "l'errore del segno", definito γ e usualmente trascurabile, che corrisponde all'area alla coda opposta a quella del corretto verso della differenza su cui si focalizza l'interesse del ricercatore. Non ritenendo appropriato e perfino paradossale includere l'area dell'errore γ nel calcolo della "potenza bilaterale", non seguiamo l'approccio di sommare entrambe le aree alle due code come fa PASS[14]. Tra l'altro PASS[14] segue l'approccio, non condivisibile, per il test bilaterale di sommare le aree alle due code solo quando entrambe sono ≤ 0.025, riportando ingiustificatamente numerosità maggiori.
Nell'articolo di A'Hern[12] si evince che l'intervallo di confidenza (IC) usato è quello proposto da Clopper and Pearson[15]. Si ricorda che tale IC può essere ottenuto direttamente grazie alla relazione tra le funzioni cumulative della distribuzione Binomiale e F attraverso i termini della funzione Beta incompleta in accordo a Kendall[16] e alla notazione semplificata di Miettinen[17]. Per il limite inferiore (L_L) e quello superiore (L_U), abbiamo:

$$L_L = \frac{k}{\left[k + (N-k+1)\cdot F_L\right]}; \quad L_U = \frac{k+1}{\left[(k+1) + (N-k)\cdot F_U\right]}$$

Essendo $F_L = F_{1-\alpha/2;2(N-k+1);2k}$, e $F_U = F_{\alpha/2;2(N-k);2(k+1)}$ gli $\alpha/2$ e $1-\alpha/2$ quantili della distribuzione F con gradi di libertà al numeratore (ν_1) pari rispettivamente a 2(N-k+1) e 2(N-k) e al denominatore (ν_2) pari a 2k e 2(k+1). In

questo caso, N e k corrispondono alla dimensione campionaria e al numero di successi. Infine, per i casi estremi k=0 e k=N, si ha:

$$k = 0 : L_L = 0;\ \ L_U = 1 - \left(\frac{\alpha}{2}\right)^{\frac{1}{N}} \text{ e } \ k = n : L_L = \left(\frac{\alpha}{2}\right)^{\frac{1}{N}};\ L_U = 1$$

Le suddette formule non sono semplici; infatti quella del limite inferiore è sbagliata in Armitage e Berry [18], in Armitage *et al*. [13] e in Cesana *et al*. [19] che l'hanno ripresa. Julious [20] e Julious e Campbell [21] riportano quelle meno complicate basate sulla distribuzione Beta.

Risultati

Nella Tabella 1 dell'articolo di A'Hern [12] sono riportate delle numerosità di poche unità che anche se correttamente calcolate per delle enormi differenze non clinicamente plausibili rischiano di essere fuorvianti per il ricercatore clinico. Inoltre, sono presenti delle numerosità errate o per problemi di arrotondamento o perché calcolate per il test zeta approssimato, stante l'insuccesso della funzione utilizzata ("BINOMDIST" di Excel®).
In Tabella 1 si riportano i valori corretti (N-corretto) unitamente a quelli non corretti (N-A'Hern) e alle rispettive differenze tralasciando le dimensioni campionarie calcolate per α=0.01 e quelle inferiori a 3.
Si noti che la Table 1[12] potrebbe anche essere utilizzata per un disegno di non-inferiorità con probabilità di successo ponendo $\pi_0=\pi_1$ per la condizione di equipoise e il margine di non inferiorità pari alla differenza tabulata. Poi, nel caso della probabilità di insuccesso, π_0 e π_1 devono invertire la loro posizione in quanto la specularità tra 0.10 e 0.05 non è quella tra 0.95 e 0.90, ma bensì tra 0.90 e 0.95; in generale, tra $\pi_0>\pi_1$ e $(1-\pi_0)<(1-\pi_1)$.
Questo aspetto è particolarmente importante in fase di pianificazione di uno studio in quanto i valori di cut-off che permetterebbero di interromperlo per efficacia o futilità sono nettamente diversi. Ad esempio, per saggiare π_0=0.95 vs. π_1=0.90, con α=0.025 and 1-β=0.80 o π_0=0.05 e π_1=0.10, la dimensione campionaria è sempre pari a 213, ma il cut-off per la superiorità è 195 invece di 18. È evidente come il primo caso sia assolutamente inefficiente nei termini di pazienti arruolati con conseguenti importanti implicazioni etiche, logistiche, temporali ed economiche. Quindi l'obiettivo (insuccesso o successo) deve essere appropriatamente scelto.
Considerando 171 situazioni (da π_0=0.05 a 0.90 con passo 0.05 con $\pi_1=\pi_0$+0.05 a 0.95) e tralasciando le numerosità con N<5 in 26 casi, il calcolo della dimensione campionaria per il test bilaterale mediante SPVM richiede in media meno soggetti rispetto al metodo della somma delle due aree (differenza: media=-1.9, mediana=0, minimo=-22, massimo=+4); inoltre, il test bilaterale richiede in media più pazienti rispetto a quello unilaterale (differenza: medie=23.9;25.7, mediane=5;7, minimi=0;0, massimi=174;188).
Inoltre il test binomiale unilaterale richiede mediamente più soggetti rispetto al test approssimato senza la correzione di continuità (differenza: media=3.5, mediana=3, minimo=-4, massimo=20), mentre è mediamente simile nel caso della correzione per la continuità.
Inoltre il test binomiale bilaterale richiede mediamente più soggetti rispetto al test approssimato bilaterale senza la correzione di continuità (differenza: medie=27.4;29.3, mediane=8;11, minimi=1;3, massimi=181;192). Le numerosità sono similari, nel caso del test approssimato con la correzione per la continuità.

Conclusioni

Nonostante che il test binomiale richieda maggiori numerosità, è stato anche proposto da Chu e Cole [22] di incrementare la numerosità fino ai valori da cui la funzione potenza con un andamento a "denti di sega" Cesana *et al*.[19] rimane stabilmente sopra la soglia richiesta. Tuttavia, tale approccio non è stato considerato eticamente accettabile da Cesana e Antonelli [23] ribadendo che il calcolo della dimensione campionaria deve basarsi su principi etici [24]. Piuttosto, risulta condivisibile la proposta di Khan *et al*.[25] di accettare dei valori di significatività e di potenza che siano molto vicini ai valori convenzionalmente accettati al fine di ottenere delle riduzioni della dimensione campionaria, particolrmente negli studi che hanno una limitata numerosità (diciamo, <50).
Si deve puntualizzare che l'impiego della distribuzione binomiale per un test esatto "chiuso" può in parte vicariare una procedura a due stadi dal momento che lo studio può essere interrotto rifiutando l'ipotesi nulla non appena il numero di successi ha superato il valore critico o, viceversa per una situazione di futilità, quando il numero di insuccessi è tale che non sarà più possibile ottenere sui restanti pazienti i successi necessari per rifiutare H_0.
Ad esempio, per uno studio pianificato con α=0.025, 1-β=0.80, π_0=0.20, e π_1=0.40, si richiedono 41 (N) pazienti con 14 successi per rifiutare π_0 (effettivamente: α=0.0241 e 1-β=0.8220). Allora, se 28 insuccessi sono osservati su, diciamo, i primi 29 o 30 pazienti, non sarà più possibile ottenere i 14 successi necessari for rifiutare H_0 dai rimanenti 12 o 11 pazienti. Quindi, lo studio potrà essere chiuso per futilità o la dimensione campionaria potrà essere ricalcolata (disegno adattativo) se è clinicamente accettabile. Ovviamente, l'aver ottenuto 14 successi prima di aver arruolato tutti i pazienti richiesti può portare a un precoce rifiuto di H_0 al prezzo di una minore precisione delle stime.
Tuttavia, è praticamente sicuro che questa procedura di "test troncato" sia meno efficiente di un reale disegno a due stadi che ha un minor numero medio atteso di pazienti da arruolare.
In ogni caso, il disegno a un solo stadio non deve essere effettuato, stante che non permette un soddisfacente monitoraggio dello studio basato su formali regole di arresto.
In conclusione, abbiamo corretto le numerosità campionarie della Table 1[12], abbiamo evidenziato come questa tabella potrebbe essere utilizzata anche per disegni di non inferiorità per il successo e per l'insuccesso, abbiamo confermato la conservatività del test binomiale [26] che può quantificarsi in circa il 3-5% per l'errore di Tipo I e dello 0.2-0.8% per la potenza (dati da studi di simulazione non riportati nel dettaglio).
Consegue che il test binomiale nell'ambito di questo disegno non sia da raccomandarsi, così come l'associata procedura di stima intervallare di Clopper-Pearson, stante la scarsa coverage [26]. Anche se i metodi statistici esatti sono generalmente preferibili [27], particolarmente quando un p-value esatto può essere calcolato in modo relativamente facile senza costi proibitivi [28], si deve tener conto che, nella ricerca biomedica i costi sono soprattutto rappresentati da esseri umani e, quindi, le dimensioni campionarie devono essere le più parsimoniose possibili.

Tabella 1.

Probabilità		α=0.05; 1−β=0.8			α=0.05; 1−β=0.9		
π_0	π_1	Diff.	N corretto	N A'Hern	Diff.	N corretto	N A'Hern
0.05	0.15	0	52	52	1	77	76
0.05	0.35	1	12	11	0	14	14
0.15	0.20	5	360	355	0	500	500
0.15	0.40	1	22	21	0	27	27
0.30	0.40	3	144	141	0	193	193
0.50	0.55	2	620	618	0	866	866
0.60	0.80	3	36	33	0	45	45
0.70	0.85	0	49	49	4	69	65
0.80	0.85	6	365	359	6	501	495

Bibliografia

[1]ClinicalTrial.govhttps://clinicaltrials.gov/ct2/results?recrs=&cond=&term=phase+ii&cntry=&state=&city=&dist= (accesso il 4 Giugno 2019)

[2] ClinicalTrial.govhttps://clinicaltrials.gov/ct2/results?cond=&term=phase+II+one+stage&cntry=&state=&city=&dist=&Search=Search&flds=aby (accesso il 4 Giugno 2019).

[3] Gehan E.A. The determination of the number of patients required in a follow-up trial of a new chemotherapeutic agent. Journal of Chronic Diseases 1961;13:346-353.

[4] Herson J. Predictive Probability Early Termination Plans For Phase II Clinical Trials. Biometrics, 1979;35(4):775-783.

[5] Fleming T.R. One-Sample Multiple Testing Procedure for Phase II Clinical Trials. Biometrics 1982;38(1):143-151.

[6] Schultz J.R, Nichol F.R., Elfring G.L., et al. Multiple stage procedures for drug screening. Biometrics 1973;29: 293-300.

[7] Aroian L.A. Sequential analysis, direct method. Technometrics, 1968;10:125-32.

[8] Simon R. Optimal two-stage designs for phase II clinical trials. Controlled Clinical Trials. 1989;10:1-10.

[9] Thall P.F., Simon R. Practical Bayesian Guidelines for Phase IIB Clinical Trials Biometrics, 1994;50(2):337-349.

[10] Mariani L., Marubini E. Design and Analysis of Phase II Cancer Trials: A Review of Statistical Methods and Guidelines for Medical Researchers International Statistical Review / Revue Internationale de Statistique. 1996;64(1):61-88.

[11] Green S. Overview of Phase II Clinical Trials, Chapter 7 in Crowley J., Hoering A. (eds.), Handbook of Statistics in Clinical Oncology. 3rd 2012 CRC Press Ed.Taylor & Francis Group Boca Raton.

[12] A'Hern R.P. Sample size tables for exact single-stage phase II designs. Statistic in Medicine 2001; 20:859-66.

[13] Armitage P, Berry G. Matthews JNS. Statistical Methods in Medical Research. 4rth ed. Blackwell Scientific Publications: Oxford, 2002. pages 113-114, 117.

[14] Hintze, J. PASS 13. Chapter 120 Single-Stage Phase II Clinical Trials. NCSS, LLC. Kaysville, Utah, USA, 2014. www.ncss.com.

[15] Clopper C., Pearson E.S. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 1934;26:404–13.

[16] Kendall M.G. The Advanced Theory of Statistics. Vol. 1. J.B. Lippincott Co.: London, 1943, page 120.

[17] Miettinen O.S. Estimation of Relative Risk from Individually Matched Series. Biometrics 1970;26:75-86.

[18] Armitage P, Berry G. Statistical Methods in Medical Research. 2nd ed. Blackwell Scientific Publications: Oxford, 1987. page 117

[19] Cesana B.M., Reina G., Marubini E. Sample Size for Testing a Proportion in Clinical Trials: A "Two-Step" Procedure Combining Power and Confidence Interval Expected Width. The American Statistician 2001;55: 288-92.

[20] Julious SA. Letter to the editor Two-sided confidence intervals for the single proportion: comparison of seven methods by Robert G. Newcombe. Statistics in Medicine 1998;17:857–872. 20. Statistics in Medicine 2005;24:3383–3384

[21] Julious SA, Campbell MJ. Tutorial in biostatistics: sample sizes for parallel group clinical trials with binary data. Statistic in Medicine 2012; 31:2904–36.

[22] Chu H., Cole S.R. Sample size calculation using exact methods in diagnostic test studies. Journal of Clinical Epidemiology 2007;60:1201-2.

[23] Cesana B.M., Antonelli P. Sample size calculations need to be adequate and parsimonious. Journal of Clinical Epidemiology(2018) doi: https://doi.org/10.1016/j.jclinepi.2018.11.018

[24] Cesana B.M., Antonelli P. Sample size calculations in clinical research should also be based on ethical principles. Trials 2016;17:149 DOI 10.1186/s13063-016-1277-5.

[25] Khan I., Sarker S-J, Hackshaw A. Smaller sample sizes for phase II trials based on exact tests with actual error rates by trading-off their nominal levels of significance and power British Journal of Cancer. 2012;107:1801–9.

[26] Agresti A., Coull B.A. Approximate is Better than “Exact” for Interval Estimation of Binomial Proportions. The American Statistician 1998;52:2,119-26.

[27] Mehta C.R., Patel N.R. StatXact 3 for Windows: Statistical software for exact nonparametric inference: User manual. Cambridge, MA: Cytel Software Corporation; 1995.

[28] Berger V.W. An empirical demonstration of the need for exact tests. Journal of Modern Applied Statistical Methods 2017;16:34-50.

CHARACTERISING RATINGS OF 12 COMMONLY CONSUMED FOODS WITH RESPECT TO A SET OF DESIRABILITY ATTRIBUTES IN A SURVEY ON WOMEN'S DIETARY HABITS IN RURAL UGANDA.

Luigi Palla[1], Andrea Spray[1], AnnaRita Vestri[2], Gwen Varley[3], Joweria Nambooze[4], Kate Wellard[3], Elaine Ferguson[1]

[1] *London School of Hygiene & Tropical Medicine, London, UK*

[2] *La Sapienza University, Rome, Italy*

[3] *Natural Resources Institute, University of Greenwich, UK*

[4] *Africa Innovations Institute, Bukoto, Uganda*

Introduction

Research, in high income countries, shows that high maternal workloads have an important influence on food choice [1]. The impact of women's workload on food choice among rural women living in LMICs, however, is largely unknown. Investigating such food choices motivated a research project funded by Drivers of Food Choice grant programme that surveyed a sample of women from rural villages in Uganda for dietary and time-use habits. The questionnaire also gave the opportunity to collect information on women's own view regarding some foods commonly consumed in the local setting. In particular, the study participants in the Drivers of Food Choice study have been presented with a survey questionnaire asking them to rate 12 foods selected from typical diet in rural Uganda: yellow sweet potato, beans, jackfruit, doodo (amaranthus), maize, tomato, milk, mango, mukene (silver cyprinid also known as Lake Victoria sardines), groundnuts, egg and matooke (East African highland banana), with respect to 6 attributes: cheapness, accessibility, deliciousness, healthiness, quickness to prepare, quickness to cook.

Objective

The aim of this study is to characterise the food ratings in this survey of Ugandan women in rural settings, at population level. In particular, specific objectives are: to find similarities in food rating profiles across attributes and in attributes profiles across foods, as well as finding association between foods and attributes.

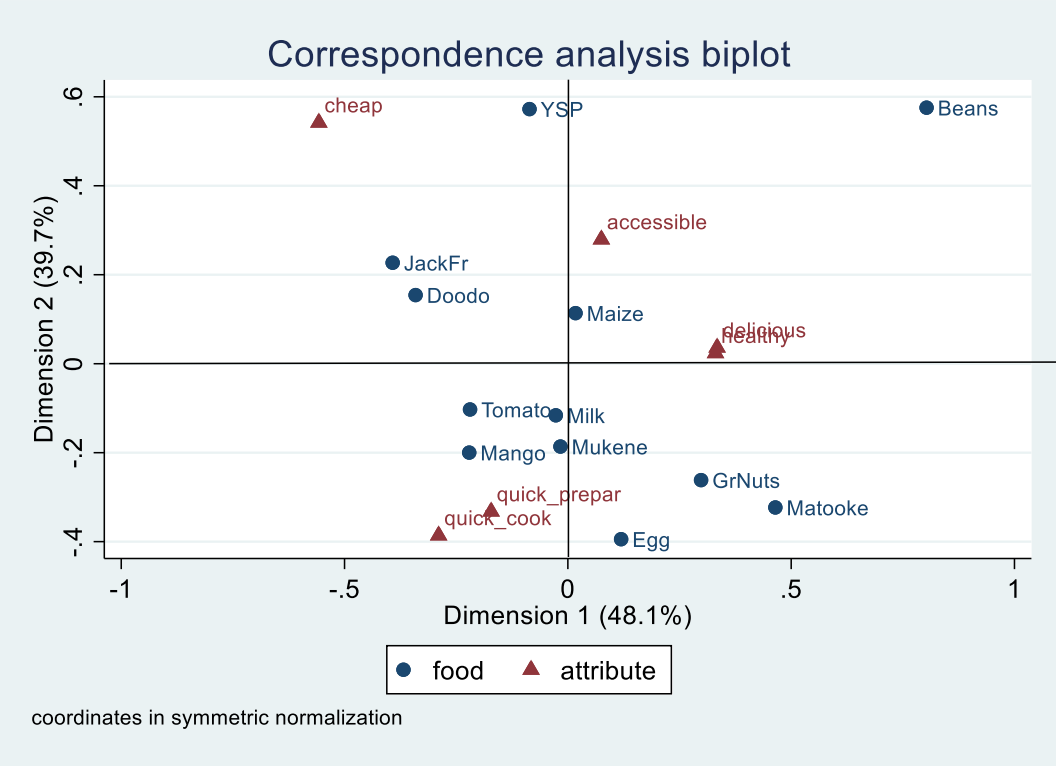
Methods

The data were collected in the context of a Drivers of Food Choice project whereby a cross-sectional survey of mothers and children age 12 to 23 months (n=226) was conducted between January and February 2018 from 18 villages in the Bugiri and Kamuli Districts, Eastern Region, Uganda. Within a general questionnaire survey administered by trained interviewers, 206 women rated twelve foods from 1 to 4 (corresponding to definitely (4), moderately (3), moderately not (2), definitely not (1)) for each of 6 attributes: cheapness, accessibility, deliciousness, healthiness, quickness to prepare, quickness to cook. Average rating for each combination of foods and attribute across women were then obtained and the resulting 12*6 matrix (i.e. 12 foods by 6 attributes) was the basis for the following analysis.
Correspondence analysis [2], [3], [4] was applied to such matrix first. This technique is often applied to analyse contingency tables, i.e. matrix of counts, but may also be utilised when the character in the cells is a different type of measure, like the ratings of an attribute. This allows us to study the relationship between rows and columns (foods and attributes) by a visual technique, the biplot, that displays in the same graph a projection of categories of both variables (foods and attributes) on the axes of principal inertia (deviation

from the condition of independence of the two variables). The origin of the plot corresponds to the average profile for both rows (foods) and columns (attributes). Similarity between food profiles or attribute profiles is reflected in their respective proximity in the graph. Results are displayed in a symmetric coordinates biplot whereby foods and attributes that are furthest from the origin and mostly aligned on the same direction away from the origin can be regarded as most positively related.

Results

Figure 1. *Correspondence Analysis Biplot, accounting for nearly 87.8% of total system inertia.*



From Correspondence Analysis biplot (Figure 1), representing 87.8% of the system inertia, the pair of attributes that have very similar ratings across foods appear to be: quickness to prepare and quickness to cook (lower left quadrant; deliciousness and healthiness (upper right quadrant) which almost completely overlap. Furthermore, the pair of foods that appear to have most similar ratings across attributes are: 1) jackfruit and doodo (upper left); 2) tomato and mango (lower left); 3) groundnuts and matooke (lower right); 4) milk and mukene (lower left). As regards the association between categories of different variables, jackfruit and doodo appear to be associated with cheapness and mango (and to some extent also milk, mukene and tomato) with quickness to prepare/cook.

Discussion

A multivariate analysis of healthy foods rated on a scale 1-4 for desirable attributes by women in rural settings in Uganda (as part of a Drivers of Food Choice study) highlighted, at population level, very little variability in the delicious and healthy quality of the foods, which indicates that women in those settings tend to recognise the healthiness of the foods and associate healthiness with a pleasant taste or alternatively

attribute healthiness to foods that they find delicious: the causal direction of this association remains to be determined. The variation in the average ratings appear mostly linked to the ease in cultivating locally or retrieving at local markets such foods and in the economic capacity to buy them in settings with very limited financial resources. In particular it appears that the vegetables doodo and jackfruit are the products with the best ratings across all attributes, being especially cheap (unlike tomato which shares with them similar ratings). It should be noted that doodo, jackfruit and mango are available as foods that grow naturally in the wild and so might have no economic costs at all for households in rural settings. Animal products like milk and mukene have similar ratings across all "practical" attributes but they score much lower for cheapness. Even worse for cheapness are eggs which, like groundnuts, however, have a somewhat lower accessibility in women's view. Finally some non–cheap plant-based products like beans, which are rich in both proteins and micronutrients, are also deemed long to prepare/cook.

Conclusion

The access to nutrient-dense foods that can be cultivated in the household or easily accessed and prepared remains the cheapest channel towards healthy food consumption. However to promote a balanced and diverse diet, policies should be put in place to make nutrient-dense products, like eggs, mukene and beans cheaper, more accessible and/or easier to prepare in rural households in Uganda. In particular Social and Behaviour Change Communication programmes should be put in place to raise awareness of nutritious foods and promote consumption for most vulnerable populations (like pregnant and lactating women and infants), as well as programmes for the identification and distribution of other nutrition-dense foods not currently available in this food environment, such as orange-fleshed sweet potato, iron-rich beans.
Finally, in order to better elucidate the determinants of food choice at the individual, rather than population level, we next plan to study the relationship between the ratings expressed and the actual reported weekly consumption for the twelve rated foods, as well as the use of time to prepare such foods.

References

[1] Devine C.M., Farrell T.J., Blake C.E. et al, Work conditions and the food choice coping strategies of employed parents, J Nutr Educ Behav, 2009, 41-5.

[2] Greenacre M. Correspondence analysis in practice, 2016, Chapman and Hall/CRC.

[3] Beh E., Lombardo R. Correspondence Analysis: Theory, Practice and New Strategies, 2014, Wiley Series in Probability and Statistics

[4] Bower J.A., Statistical methods for Food Science, 2013, John Wiley and Sons.

ANALISI BAYESIANA DEL RISCHIO DI RICOVERO PER MALFORMAZIONI CONGENITE IN PUGLIA. ANNI 2001-2013.

Bartolomeo Nicola[1], Mincuzzi Antonella[2], Leogrande Simona[2], Serio Gabriella[1], Minerba Sante[2], Trerotoli Paolo[1]

[1] *Dipartimento di Scienze Biomediche ed Oncologia Umana, Università degli Studi di Bari Aldo Moro*

[2] *SC Statistica Epidemiologia, ASL Taranto*

Introduzione

Le malformazioni congenite (MC) sono errori della morfogenesi, determinati solo in parte da fattori genetici (25% dei casi). L'esposizione della madre e del feto a fattori teratogeni noti (infettivi, fisici, chimici, patologie materne) causa circa il 9-10% dei difetti, mentre il 65% ha un'eziologia non nota, forse correlata a complesse interazioni tra i geni e l'ambiente.
Il Ministero della salute stima che, in Italia, su 500.000 nati mediamente ogni anno, circa 25.000 presentino almeno una MC. Sempre in Italia, circa il 25% della mortalità infantile è dovuta ad MC e circa il 50% a stati morbosi perinatali, quasi sempre di origine prenatale.
Negli ultimi quindici anni, vi è stata una evidenza crescente circa l'associazione tra l'esposizione all'inquinamento atmosferico e un aumentato rischio di malformazioni congenite. Uno studio di coorte retrospettivo condotto tra il 1997 e 2004 su una popolazione di 216730 bambini israeliani ha dimostrato che l'esposizione ad alti livelli di PM10 e NOx durante la gravidanza era associato ad un aumentato rischio di malformazioni congenite [1]. Uno studio caso-controllo effettuato nell'area di Brindisi (Italia) su bambini nati nel periodo 2001-2010 ha evidenziato che l'esposizione materna al biossido di zolfo aumenta il rischio di cardiopatia congenita [2]. Comunque, limitazioni metodologiche negli studi epidemiologici non hanno consentito di evidenziare in maniera definitiva la relazione tra inquinamento e incidenza o rischio di malformazioni congenite. Nel 2018, Baldacci et al. [3] hanno prodotto una revisione sistematica delle evidenze epidemiologiche sul rischio di MC associato alle esposizioni ambientali e ai fattori di rischio socioeconomici e individuali, dalla quale emerge che l'evidenza epidemiologica relativa all'associazione tra esito di MC ed esposizione a siti industriali è da considerarsi limitata; così come è limitata l'evidenza relativa all'associazione tra esposizione a inquinanti atmosferici e rischio di MC. Gli autori auspicano migliori metodi di valutazione dell'esposizione, in particolare misure o modelli spaziali più accurati, una definizione standardizzata dei casi, un campione più ampio e un controllo più accurato dei fattori confondenti principali o presunti.
La regione Puglia presenta alcune aree a rischio ambientale, già oggetto di studi specifici sulla distribuzione di patologie neoplastiche. Risulta di interesse la valutazione, in queste aree fortemente industrializzate, della presenza e distribuzione delle malformazioni congenite.

Obiettivi

L'obiettivo di questo studio è di valutare, attraverso l'applicazione di modelli spaziali bayesiani applicati a dati provenienti da database sanitari amministrativi, la distribuzione dei ricoveri per anomalie congenite in Puglia e nella città di Taranto in particolare, la quale risulta ad alto rischio ambientale rispetto al contesto regionale.

Metodi

L'analisi si basa sui dati di ospedalizzazione rilevati dall'archivio delle schede di dimissione ospedaliera (SDO) della Puglia negli anni 2001-2013, codificati secondo la Classificazione internazionale delle malattie

- nona revisione con modificazioni cliniche (ICD IX-CM). Sono stati considerati i ricoveri avvenuti entro e fuori regione dei residenti nell'area in studio. Le malformazioni congenite sono state individuate, tra i nati vivi, dal codice di diagnosi principale e/o secondaria compreso tra 740 e 759. Oltre all'analisi globale di primo livello dei Comuni pugliesi, è stata effettuata un'analisi di dettaglio sul Comune di Taranto utilizzando come aree di secondo livello le sezioni di Censimento Istat, poi aggregate in quartieri. L'attribuzione dei casi alle sezioni di censimento per l'analisi di secondo livello è avvenuta attraverso una procedura di record linkage (RL) tra le banche dati sanitarie e il db dell'Anagrafe Comunale di Taranto. Il RL è avvenuto a step successivi, iniziando con quello deterministico seguito da RL probabilistici sui residui non linkati.
Sono stati, quindi, determinati i casi attesi per comune e per sezione di censimento utilizzando un metodo di standardizzazione indiretta, usando come riferimento i tassi regionali della regione Puglia.
Casi osservati e casi attesi sono stati modellati per la stima dei rischi relativi, utilizzando l'indice di deprivazione socio-economica di ciascuna area geografica per controllare il possibile confondimento dovuto ai fattori socioeconomici e assumendo la presenza di una struttura spaziale dei rischi (*smoothing*).
In un'analisi metodologica esplorativa è stata valutata l'applicazione di due diversi modelli bayesiani di Disease Mapping: il modello di Besag, York e Mollié (BYM) [3,4] e il modello Zero Inflated Poisson (ZIP) [5]. È stato scelto il primo (BYM) dopo aver valutato sia la convergenza delle distribuzioni a posteriori dei parametri attraverso la statistica di Gelman-Rubin, sia il *fitting* del modello attraverso alcune misure quali il DIC (Deviance Information Criterion), lo pseudo-R^2 e la devianza assoluta dei residui rispetto ai valori di rischio stimati dal modello.
Nel modello BYM, il numero di casi osservati (O_i) nell'area geografica (comune di residenza oppure sezione di censimento) è assunto seguire una distribuzione di Poisson con media $\mu_i = E_i\, RR_i$:

$$O_i \sim \text{Poisson}(\mu_i)$$

$$\log RR_i = \gamma_0 + \gamma_1\, \text{depriv}_i + b_i + h_i$$

dove E_i rappresenta il numero di casi attesi nella i-esima unità geografica, RR_i il "vero" ma non conosciuto rischio relativo nell'area i da stimare, depriv_i è l'indice di deprivazione nell'area i-esima, b_i è la componente strutturata degli effetti spaziali random attribuita a priori secondo il modello CAR (Conditional Auto Regressive), h_i è la componente non strutturata degli effetti spaziali random per la quale si assume una distribuzione a priori Normale $(0;\tau_h)$.
L'indice di deprivazione, utilizzato come covariata nel modello, è stato costruito sulla base di 5 indicatori che rappresentano: basso livello di istruzione, disoccupazione, mancato possesso dell'abitazione, famiglia monogenitoriale e densità abitativa.
Per lo *smoothing* dei rischi la componente strutturata del modello di BYM è rappresentata da una matrice di adiacenza che considera, nell'analisi di primo livello, ciascun comune influenzato dai comuni adiacenti, mentre nell'analisi di secondo livello ciascuna sezione di censimento influenzata dalle sezioni che ricadono in un raggio di 500 metri.
L'analisi è stata effettuata utilizzando il software SAS 9.4 for PC per la gestione dei databases, il RL e la standardizzazione, Winbugs 14 per la stima dei rischi con il modello Bayesiano ed il software QGIS 2.8.2 per la mappatura dei rischi e la gestione della georeferenziazione.

Risultati

Nel periodo 2001-2013, in Puglia, sono stati osservati 13073 casi incidenti di MC, 7884 maschi (60.3%) e 5189 femmine (39.7%), con un tasso medio nel periodo di 26.5 bambini con MC ogni 1000 nati vivi. Il tasso annuale, variabile nel range 22.0-26.4 negli anni tra il 2001 e il 2009, è sensibilmente in cresciuto negli ultimi anni esaminati: 27.9 nel 2010, 28.2 nel 2011, 30.6 nel 2012 fino ad arrivare a 34.9 bambini con MC ogni

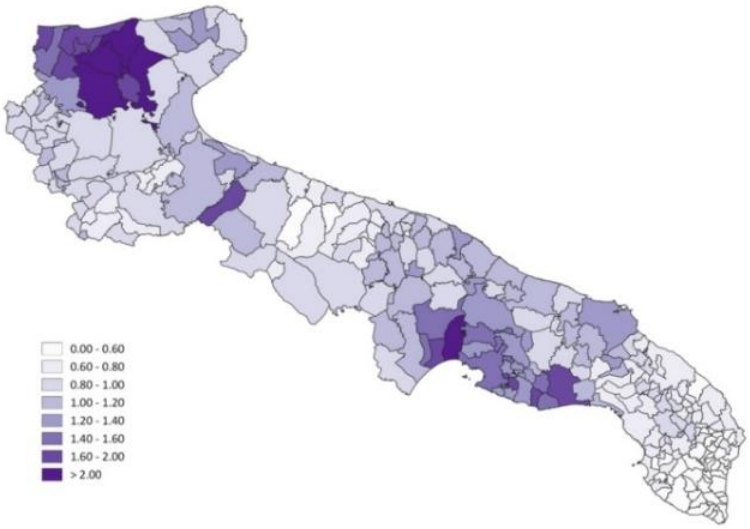
1000 nati vivi nel 2013. L'analisi della mappa regionale mette in evidenza la presenza di due zone nelle quali è risultato più elevato il rischio relativo di ricovero per MC: l'area più a nord del "Tavoliere di Foggia" e l'arco ionico-tarantino con il nord-salento. Nella provincia di Foggia i comuni che rientrano nel cluster evidenziato hanno dei rischi superiori a 1.60 con diversi casi di RR superiore a 2.00. Nella zona ionico-salentina un solo comune ha un RR superiore a 2.00, mentre tutti gli altri comuni hanno un RR che varia nel range 1.20-2.00. Sebbene meno marcatamente, appare più a rischio rispetto al riferimento medio regionale anche la zona costiera che si estende dal sud di Bari fino a Brindisi, dove i rischi relativi variano tra 1.00 ed 1.40.
Nel comune di Taranto sono stati osservati 971 casi di MC tra il 2001 ed il 2013, con un RR di ricovero pari a 1.60 [I.Cr. 95%, 1.50-1.70]. Il tasso medio di periodo è stato di 43.0 bambini con MC ogni 1000 nati vivi; tasso comunale anch'esso in crescita negli ultimi anni essendo passato dal 38.1x1000 nel periodo 2001-2010, al 53.1 del 2011, 61.6 del 2012 e 62.9 bambini con MC ogni 1000 nati vivi nel 2013. Scendendo nel dettaglio all'analisi per quartiere, i RR sono tutti significativi e superiori a 1.30, con i quartieri a ridosso dell'area industriale che rientrano nella classe 1.60-1.80.

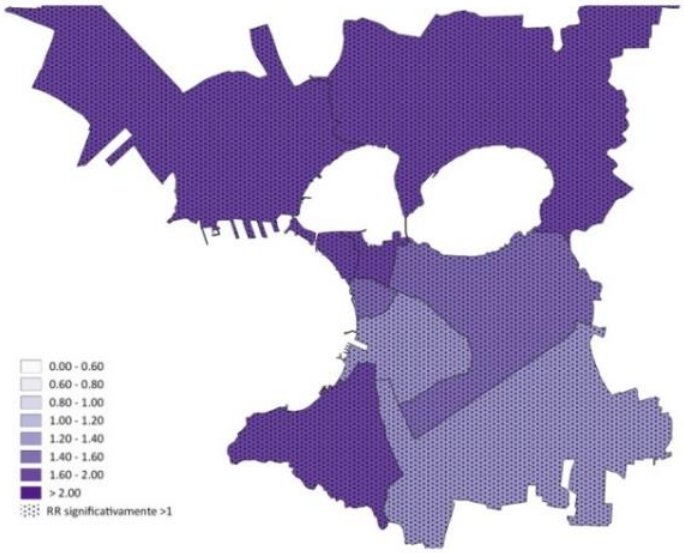
Conclusioni

Ci sono alcune aree in Puglia che mostrano una evidenza di rischio elevato di ricovero per malformazioni congenite. Lo smoothing bayesiano del rischio relativo ha permesso di delinearne i contorni quasi ad elevarli al rango di "aggregati" o "cluster". Si tratta della zona a nord della provincia di Foggia e quella, meno marcata ma più estesa, che ha come centromero il comune di Taranto. La prima area non rientra tra quelle che l'Istituto Superiore di Sanità, attraverso il progetto "Sentieri", ha individuato come Siti di Interesse Nazionale (SIN) perché esposti a rischio da inquinamento; mentre l'area ionica, con i comuni di Taranto e Statte, ne fa parte e da anni è al centro dell'attenzione per gli effetti sulla salute di un insediamento industriale.
Anche nel V rapporto "Sentieri", pubblicato nel 2019 [6], il rapporto standardizzato calcolato per le MC a Taranto è significativamente superiore al riferimento regionale, nel periodo 2002-2015.
La mappa dettagliata per quartieri sembra avvalorare la tesi di una maggior rischio di MC legato all'esposizione diretta alle fonti di inquinamento.
Anche il Comune di Brindisi rientra nell'elenco dei SIN ed anche in tale area la nostra analisi ha evidenziato un rischio aumentato di ricovero per MC. Già nello studio del 2012 effettuato da Gianicolo et al. [7] emergeva come nella città di Brindisi, nel periodo 2001-2010, ci fosse un incremento della prevalenza di anomalie congenite (specialmente le malattie cardiache congenite), 1.2 volte superiore a quello riportato dal pool di registri EUROCAT. In controtendenza i risultati dell'ultimo studio "Sentieri" secondo il quale la prevalenza totale delle MC risulta sostanzialmente in linea con quella regionale. Il confronto tra i risultati in controtendenza è limitato dai due differenti periodi di osservazione.
Negli altri due SIN pugliesi, Manfredonia (FG) e Bari, i risultati tra il presente studio e lo studio "Sentieri" indicano coerentemente un leggero eccesso di rischio nel comune foggiano ed una prevalenza più bassa, rispetto alla regione, nel comune di Bari. La sistematicità delle stime in difetto nel capoluogo di regione sembra poco plausibile e suggerisce una possibile anomalia nei dati delle schede di dimissione ospedaliera utilizzati per il calcolo degli indicatori, laddove vi è probabilmente una bassa qualità di codifica delle MC.
L'aumento di casi di MC registrato in tutta la regione negli ultimi anni esaminati potrebbe essere stato favorito da una maggiore attenzione nella compilazione della SDO.
Sicuramente, rispetto al passato, ci sono maggiori evidenze circa la relazione tra inquinamento ambientale e rischio aumentato di malformazioni congenite. In particolare sono state trovate evidenze riguardanti l'effetto dell'inquinamento atmosferico sul rischio di anomalia cardiaca congenita [8]. Si è visto, ad esempio, che l'aumento dell'esposizione a PM2,5 nel periodo di periconcezionale è associato ad alcuni modesti aumenti del rischio di malformazioni congenite. Però, sebbene l'aumento del rischio sia modesto, il

potenziale impatto sulla popolazione diventa rilevante quando un numero maggiore di donne ha un certo grado di esposizione [9].
L'aver introdotto, come fattore di aggiustamento del rischio, l'indice di deprivazione ha sicuramente evitato il bias del confondimento dovuto alla situazione socioeconomica della famiglia in cui si registra la nascita di un bambino con MC.
Permangono due criticità, entrambe legate all'impossibilità di collegare la scheda di dimissione del neonato a quella della madre: infatti non è stato possibile aggiustare il rischio di MC per l'età della madre e abbiamo dovuto assumere, soprattutto nell'analisi di geografica di secondo livello, che la residenza della madre durante la gestazione corrispondesse a quella al momento della registrazione anagrafica del nascituro. Sicuramente, quindi, tra gli sviluppi futuri vi è quello di utilizzare banche dati complementari alle SDO, quali i Certificati di Assistenza al parto e, laddove esistenti, i Registri delle Malformazioni.

Figura 1. *Distribuzione del RR di ricovero per AC in Puglia e nel Comune di Taranto, anni 2001-2013.*





Bibliografia

[1] Farhi A., Boyko V., Almagor J. et al., The possible association between exposure to air pollution and the risk for congenital malformations. Environ Res., 2014;135:173–180.

[2] Gianicolo E.A., Mangia C., Cervino M. et al., Congenital anomalies among live births in a high environmental risk area--a case-control study in Brindisi (southern Italy). Environ Res., 2014;128:9–14.

[3] Clayton D., Kaldor J., Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. Biometrics, 1987;43:671–91.

[4] Besag J., York J., Mollie A., Bayesian image restoration with two applications in spatial statistics. Ann Inst Stat Math., 1991;43:1–21.

[5] Lambert D., Zero-Inflated Poisson Regression, with An Application to Defects in Manufacturing. Technometrics, 1992;34:1-14

[6] Zona A., Iavarone I., Buzzoni C. et al., SENTIERI - Studio epidemiologico nazionale dei territori e degli insediamenti esposti a rischio da inquinamento. Quinto Rapporto. Epidemiol Prev. 2019; 43 (2-3), Suppl.1:1-208

[7] Gianicolo E.A., Bruni A., Rosati E., et al., Congenital anomalies among live births in a polluted area. A ten-year retrospective study. BMC Pregnancy Childbirth., 2012 Dec 27;12:165.

[8] Vrijheid M., Martinez D., Manzanares S. et al., Ambient Air Pollution and Risk of Congenital Anomalies: A Systematic Review and Meta-analysis. Environ Health Perspect., 2011 May;119(5):598–606.

[9] Ren S., Haynes E., Hall E. et al., Periconception Exposure to Air Pollution and Risk of Congenital Malformations. J Pediatr., 2018 Feb;193:76-84.

VALUTAZIONE DELL'INQUINAMENTO QUALE FATTORE DI RISCHIO NELLA CANCEROGENESI DEL SISTEMA EMOLINFOPOIETICO: ESPERIENZA IN TERRITORI MOLISANI E ABRUZZESI A RISCHIO – RISULTATI PRELIMINARI

De Gregorio Daniela[1], Recchia Laura[1], Nofroni Italo[2], Alunni Fegatelli Danilo[2], de Michele Isabella[3], Vitullo Felice[4], Cantò Chiara[4], Storti Sergio[5], Vestri Anna Rita[2]

[1] Dip. di Medicina e Scienze della salute, Università degli Studi del Molise, Campobasso

[2] Dip. Sanità pubblica e Malattie infettive, Università Sapienza, Roma

[3] Psicologa, Centro trasfusionale Ematologia, Ospedale San Timoteo, Termoli, Campobasso

[4] Medico epidemiologo, Ospedale Spirito Santo, Pescara

[5] Università Cattolica, Roma. Direttore U.O.C di Oncoematologia, Fondazione di Ricerca e Cura "Giovanni Paolo II", Campobasso

Introduzione

L'area del basso Molise si caratterizza da parecchi decenni per una elevata prevalenza di patologie leucemiche. L'inquinamento ambientale dovuto alla presenza di un nucleo industriale nella città di Termoli, lo sversamento di reflui in canali e corsi d'acqua, la presenza di discariche abusive di scarti tossici non adeguatamente smaltiti, l'interramento di rifiuti pericolosi nei terreni agricoli ed il controllo, probabilmente insufficiente, delle emissioni e dei fumi delle industrie, non hanno favorito la salvaguardia ambientale del territorio né la tutela della salute pubblica. Per le stesse motivazioni, anche alcuni territori abruzzesi situati nelle province di Pescara e di Chieti e caratterizzati dai suddetti rischi (Comba et al., 2006, 2009) sono stati presi in considerazione.
Le patologie oncoematologiche nel loro complesso arrivano a rappresentare in Italia circa il 10% di tutti i tumori, facendo registrare 40.000 nuove diagnosi ogni anno. Le loro cause non sono ancora del tutto chiare, anche se ormai è evidente che tale patologie, come per altri tumori, costituiscono il risultato finale di una complessa serie di eventi che coinvolgono sia fattori genetici costituzionali, sia fattori epigenetici ambientali e funzionali come radiazioni ionizzanti, sostanze tossiche, derivati del benzene, agenti infettivi.

Obiettivo

Lo studio in corso si propone di verificare se e in che modo l'esposizione ambientale ad agenti inquinanti di vario tipo possa costituire un fattore di rischio per l'insorgenza di patologie del sistema emolinfopoietico (leucemie e linfomi) per la popolazione residente nel territorio del Basso Molise e in quello Abruzzese delle province di Chieti e di Pescara.

Materiali e metodi

Il presente progetto si configura come studio caso-controllo finalizzato a valutare la maggior pregressa esposizione dei soggetti affetti da neoplasie del sistema emolinfopoietico, ricoverati nei reparti di Ematologia dei centri ospedalieri di Termoli (Ospedale San Timoteo), di Campobasso (Fondazione di Ricerca e Cura "Giovanni Paolo II" -Università Cattolica) e di Pescara (Ospedale Spirito Santo). Nella popolazione residente verranno identificati, quindi, due gruppi di soggetti: i casi, pazienti con la patologia linfoematopoietica in atto, e i controlli, soggetti selezionati fra coloro che non sono affetti dalle patologie studiate, né da altre patologie "inquinamento correlate".

La valutazione dell'intensità dell'esposizione dei soggetti residenti nell'area di studio, sarà effettuata mediante la variabile "distanza in km" tra i luoghi abitualmente frequentati da ciascun individuo (residenza, posto di lavoro, di studio e di svago) e i siti considerati principali fonti di inquinamento. Al tal fine, attraverso una procedura di georeferenziazione effettuata utilizzando il Software ArcGIS 10.3, sarà possibile caratterizzare l'area in studio e costruire una mappatura del territorio considerato sulla base di diversi livelli di inquinamento.
Assumendo la probabilità di esposizione tra i controlli pari a 0.2 e un coefficiente di correlazione per l'esposizione tra i casi e i controlli appaiati pari a 0.2, prefissando un OR = 2.2, un alfa pari a 0.05 e una potenza del test di 0.85, per individuare una differenza statisticamente significativa, occorreranno 135 casi e 270 controlli per un numero totale di 405 soggetti da arruolare. Prevedendo un 10% di questionari inaffidabili o solo parzialmente compilati, si prevede di arruolare altri 45 soggetti (15 casi e 30 controlli), per un totale di 450.
Sui dati raccolti sarà eseguita l'analisi statistica descrittiva, tramite il calcolo di indici di tendenza centrale con i relativi intervalli di confidenza al 95%. Per valutare la presenza di associazione tra variabili di tipo categorico sarà eseguito il test statistico chi quadro con analisi di sottogruppi in base a sesso, età, tipo di tumore, attività lavorativa, livello di esposizione. Per le variabili continue, in funzione del tipo di distribuzione dei dati, saranno eseguiti test parametrici o non parametrici. Il livello di significatività verrà fissato a 0.05.
La forza dell'associazione tra le variabili indipendenti e l'insorgenza della patologia sarà misurata attraverso il calcolo degli odds ratio con i relativi intervalli di confidenza al 95% (Wilson method). Tutte le variabili significative all'analisi univariata saranno inserite in un modello di regressione logistica multivariata con procedura ENTER e/o STEPWISE di selezione delle variabili, al fine di individuare le covariate indipendenti come eventuali predittori dell'insorgenza della patologia, eventualmente con "aggiustamento" delle variabili confondenti.
L'analisi dei dati georeferenziati all'interno della regione geografica di indagine, consentirà sia l'individuazione di eventuali cluster, che il calcolo dello stimatore Kernel (Goovaerts, 2005) per il rischio di ammalarsi di leucemie nei comuni oggetto dello studio.

Risultati preliminari

Al momento sono stati raccolti 53 casi e 53 controlli. Le elaborazioni preliminari non consentono, per l'esiguità dell'attuale campione, analisi esaustive ed approfondite. Ma già si può trarre qualche importante conclusione: in un modello di regressione logistica, metodo enter, la variabile età appare fin da subito avere un ruolo significativamente predominante per l'insorgenza delle patologie del sistema ematopoietico con un OR di 1.074 (95% CI: 1.041 – 1.109). Non entrano nel modello le variabili Sesso, Zona di residenza, espressa dicotomicamente in area ad inquinamento basso o alto, e anzianità nell'attuale residenza. Di quest'ultimo rilievo appare, al momento, difficile dare una interpretazione sensata; solo il procedere dello studio ci consentirà, probabilmente, di comprendere meglio il meccanismo di insorgenza di queste patologie.

Bibliografia

Comba P.Trinca S, Viviano G., et al. Realizzazione di un'indagine epidemiologica su ambiente e salute nell'area di Termoli. Rapporto finale (2009). ISS, Roma.

Comba P., Conti S., Iavarone I., et al. Ambiente e salute a Taranto: evidenze disponibili e indicazioni di sanità pubblica. Considerazioni conclusive (2012). http://www.salute.gov.it.

Goovaerts P. Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging. International Journal of Health Geographics (2005), 4:31.

TOWARDS A HYBRID HUMAN-MACHINE EMERGENCY TRIAGE: PREDICTING CLINICAL EVALUATIONS FROM EMERGENCY DEPARTMENT TRIAGE INFORMATION THROUGH MACHINE LEARNING MODELS IN ORDER TO OPTIMIZE ACCESS TO CARE PRIORITY LISTS

Milanese Alberto[1], Vestri Annarita[1]

[1] *Department of Public Health and Infectious Diseases, Sapienza University of Rome, Rome, Italy.*

Background

In the Italian emergency departments (ED), the assignment of the priority code (currently, a 4-levels colour code of progressively decreasing urgency for medical care: “red”, “yellow”, ”green”,” white”; in the near future to be replaced by a 5-levels numeric code of progressively decreasing urgency from 1 to 5) is the outcome of the nursing decision formulated within the scope of triage's activity and is based on the elements detected in the subjective and objective assessment phases. This process determines the priority of access to care in relation to the patient's clinical conditions, the risk of their progression in time and the availability of resources. The priority code assigned by the triage personnel not necessarily corresponds to the clinical severity of the patient: only the integration with the subsequent phase of evaluation, carried out during the medical examination, will produce the essential elements for the correct formulation of the clinical gravity judgment of the case.
At the end of the medical service path, the ED doctor must identify the patient's severity level by assigning an exit code, which is currently represented by colour codes similar to those assigned during triage (4-levels colour code: “red”, “yellow”, ”green”, ”white”; in the near future to be replaced by a 4-levels letter code: "C"-Critical Patient-, "A" -Acute Patient-, "D" -Urgent Differable Patient-, "N"-Non Urgent Patient-) and has the purpose of evaluating the overall clinical severity of the patient, resulting an indicator of appropriateness of the access to emergency department medical care [1].
Whereas the priority code could not be used to predict the overall gravity of the patient, it might be reasonable to use the exit code information together with the destination of the patient subsequent to the emergency department assistance as a mean of assessing priority of access to medical care in ED.
Specifically, in the case of emergency department triage, it is known that overcrowding and boarding in the emergency department have been identified as a healthcare problem in many countries. This is especially significant as triage main aim is to identify vulnerable and high-risk patients through a rapid assessment among a large number of less urgent patients, therefore ED overcrowding, boarding and delay can determine increased mortality in many situations [2-4].
Regarding the triage priority score, although standardized, it heavily relies on provider judgment and experience, hence is subject to high variation. Moreover, contrary to the objective of triage itself, the system struggles in differentiating especially the "green" codes (situation analogous to the international level 3 of ESI -Emergency Severity Index-), a large diversely ill patient group which accounts for the greatest proportion of all the attributed codes [5]. Such difficulty to differentiate may pose safety risks to patients, who could be critically ill and undertriaged, as well as influence the efficiency of ED resources allocation, as low-acuity patients are often overtriaged for several reasons [6,7].
It has been proved that machine learning models, compared to the conventional approaches, demonstrate a superior performance to predict critical care and hospitalization outcomes, as well as it has been speculated that in the future the extensive use of such systems might enhance clinicians’ triage decision making, thereby achieving better clinical care and optimal resource utilization [8].
The patient safety associated with triage in crowded ED limitations could therefore find a solution for the need for accurate risk assessment thanks to a hybrid triage system that exploits machine-learning capabilities.

Objectives

The study aims to build and to evaluate the performance of several machine-learning-based triage models in predicting clinical gravity of adult patients (≥18 years) accessing to the emergency department from routinely collected triage data (e.g. primary chief complaint, mode of arrival, basic demographics, vital signs, comorbidities).
Clinical gravity will be assessed and evaluated according to two different outputs: the exit code classification of the patient overall gravity (4-levels colour code, currently analogous to triage priority code) and a 3-levels emergency clinical outcome built according to the patient destination subsequent to the emergency department medical care ("High" for inhospital mortality, direct admission to ICU, emergency procedure, "Medium" for direct hospital admission or transfer to other structure, "Low" for home discharge).
The purpose of this project is twofold: on the one hand, we want to validate for the Italian reality a state of the art procedure of machine-learning-based triage to be compared to existing human-based triage as well as assess among the several machine learning models to be developed which of the aforementioned performs the best on test data, on the other hand, we want the project to be a potential laying foundation for the future development of an AI-assisted triage system for the emergency department, aimed to real-time assist the nursing personnel in everyday triage work, potentially being able to reduce incident and near miss rates, suggest patient reevaluations and optimize access to care priority lists.

Methods

The study will be structured as a retrospective study based on the last 5 years emergency department accesses of adult patients (≥18 years); the data will be provided by Policlinico Umberto I triage records administrative database.
In order to analyze these data, we will exploit the capabilities of machine-learning algorithms, which allow computers to learn by themselves from data without explicit programming: in our research we will use several models in order to realize our predictions and we will then compare their performances.
Some of the models we will use are the following:
Logistic regression, a model that is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, given a set of independent variables [9].
Random forest classification, an ensemble learning method that operates by constructing a large number of decision trees and outputting the class that is the mode of the classes of the individual trees [5].
Deep artificial neural networks, supervised learning computational models inspired by the biological neural networks that constitute animal cortex and able to learn tasks by considering examples without task-specific programming. Neural networks are nonlinear statistical data modeling tools which are able to find complex patterns between inputs and outputs. Artificial neurons are aggregated into several layers, so that different layers may perform different kinds of transformations on their inputs, with signals travelling from the first layer (the input layer) to the last layer (the output layer) multiple times in both forward and backward passes: the connections between artificial neurons (a.k.a. nodes) have a weight that adjusts as learning proceeds [10].
Machine learning models will use routinely available triage data as predictors (e.g. primary chief complaint, mode of arrival, basic demographics, vital signs, comorbidities) and will be trained to predict, via supervised learning, two different outputs: on the one hand, the medical exit code classification of the patient overall gravity (4-levels colour code, analogous to triage priority code), on the other hand, a 3-levels emergency clinical outcome built according to the patient destination after the emergency department care ("High" for inhospital mortality, direct admission to ICU, emergency procedure, "Medium" for direct hospital admission or transfer to other structure, "Low" for home discharge).

In order to validate the procedure and assess which model performs the best on test data, several machine learning models will be implemented; the collected data will be randomly partitioned into training, validation and test sets, then we will finally measure the predictive performance of the models [11].

References

[1] Ministero della Salute, Schema di Accordo tra il Ministro della Salute, le Regioni e le Province Autonome di Trento e di Bolzano sui Documenti "Linee di indirizzo nazionali sul Triage Intraospedaliero", "Linee di indirizzo nazionali sull'Osservazione Breve Intensiva" e "Linee di indirizzo nazionali per lo sviluppo del Piano di gestione del sovraffollamento in Pronto Soccorso". 2019 May.

[2] Bernstein S., Aronsky D., Duseja R. et al., The effect of emergency department crowding on clinically oriented outcomes. Acad Emerg Med. 2009 Jan;16(1):1-10.

[3] Sun B., Hsia R., Weiss R. et al., Effect of emergency department crowding on outcomes of admitted patients. Ann Emerg Med. 2013 Jun;61(6):605-611.e6.

[4] Carter E., Pouch S., Larson E. The relationship between emergency department crowding and patient outcomes: a systematic review. J Nurs Scholarsh. 2014 Mar;46(2):106-15.

[5] Levin S., Toerper M., Hamrock E. et al., Machine-Learning-Based Electronic Triage More Accurately Differentiates Patients With Respect to Clinical Outcomes Compared With the Emergency Severity Index. Ann Emerg Med. 2018 May;71(5):565-574.e2.

[6] Goto T., Camargo C., Faridi M., Machine Learning–Based Prediction of Clinical Outcomes for Children During Emergency Department Triage. JAMA Netw Open. 2019 Jan 4;2(1):e186937.

[7] Newgard C., Staudenmayer K., Hsia R. et al., The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers. Health Aff (Millwood). 2013 Sep;32(9):1591-9.

[8] Raita Y., Goto T., Faridi M. et al., Emergency department triage prediction of clinical outcomes using machine learning models. Crit Care. 2019 Feb 22;23(1):64.

[9] Dugas A., Kirsch T., Toerper M. et al.,An Electronic Emergency Triage System to Improve Patient Distribution by Critical Outcomes. J Emerg Med. 2016 Jun;50(6):910-8.

[10] Joon-myoung Kwon, Youngnam Lee, Yeha Lee et al., Validation of deep-learning-based triage and acuity score using a large national dataset. PLoS One. 2018; 13(10): e0205836.

[11] Woo Suk Hong, Adrian Daniel Haimovich, R. Andrew Taylor, Predicting hospital admission at emergency department triage using machine learning. PLoS One. 2018; 13(7): e0201016.

STUDYING THE EFFECT OF A BRIEF MEDITATION TRAINING ON PSYCHOLOGICAL INDICATORS OF WELL-BEING.

Teresa Fazia[1], Francesco Bubbico[1], Giovanni Berzuini[1], Salvatore Bruno[2], Luisa Bernardinelli[1]

[1] *Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy.*

[2] *Istituto di Psicosintesi, Milano, Italy*

Introduction

Meditation is an ancient practice which has its roots in India about 2500 years ago and it has been perpetuated until today in various form and through different tradition. Despite this, only in the last few decades the meditation has gained consideration in the academic world, given the presence of scientific evidences with regards to its beneficial effects on psychological, neurological, endocrine and immune variables, as well as positive influence on well-being and a broad range of biological processes [1-6]. For this reason, the meditation practise is today frequently integrated in treatment in contemporary psychotherapy and it is used to obtain enhancement in several medical conditions.

Objectives

The present study aims to assess the beneficial effects of a short-term intervention (a combination of 11 practical sessions of meditation one-per-week of 1-hour each), we called *Imaging Meditation*, on psychological indicators in healthy subjects. The choice to focus on healthy subjects was motivated by the observation that even though the clinical aspects of meditation are obviously more attractive to study and more powerful, a part of academic world started to investigate what could be the effects of meditation on non-clinical population.

Methods

Healthy women and men aged 18–67 were recruited for the study (n=220) and randomized to the intervention (n=116) arm or control arm (n=104). All study participants underwent an identical set of assessments, i.e completing all the study psychometric questionnaires both at t_0 and at t_1.
Psychometric questionnaires included: Core Outcomes in Routine Evaluation-Outcome measure (CORE-OM), Five Facets mindfulness Questionnaire (FFMQ), Life Satisfaction Scale (LSS), Positive Affect and Negative Affect Scales (PANAS), Perceived Stress Scale (PSS), Self-Compassion Scale (SCS), Subjective Happiness Scale (SHS), Warwick-Edinburgh Mental Well-Being Scale (WEMWBS).To evaluate pre - post changes between the two groups, each outcome was analyzed using linear mixed model effect with group (i.e. control and treatment), time (i.e., t_0 and t_1) and their interaction (time*group) as fixed effects. All models were adjusted for sex, age and previous meditation experience. Missing data were imputed and internal consistency of each questionnaire was assessed via Cronbach's α coefficients.

Results

These results highlighted a beneficial effect of meditation on: i) reducing psychological distress as measured using the CORE-OM questionnaires and its subscales (*all items* β=-0.33, p-value=1.62e^{-04}; *well-being* β=-0.55, p-value=1.94e^{-04}; *functioning* β=-0.32, p-value=5.54e^{-04}; *symptoms/problems* β=-0.38, p-value=6.35e^{-04}), ii) improving the level of mindfulness as measured using the FFMQ questionnaires and its subscales (*all items* β=0.27, p-value=5.79e^{-05}; *observe* β=0.31, p-value=6.90e^{-04}; *describe* β=0.34, p-value=2.12 e^{-04}; *non-judge* β=0.21, p-value=5.44e^{-02}; *non-react* β=0.29, p-value=2.21e^{-03}; *act with awareness* β=0.20, p-value=3.33e^{-02}), iii) improving subjective happiness as measured using the SHS questionnaires (β=0.36, p-

value=7.88e^{-03}), iv) improving the person's satisfaction as measured using the LSS questionnaires (β=2.59, p-value=1e^{-03}), improving the well-being in the sense of reducing the negative effect measured using the PANAS questionnaire's specific subscale (β=-2.44, p-value=2.26e^{-02}), and increasing the positive effect measured using the PANAS questionnaire's specific subscale (β=2.52, p-value=3.89e^{-03}), v) reducing the perceived stress as measured using the PSS questionnaires (β=-4.36, p-value=1.5e^{-05}), and lastly vi) improving the mental well-being as measured using the WEMWBS questionnaires (β=5.36, p-value=1.09e^{-05}).
No significant differences ($p<0.05$) were instead observed between the intervention and control group in the analysis of SCS questionnaire either considering all items or the specific subscales; thus resulting that our intervention does not act in improving the level of compassion.

Conclusions

Our results give evidence of a beneficial psychological effects of the short-term intervention in a healthy population. So practicing meditation in a big group have a very positive effects such as to improve the quality of life and reduce stress that is thought to be at the basis of numerous medical conditions.

References

[1] Aherne D, Farrant K, Hickey L, et al. Mindfulness based stress reduction for medical students: optimising student satisfaction and engagement. BMC Med Educ. 2016;16(1):209. 2.

[2] Bai Z, Chang J, Chen C, et al. Investigating the effect of transcendental meditation on blood pressure: a systematic review and meta-analysis. J Hum Hypertens. 2015;29(11):653–662.

[3] Chiesa A, Fazia T, Bernardinelli L, Morandi G. Citation patterns and trends of systematic reviews about mindfulness. Complement Ther Clin Pract. 2017;28:26–37.

[4] Paul-Labrador M, Polk D, Dwyer JH, et al. Effects of a Randomized Controlled Trial of Transcendental Meditation on Components of the Metabolic Syndrome in Subjects with Coronary Heart Disease. Arch Intern Med. 2006;166(11):1218.

[5] Rubik B. Neurofeedback-Enhanced Gamma Brainwaves from the Prefrontal Cortical Region of Meditators and Non-Meditators and Associated Subjective Experiences. J Altern Complement Med. 2011;17(2):109–115

[6] Wood K, Lawrence M, Jani B, et al. Mindfulness-based interventions in epilepsy: a systematic review. BMC Neurol. BioMed Central; 2017;17(1):52.

STUDIO RETROSPETTIVO DI UNA COORTE DI SOGGETTI ESPOSTI ALL'AMIANTO RESIDENTI IN PROVINCIA DI PAVIA

Silvia D. Visonà[1], Laura Gavarini[2], Simona Villani[2], Antonio Osculati[1]

[1] Unità di Medicina Legale e Scienze Forensi "Antonio Fornari", Dipartimento di Sanità Pubblica, Medicina sperimentale e Forense, Università degli Studi di Pavia

[2] Unità di Biostatistica ed Epidemiologia Clinica, Dipartimento di Sanità Pubblica, Medicina sperimentale e Forense, Università degli Studi di Pavia

Introduzione

Con il termine "amianto" si intende indicare un gruppo eterogeneo di minerali che, in seguito a frammentazione, danno luogo alla produzione di fibre [1]. E' noto che l'esposizione ad amianto determina l'insorgenza di alcune patologie che nel loro complesso sono definite "patologie amianto correlate" o "asbestos related diseases" (ARDs)[2]. Le manifestazioni patologiche ascrivibili all'amianto sono suddividibili in due gruppi: un primo gruppo, rappresentato da manifestazioni non neoplastiche, comprendente la caratteristica fibrosi polmonare nota come "asbestosi", e una serie di manifestazioni benigne a carico della pleura, le più frequenti delle quali sono rappresentate da tipici ispessimenti a placca ("placche pleuriche"). Il secondo gruppo è rappresentato da manifestazioni neoplastiche maligne, comprendente il mesotelioma della pleura e di altre membrane sierose ed il tumore del polmone. Dopo i primi studi epidemiologici che dimostrarono come l'insorgenza di patologie asbesto-correlate (in particolare il mesotelioma maligno) fosse correlata ad esposizioni ad amianto di tipo occupazionale [3-4], se ne ebbero altri nei quali si evidenziò una relazione anche con casi occorsi in familiari di lavoratori o residenti nelle zone limitrofe ai siti produttivi dei manufatti d'amianto [5].

Obiettivi

L'obiettivo dello studio è indagare la relazione tra la sopravvivenza al mesotelioma e l'esposizione all'amianto in una coorte di soggetti residenti a Broni (Pavia), Sito di Interesse Nazionale (SIN), e hinterland, in base al tipo di esposizione.

Metodi

È stato condotto uno studio retrospettivo su 188 individui morti per patologie asbesto correlate, accertate mediante autopsia giudiziaria presso la Sezione di Medicina Legale e Scienze Forensi dell'Università di Pavia fra il 2000 e il 2017. Una delle numerose industrie produttrici di amianto fu la Società Fibronit, ubicata a Broni, la quale intraprese la lavorazione dei manufatti in amianto dal 1932 al giugno 1993. La lavorazione prevedeva l'utilizzo di diversi tipi di amianto, che comportava l'immissione nell'ambiente di notevoli quantità di fibre definite, in base alle loro caratteristiche, respirabili; esse venivano, pertanto, inalate dai lavoratori e da coloro che risiedevano nella zona circostante lo stabilimento. La fonte dei dati è stata l'archivio cartaceo del settorato autoptico dell'UO di Medicina legale. Dall'archivio cartaceo sono stati estratte e implementate in un database le seguenti variabili: sesso, età della morte, residenza (Broni, hinterland, altro), abitudine al fumo, tipo di lavoro e mansione (laddove disponibile), causa di morte (mesotelioma, asbestosi, tumore del polmone, altri tumori e patologie cardiopolmonari) accertata autopticamente, tipo di esposizione all'amianto (occupazionale, familiare e ambientale), data di morte, data della diagnosi, data della prima esposizione (laddove disponibile).
Partendo dalla variabile causa di morte è stata costruita la variabile morte per mesotelioma, codificata 1 in caso di mesotelioma, 0 altrimenti. Il tempo di sopravvivenza dalla diagnosi è stato definito come il tempo tra

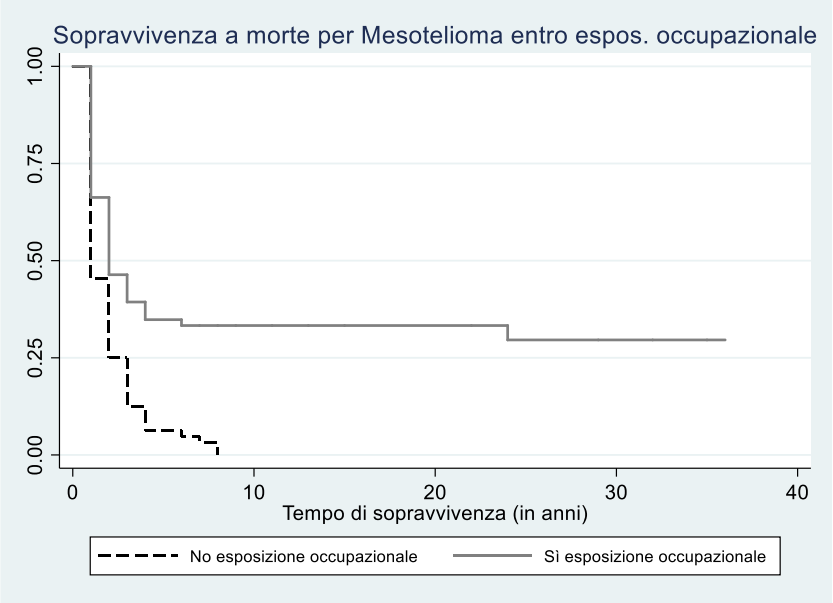
la data di morte e la data prima diagnosi (se concordante con la causa di morte) e quello di latenza come tempo tra la prima esposizione e la diagnosi.
Per le variabili quantitative sono stati usati come indici descrittivi quali la media e la deviazione standard; per le variabili qualitative sono state utilizzate le frequenze percentuali. Il test t Student per gruppi indipendenti è stato utilizzato per valutare le differenze in base al sesso e al tipo di esposizione nel caso di variabili quantitative. Il test Chi-quadrato o il test esatto di Fisher per valutare le differenze di variabili qualitative in base all'esposizione. L'analisi di sopravvivenza è stata stimata secondo Kaplan-Meier in base al tipo di esposizione. Il livello di significatività è stato fissato a 0.05. Le analisi sono state eseguite usando software STATA 15.

Risultati

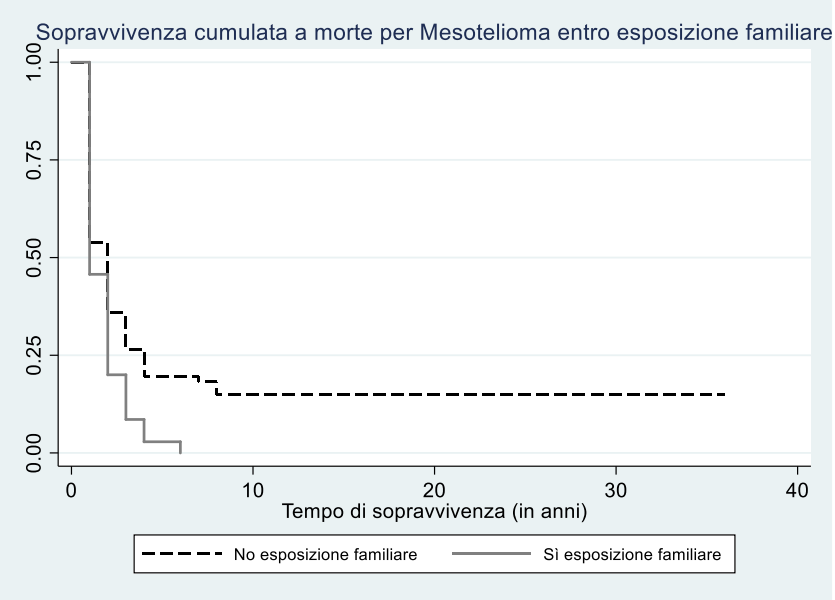
Il 66% dei soggetti erano maschi, l'età media alla morte era di 71.4 anni (±11.8 anni), significativamente più alta nei maschi che nelle femmine (72.8 anni±11.2 anni vs. 68.6 ± 12.8 anni) (t=2.34, p=0.02). La maggior parte dei soggetti al momento del decesso (97.3%) viveva in provincia di Pavia, era ex-fumatore o fumatore (64.6%).
L'84% dei soggetti è risultato avere avuto un'esposizione di tipo ambientale ad amianto, il 60% di tipo lavorativo, e il 27% di tipo familiare, in quanto congiunto di un ex lavoratore Fibronit. Le esposizioni non sono tra loro mutuamente esclusive: alcuni soggetti presentavano più di un'esposizione. Fra i soggetti con esposizione lavorativa il 90% sono risultati uomini contro il 28.8% fra quelli senza tale esposizione (Chi2=72.8%, p<0.001). Per quanto concerne i 110 soggetti con esposizione di tipo lavorativo ad amianto, il 70% erano stati occupati alla Fibronit, e nella quasi totalità erano maschi. Tra i soggetti con esposizione di tipo familiare si è trovata una frequenza di donne doppia rispetto al gruppo senza tale esposizione (62.2% vs 35%) (Chi2=8.2, p=0.004). La frequenza di donne non è risultata significativamente diversa tra chi è stato classificato come esposto a livello ambientale e chi è stato classificato come non esposto (38.8% vs 21.4%) (Chi2=3.1, p=0.08).
La maggior parte dei soggetti è deceduta per mesotelioma (79.7%), il 10.7% per asbestosi e nei restanti casi patologie asbesto-correlate. Fra i soggetti con esposizione di tipo lavorativo il 67% è morto per mesotelioma contro la totalità di quelli per i quali tale esposizione non si è verificata (Chi2=30.1, p<0.001). Tra quelli con esposizione familiare la totalità è morta per mesotelioma contro l'82.1% di quelli per i quali tale esposizione non è stata riscontrata (p esatto di Fisher = 0.011). Non sono state trovate differenze nell'occorrenza di morte per mesotelioma in base all'esposizione di tipo ambientale (82.1% vs 81.5%, p esatto di Fisher>0.90).
In generale i maschi presentano tempo di latenza maggiore delle femmine, ma la differenza tra i due sessi non è statisticamente significativa (t=0.88, p=0.38).
La sopravvivenza alla morte per una forma di mesotelioma dal momento della diagnosi è significativamente superiore nei soggetti che avevano avuto una esposizione di tipo lavorativo rispetto a quelli per i quali tale esposizione è stata esclusa (test di Tarone-Ware=16.49; p<0.0001) (Figura 1).
Nel caso di esposizione indiretta da familiare i soggetti per i quali è stato possibile accertarla sono risultati avere una sopravvivenza inferiore a quella dei soggetti per cui è stata esclusa (Figura 1), per quanto al limite della significatività statistica (test di Tarone-Ware= 3.87; p=0.05).

Figura 1. *Stima della sopravvivenza cumulata a morte per mesotelioma in base al tipo di esposizione.*

a)



b)



Conclusioni

Lo studio ha evidenziato come nei soggetti lavoratori a contatto con l'amianto, sebbene la maggior parte (67% circa) sia deceduta per una forma di mesotelioma, una considerevole porzione del sottogruppo non ha sviluppato mesotelioma, ma è invece deceduto per asbestosi (17.4% dei casi) o per altre patologie. Sorprendentemente, invece, tutti i soggetti privi di esposizione lavorativa, e, pertanto, esposti a basse dosi di amianto, sono deceduti per mesotelioma.
Considerando che la concentrazione di fibre aerodisperse all'interno di una fabbrica dove viene lavorato l'amianto [3] non è nemmeno lontanamente paragonabile, quantitativamente, a quella rilevabile nell'ambiente circostante. Risulta evidente come i soggetti che non lavorarono mai all'interno della Fibronit furono esposti a dosi di amianto sicuramente più basse. Ciò dovrà trovare conferma nella sistematica analisi in microscopia elettronica a scansione (SEM EDS) del residuo inorganico nei reperti autoptici dei polmoni dei soggetti, (attualmente in corso).
Il successivo risultato interessante deriva dall'analisi della sopravvivenza. Quanto emerso suggerisce che, gli individui esposti a basse dosi di amianto siano andati incontro ad un decorso più aggressivo della patologia. Siffatte osservazioni portano a corroborare l'ipotesi dell'esistenza di elementi di suscettibilità individuale, forse di natura genetica, che incidono significativamente nell'eziopatogenesi del mesotelioma.
Infine, analogamente, la pressoché totalità delle donne appartenenti al campione, pur essendo state sottoposte a dosi molto ridotte di amianto, ha sviluppato il mesotelioma (dato non riportato).
In conclusione, il ruolo dell'amianto nell'eziopatogenesi delle malattie amianto correlate è tuttora oggetto di acceso dibattito scientifico e molti aspetti della questione devono essere, ancora, indagati.

Bibliografia

[1] American Thoracic Society. Diagnosis and Initial Management of Nonmalignant Diseases Related to Asbestos. American Journal of Respiratory and Critical Care Medicine, 2004 Sep 15; 170 (6): 691–715.

[2] IARC (International Agency for Research on Cancer). Asbestos (chrysotile, amosite, crocidolite, tremolite, actinolite, and anthophyllite). Monography 100C, 2012.

[3] Selikoff I.J., Churg J., Hammond E.C. Asbestos Exposure and Neoplasia. JAMA, 1964 Apr; 6:188:22-6.

[4] Wagner J.C., Sleggs C.A., and Marchand P. Diffuse Pleural Mesothelioma and Asbestos Exposure in the North Western Cape Province. British Journal of Industrial Medicine, 1960 October; 17: 260–71.

[5] Churg A. Current Issues in the Pathologic and Mineralogic Diagnosis of Asbestos-Induced Disease. Chest, 1983 Sep; 84 (3): 275–80.

EXPOSURE TO CLEANING AGENTS AT HOME OR AT WORK DECREASES LEVELS OF FRACTIONAL EXHALED NITRIC OXIDE (FeNO). THE GEIRD STUDY.

Pierpaolo Marchetti[1], Lorena Torroni[1], Francesca Locatelli[1], Mario Olivieri[2], Giuseppe Verlato[1]

[1]*Unit of Epidemiology and Medical Statistics, Department of Diagnostics and Public Health, University of Verona, Verona, Italy.*

[2]*Unit of Occupational Medicine, Department of Diagnostics and Public Health, University of Verona, Verona, Italy.*

Introduction

Frequent use of cleaning products has been associated with respiratory symptoms, asthma diagnosis and lung function decline in housewives and professional cleaners [1]. Fractional exhaled Nitric Oxide (FeNO) is a biomarker of airways inflammation.

Aims

To investigate whether FeNO is affected by exposure to cleaning products at home and in occupational cleaners.

Methods

Information about occupation was collected in a general population sample (n=1227), enrolled in Verona within the clinical phase of GEIRD (Gene Environment Interaction in Respiratory Diseases) case-control study. Based on the answer to a "filtering question" for the 4 modules (Cleaning and/or washing in the home, all types of cleaners, all types of nurses, disinfection worker), we classified subjects in three groups: 1) not cleaning if they answered "no" to all filtering questions of occupational modules (n=558); 2) cleaning at home if they answered "yes" only to "cleaning and/or washing in the home" module (n=454); 3) occupational cleaners if they answered "yes" to at least one of three modules "all types of cleaners", "all types of nurses" and "disinfection workers" (n=215).
Airway inflammation was assessed by FeNO, measured in accordance with ATS/ERS guidelines (ppb at 50 ml/sec), while lung function by FEV_1/FVC percent.
Frequency of cleaning products (detergents and sprays) was expressed as a semi-quantitative total score [2]. The influence of cleaning product use on FeNO and FEV_1/FVC percent was investigated by median regression, controlling for sex, age, BMI, level of education (high/low), smoking habits (no/former smoker/current smoker), atopy (yes/no), respiratory disease (none/rhinitis/asthma/chronic bronchitis).

Results

Overall 1227 subjects were included in the analysis. In the two groups of cleaners, the proportion of women were similar (80% in cleaning at home and 70% in occupational cleaners) while being 21% in the not cleaning group ($p<0.001$). Low level of education was more frequent in occupational cleaners than in cleaning at home and not cleaning subjects (26.4% vs 14.8% and 12.5% respectively; $p<0.001$). The levels of FeNO were significantly lower in the cleaning groups than in not cleaning subjects (median [IQR] 14.9 [10.2-25.6] ppb in occupational cleaners, 15.8 [10.5-23.8] ppb in cleaning at home and 18.9 [12.2-33.3] ppb in not cleaning; $p<0.001$).
In a multivariable analysis on subjects cleaning at home, using at least one detergent per ≥1day/week was associated with a median decrease in FEV_1/FVC percent of 3.21 (0.67-5.76; $p=0.01$) (Table 1). Median

FeNO decreased by 0.23 ppb (95% CI 0.01-0.45, p=0.039) and by 0.17 ppb (95% CI 0.06-0.29, p=0.004) respectively, when the weekly use of cleaning sprays and all cleaning products increased by one day. Among occupational cleaners, FeNO and FEV_1/FVC percent were not significantly affected by frequency of cleaning product use.

Conclusion

Exposure to cleaning products, both at home and in the workplace, is associated with lower FeNO levels, suggesting an irritant effect on the airways. Moreover, frequent use of cleaning products at home is associated with lower levels of both FeNO and FEV_1/FVC percent.

Table 1. *Effect of frequency of cleaning product use on FeNO and* ***FEV_1/FVC,*** *assessed by quantile regression†*

	Cleaning at home only (n=454)			**Occupational cleaners (n=215)**		
	n(%) users	FeNO (ppb)	FEV_1/FVC (%)	n(%) users	FeNO (ppb)	FEV_1/FVC (%)
At least 1 product per ≥1d/w						
Detergents	414 (91.2)	-0.10 (-3.87;3.66)	-3.21 (-5.76; -0.67)*	180 (83.7)	1.81 (-6.65;10.27)	-1.71 (-4.84;1.42)
Sprays	233 (51.3)	-1.41 (-3.66;0.85)	-0.69 (-2.23;0.86)	142 (66.1)	0.65 (-6.24;7.54)	-0.84 (-3.21;1.53)
Cumulative days use (Per one day increase)						
Detergents		-0.18 (-0.39;0.18)	-0.13 (-0.28; 0.20)		0.07 (-0.22;0.36)	0.07 (-0.22;0.36)
Sprays		-0.23 (-0.45;-0.01)*	0.01 (-0.16;0.18)		0.06 (-0.31;0.44)	0.06 (-0.31;0.44)
All cleaning products		-0.17 (-0.29;-0.06)**	-0.07 (-0.16;0.02)		0.04 (-0.14;0.22)	-0.03 (-0.08;0.02)

*p<0.05;**p<0.01. †adjusted for sex, age, bmi, atopy, respiratory diseases, level of education, smoking habits.

Bibliography

[1] Svanes Ø., Bertelsen R.J., Lygre S.H.L., et al., Cleaning at Home and at Work in Relation to Lung Function Decline and Airway Obstruction. Am J Respir Crit Care Med., 2018;197(9):1157-1163.

[2] Casas L., Zock J.P., Torrent M., et al., Use of household cleaning products, exhaled nitric oxide and lung function in children. Eur Respir J 2013; 42:1415-1418.

UN ELEMENTO DI VERIFICA E DI VALUTAZIONE DELL'APPRENDIMENTO NEI CORSI DI LAUREA MAGISTRALE IN MEDICINA E CHIRURGIA: DAL PROGRESS TEST AL TRAINING TEST, ELABORAZIONE DATI FINALI.

Laura Recchia[1], Bruno Moncharmont[2]

[1] *Coordinatrice U.G.Q. (Unità di gestione della Qualità) e Vicepresidente del Corso di Laurea Magistrale in Medicina e Chirurgia, Università degli studi del Molise*

[2] *Vice Presidente Vicario della Conferenza Permanete dei Presidenti di Consiglio di Corso di laurea magistrale in Medicina e Chirurgia (CPPCLMM&C).*

Introduzione

Tra le azioni di maggior rilievo promosse dalla Conferenza Permanente dei Presidenti di CCLM in Medicina (CPPCLMM&C) si può inserire sicuramente la creazione ed implementazione del *progress test* (PT), un complesso di 300 domande a risposta multipla sottoposto a cadenza annuale contemporaneamente sul territorio nazionale a tutti gli studenti dei Corsi di Laurea in Medicina dal 2006 e in Odontoiatria e Protesi Dentaria dal 2017 [1].Gli straordinari risultati ottenuti hanno ispirato gli estensori del D.M. del 9 maggio 2018, n. 58 del MIUR, in cui è enunciato che l'esame di Stato '*assumerà la forma di un test con domande a scelta multipla non note in anticipo ai candidati'* [2]. Il *progress test è stato dunque 'riprogettato' per il 2018 in base a* caratteristiche formali che lo hanno parzialmente avvicinato al nuovo esame di stato: composto da 200 domande, ripartite in 50 formulate su argomenti riguardanti le conoscenze di base, nella prospettiva della loro successiva applicazione professionale e 150 allestite su argomenti riguardanti la capacità di applicare le conoscenze biomediche e cliniche alla pratica medica ed alla risoluzione di una serie di problemi clinici. Il test di quest'anno è così diventato un *training test (TT),* ovvero un allenamento in vista del futuro esame di stato.

Obiettivi

Obiettivo del *training test è stato* allenare gli studenti al superamento di un esame di stato nazionale e quindi non sono state estratte le domande attingendo a grandi e qualificate banche-dati internazionali ma è stato necessario attenersi strettamente al '*core curriculum'* dei corsi di laurea in Medicina, elaborato dalla CPPCLMM&C .A questo scopo la Conferenza ha impegnato tutti i Presidenti di CLM a fornire in modo riservato una serie di domande aderenti a ciò che effettivamente si insegna nei nostri CLM. Hanno contribuito a questo progetto ben 59 esperti disciplinari (che hanno redatto le domande a risposta multipla) e 10 esperti trasversali (che hanno confezionato il test). Il riferimento per la formulazione dei quesiti è stato il '*core curriculum'*, ovvero i contenuti dell'abituale insegnamento di ogni docente ed i quesiti sono stati selezionati, seguendo il metodo 'PUIGER', ovvero scegliendo temi rilevanti per la **P**revalenza del soggetto, o la sua **U**rgenza, o le possibilità di **I**ntervento ad esso connesse, o la sua **E**semplarità metodologica, o la sua **R**ilevanza sociale. Altra caratteristica essenziale nella formulazione delle domande è stata l'accessibilità alla risposta esatta per uno studente di 'capacità-soglia'. Il TT infatti non è stato creato per individuare e selezionare un élite di studenti particolarmente meritevoli ma quale strumento di monitoraggio della preparazione complessiva del gruppo classe e, indirettamente, sulla capacità del docente di portare l'intera classe ad un livello base di conoscenza. Si è imposta, dunque, una sostanziale modifica rispetto al classico *progress test, utile per* valutare il *progresso* dell'apprendimento dello studente (o della coorte) da un anno di corso all'altro, necessariamente basato sulle discipline insegnate nei vari anni [3]. Al contrario, il test previsto dal citato DM del 9 maggio 2018 si prefigge di valutare le *competenze in uscita* dello studente, quali la capacità di prendere decisioni e risolvere problemi e si preoccupa di controllare l'acquisizione delle scienze di base principalmente tramite la loro ricaduta sulla competenza clinica. Di conseguenza, nel TT,

rispetto al *PT,* non solo è mutato il numero e la distribuzione delle domande ma si è imposta una *contestualizzazione* clinica delle stesse. Nel commissionare le domande di base si sono quindi privilegiate le formulazioni problematiche e nel redigere le domande cliniche non ci si è attenuti a criteri "per discipline" ma per quattro 'macro-aree' di contesto clinico: *in nosocomio*, *in ambulatorio*, *in emergenza/urgenza* e *nel territorio*. Obiettivo di questo articolo è una prima elaborazione dei dati grezzi pervenuti dalle singole sedi la cui interpretazione dei risultati ci indicherà '*quanto abbiamo colto nel segno*.

Metodi

Il numero di partecipanti al PT è aumentato nei diversi anni accademici in forma esponenziale, passando da 3.496 studenti nel 2006 a 37.143 studenti in corso più 480 fuori corso per un totale di 37.623, nel 2018. Hanno infatti partecipato tutti i 49 corsi di studi erogati in lingua italiana ai quali si sono aggiunti 10 dei quattordici corsi di studi erogati in lingua inglese, con il conseguente aumento del numero delle risposte e della mole di dati da raccogliere ed elaborare. Il campione analizzato è composto da 33.394 studenti i cui risultati sono stati inviati dalle diverse sedi oggetto della prova e sintetizzati sia per l'area preclinica che clinica, i dati sono stati raccolti dalle singole sedi su due matrici studente/domanda nella quale era riportato

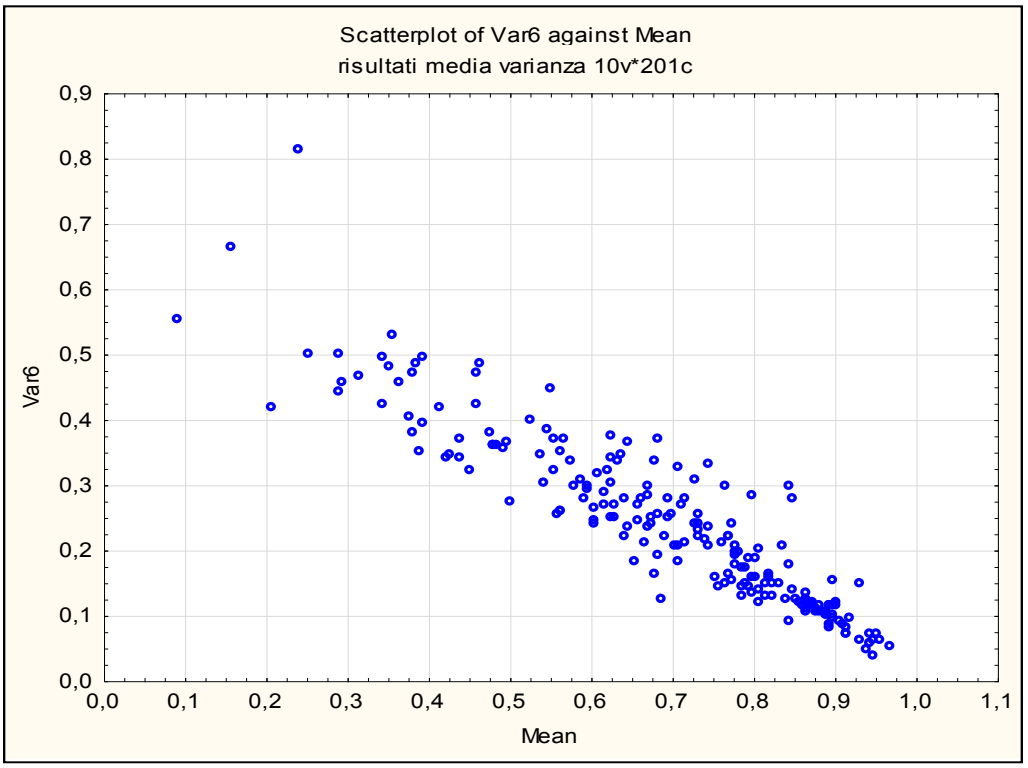
1. il risultato della correzione [1, risposta esatta; 2, risposta errata; 3, nessuna risposta];
2. la risposta data dallo studente alla domanda [A; B; C; D; E].

I files sono stati poi inviati al Presidente della CPPCLMM&C per consentirne un'adeguata elaborazione. Ai dati grezzi pervenuti sono stati applicati gli Indicatori di Facilità / Difficoltà, la formula KR-20 per la valutazione dell'affidabilità della prova [4] e per ogni domanda (Qx) sono state calcolate le percentuali di risposte esatte per tutti gli anni, dal primo al sesto e per tutte le sedi pervenute. Per ogni domanda è stata misurata la variabilità dell'indice di facilità riscontrato nelle diverse sedi. Per esempio, la domanda che ha fatto registrare la massima variabilità nel tasso di risposte esatte è stata la domanda 56 (Quale farmaco viene impiegato per la profilassi della condizione di Eclampsia?) che, nelle diverse sedi, è stata risolta da percentuali di studenti che variano dallo 0 all'87%. La variabilità di risultato è stata espressa con tre parametri diversi: la varianza, la deviazione standard e il coefficiente di variazione [5]. Abbiamo voluto verificare se la tipologia di domanda utilizzata e il suo ambito culturale avessero un'influenza sulla "facilità" con la quale gli studenti hanno risposto. Per fare questo, abbiamo isolato il sottoinsieme delle domande "difficili" includendo tra le domande "difficili" quelle relative al 25° percentile delle domande con l'indice di facilità minore. Le 50 domande con indice di facilità più basso (25% sul campione di 200 domande) variavano tra un indice di facilità di 0.089 (il più basso) ed uno di 0.588.

Risultati

Ventidue domande su 50 appartenevano alla classe di quesiti indirizzati a vagliare la capacità interpretativa dello studente, ovvero l'abilità di interpretare dei referti di laboratorio (10), radiologici (4), elettrocardiografici (2), ecografici (2), spirometrici (1), emogas -analitici (1) e perfino le risultanze dell'esame obiettivo (2). Un altro gruppo corposo (17/50) di domande risultate difficili riguardavano quesiti sulla prescrizione di interventi terapeutici (13) o di esami strumentali (4). Le rimanenti 11 domande cadute all'interno del 25° percentile potevano essere ricondotte a difficoltà differenti, come quella di interpretare nessi patogenetici (4) o di possedere rudimenti-base di statistica o epidemiologia (2). In una sola domanda è ipotizzabile che lo studente sia stato tratto in errore dalla formulazione del quesito: si chiedeva quale tra le lipoproteine avesse la densità minore ma i "distrattori" non erano stati elencati in ordine di densità, come sarebbe stato lecito attendersi. L'analisi dei dati ha dimostrato una stretta correlazione tra difficoltà di una domanda e variabilità con cui questa è stata accolta nelle diverse sedi.

Figura1. *Distribuzione dell'di facilità delle singole domande*



Conclusioni

Dall'elaborazione dei dati emerge un andamento decisamente crescente del numero delle risposte corrette all'aumentare dell'anno di corso e traspare quindi un'ottima conoscenza di base approfondita nel triennio clinico. Le domande che sono risultate più "difficili" da risolvere riguardavano i casi problematici, quelle più "facili" le domande conoscitive a carattere interdisciplinare. Ciò dimostra che gli studenti Italiani non sono ancora abbastanza allenati a ragionare "per problemi" ma nella verifica delle loro conoscenze, appaiono meglio acquisite quelle che hanno imposto un ragionamento interdisciplinare. La difficoltà incontrata dagli studenti Italiani nel rispondere a domande atte a valutare la capacità interpretativa e quella di risolvere problemi (fare prescrizioni) mette in luce un limite non ancora superato nelle metodologie di insegnamento messe in atto nei CLM Italiani, che sono ancora troppo "teoriche" e sbilanciate a favore dell'acquisizione di competenze conoscitive rispetto alle abilità interpretative e alle meta-competenze cliniche [6.] E' necessario inoltre individuare elementi per poter standardizzare e rendere omogenei i dati e procedere ad un corretto confronto tra i risultati delle diverse sedi statisticamente significativi. Bisognerebbe inoltre implementare un nuovo approccio valutativo dei dati che differisce sostanzialmente da quello in uso che si concentra sulla valutazione dei valori medi, individuando invece le "code" meno virtuose e più problematiche. Questo sistema rispetto alla valutazione dell'eccellenza, presenta maggiori caratteri di oggettività e risulterebbe indubbiamente più utile [7].

Bibliografia

[1] Crocetta C, Brindisi M, Lo Muzio L: Analisi dei Risultati del Progress Test 2017 dei Corsi di Laurea in Odontoiatria e Protesi Dentaria. Med. Chir. 78: 3487-3493, 2018.

[2] Lenzi A, Familiari G, Basili S, Moncharmont B, Consorti F, Nati G, Della Rocca C, Gallo P: Progress /Training test 2018. Quesiti e risposte Med. Chir. 80: 3546-3574, 2018.

[3] Gallo P: Cosa cambia con la laurea abilitante per la Professione medica Tra progress test e training test. Med. Chir. 79: 3524, 2018

[4] Dancey C P, Reidy J G, Rowe R: Statistica per le scienze mediche, un approccio non matematico. Piccin-Nuova Libraria, Padova, 2016.

[5] Mucciarelli G, Chattat R, Celani G: Teoria e pratica dei test. Piccin-Nuova Libraria, Padova, 2002.

[6] Tenore A, Basili S, Lenzi A: Il Progress Test 2016. Med. Chir 75: 3386-3390, 2017.

[7] V.Nesi et al. Sapienza Roma “Un elemento di valutazione della qualità della didattica” Maggio 2016

IDENTIFICAZIONE E INTERVENTO SU BAMBINI A RISCHIO DI DISTURBI DEL LINGUAGGIO E/O DEL NEUROSVILUPPO, FIGLI DI MIGRANTI: USABILITÀ DEL PVB-BIL

Simonelli Ilaria[1], Pasqualetti Patrizio[1], Rinaldi Pasquale[2], Cristiana Sanalitro[2], Maria Cristina Caselli[2], Lorena Remi[3], Paola Pettenati[3], Maria Elena La Marchina[4], Silvia Stefanini[4]

[1] Fondazione Fatebenefratelli per la Ricerca e la Formazione Sanitaria e Sociale, Divisione Ricerca, 2 ISTC-CNR, 3 ATS-Val Padana, 4 AUSL di Parma

Introduzione

L'individuazione di indici di rischio e profili atipici nello sviluppo linguistico di bambini bilingui figli di migranti risulta essere difficile a causa della mancanza di strumenti che permettono di valutare le competenze linguistiche nelle diverse lingue parlate da questi bambini. In tale ambito nasce il progetto INTEGRA, proponendosi di effettuare uno screening della comunicazione e del linguaggio di bambini bilingui (18-42 mesi) figli di migranti a rischio di un disturbo del linguaggio e/o del neuro-sviluppo, senza patologie già accertate.

Obiettivi

Lo studio vuole verificare l'usabilità della scheda "Primo Vocabolario del Bambino-Parole e Frasi-Forma Breve" ideata per bambini bilingui (PVB BIL) e la sua capacità di misurare le loro competenze linguistiche. La scheda è l'adattamento della corrispondente Forma Breve del questionario "Primo Vocabolario del Bambino – (PVB P&F) disponibile per i bambini italiani (1). La principale ipotesi di lavoro è che il numero di parole in produzione rilevate con la forma ideata per bambini bilingui sia in accordo con il numero di parole in produzione misurato con il metodo che attualmente rappresenta il gold-standard, e cioè la somministrazione di due schede, una nella Lingua di Origine e una in italiano. Se si dimostrasse un adeguato "agreement" tra le ampiezze di vocabolario in ciascuna delle due lingue calcolate con la forma per bilingui e quelle calcolate con le schede in ciascuna delle lingue, si potrebbe evitare di utilizzare le due schede (nei casi in cui non sia possibile somministrarle entrambe) e sostituirle con la sola scheda per bilingui. Se inoltre il grado di agreement non risultasse variabile in funzione della lingua d'origine (e dell'etnia), si potrebbe utilizzare la forma per bilingui anche per quelle lingue per cui la scheda di rilevazione nella lingua di origine non esiste o non è stata ancora validata.

Metodi

Sono state raccolte informazioni su 43 bambini di età media 35 (sd=5.56) mesi, per il 60% (26/43) maschi e nel 42% (18/43) la lingua d'origine è l'arabo (13 arabo marocchino, 4 arabo tunisino e 1 arabo algerino). Le competenze linguistiche sono state rilevate attraverso la compilazione contestuale dei questionari per i genitori MacArthur-Bates Communicative Development Inventories nella Lingua di Origine dei genitori (MB-CDI LO) e della Forma Breve del PVB-P&F (PVB IT). Ai genitori è stato inoltre somministrato il PVB BIL. Per misurare l'affidabilità è stato calcolato l'Intraclass coefficient (ICC) attraverso un modello ad effetti misti.

Risultati

L'analisi ha evidenziato una elevata correlazione positiva tra il numero di parole rilevate con il PVB IT e il numero di parole in italiano rilevate con il PVB BIL ($r=0.964$; $p<0.001$) anche l'affidabilità risulta essere eccellente (ICC=0.98, 95% IC 0.97-0.99). Invece, la correlazione tra il numero di parole rilevate con il MB-CDI LO e il numero di parole nella lingua d'origine rilevate con il PVB BIL è risultata pari a $r=0.370$ ($p=0.029$). In questo caso si osserva anche una scarsa affidabilità (ICC= 0.23, 95% CI 0.0-0.61).

Conclusioni

L'utilizzo del PVB BIL non comporta alcuna perdita di informazione rispetto alle competenze linguistiche in italiano. Come atteso (trattandosi di una traduzione del questionario italiano e non di un adattamento negli altri contesti culturali e linguistici), la correlazione rispetto alle competenze nella lingua d'origine è risultata più bassa, anche se statisticamente significativa e in grado di aggiungere informazioni utili per avere un profilo completo delle competenze comunicative dei bambini bilingui, laddove non sia semplice reperire strumenti che valutino l'altra lingua parlata dal bambino. Riuscire a valutare nei bambini bilingui le competenze in entrambe le lingue può permettere di distinguere i bambini che hanno un basso livello di vocabolario in italiano a causa di uno scarso input in questa lingua da quelli che invece hanno un basso livello di vocabolario in italiano a causa di un ritardo di linguaggio.

Bibliografia

[1] Caselli M.C., Bello A., Rinaldi P., Stefanini S., Pasqualetti P. (2015) ***Il Primo Vocabolario del Bambino: gesti, parole e frasi.*** Milano: Franco Angeli

STATISTICAL, CLINICAL AND ETHICAL ISSUES FOR DESIGN AND ANALYSIS OF SINGLE-ARM MULTI-STAGE PHASE II STUDIES IN ANESTESIOLOGY

Cofini Vincenza[1], Fusco Pierfrancesco[2], Petruccci Emiliano[2], Di Carlo Stefano [2], Necozione Stefano[1]

[1] *Department of Life, Health and Environmental Sciences, University of L'Aquila.*

[2] *San Salvatore Hospital, ASL n.1 Abruzzo, L'Aquila.*

Introduction

Although phase II trials have been very widely conducted and published for cancer clinical research, they have still been little used in the anesthesiology field [1].
To investigate the effects of the combination of ultrasound-guided pectoral nerve block II type and ultrasound-guided parasternal block in breast cancer surgery, a phase-two study would be necessary and adequate.

Objectives

Aim of this study was to evaluate the applicability of the Simon model in the anesthesiology and the related statistical, clinical and ethical issues.

Methods

This single arm phase-two-study was approved by Institutional Ethic Committee and registered with the US clinical trial registry (NCT03016117).
Women aged 18-75 years, undergoing unilateral quadrantectomy without axillary dissection with ASA physical status I–III, were eligible.
The primary outcome measure was the proportion of patients with sensory blockade from T2 to T6 breast dermatomes, assessed with cold and touch test [2].
Secondary outcomes were: mean duration of sensory blockade, drugs consumption and block-related complications, intraoperative and post-operative complications, incident and rest pain assessed with the Numeric Rate Scale of pain.
Statistical analysis
The sample size was 65 and it was based on Simon's optimal two-stage design[3,4]. Considering the primary outcome, the first stage required that 12 or more patients out of 13 had a confirmed response, to rule out an undesirably low response probability of 0.85 (p0) in favor of a desirable response probability of 0.95 (p1), with a type I error rate of 5% and a power of 80%, before proceeding to enroll other 52 patients (figure 1).
Data analysis
Descriptive statistics were calculated for all variables. The chi-square or Fisher's exact test was used for categorical variables. Continuous variables were analyzed using either t-test or the Wilcoxon rank-sum test. For the analysis of the primary outcome, one-sided significance level of p=0.05 was specified in the protocol to provide an estimate of the effect of the block's combination, particularly with a focus on the lower bound for this effect[5].
Incident and rest pain were analyzed with repeated-measures analysis of variance. Statistical analysis was performed using STATA 14 software.

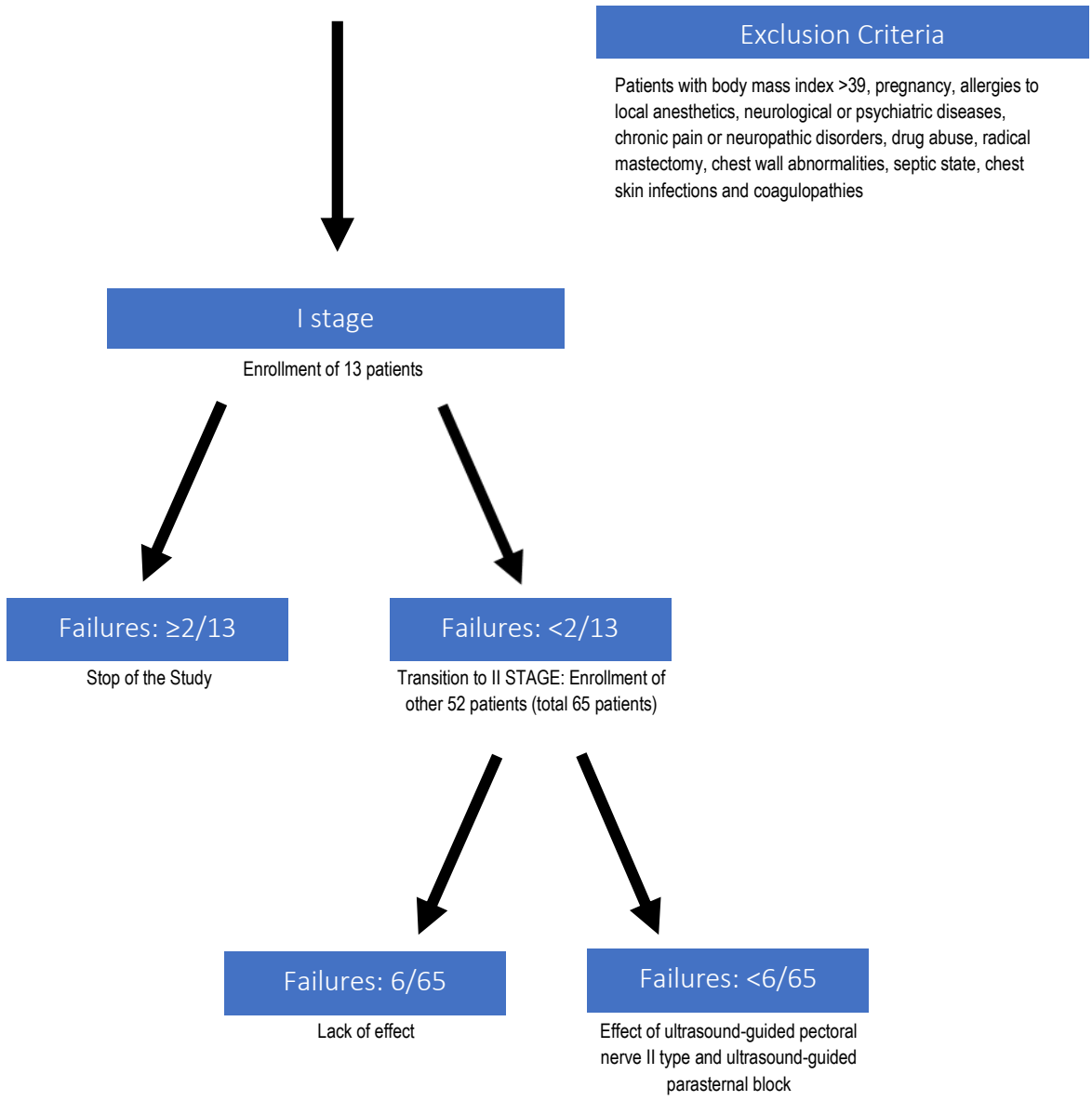
Results

Sixty-five patients were enrolled and received ultrasound-guided pectoral nerve block II type combined with ultrasound-guided parasternal block. The sensory blockade proportion of the combined blocks was 95% [62/65; 90%CI: 89%-99%].
The mean duration of sensory blockade was 9±2h. The mean score for postoperative pain eight hours after surgery was 2±2. The RM- ANOVA analysis indicated that there was a significant decrease for incident pain ($F_{time4,244}=18$, $p<0.001$) and for resting pain ($F_{time4, 244}=16$, $p<0.001$).
The total analgesic consumption, recorded in equianalgesic mg of morphine, was 126mg. 3/62 patients had a mild gastrointestinal burning. No intraoperative or post-operative complications were observed.

Conclusions

Simon's model can be a valid tool for assessing the anesthetic effect of the combination of ultrasound-guided pectoral nerve II type and ultrasound-guided parasternal block in breast surgery. The single arm multi-stage phase II study helps to respect the ethical principles and it is functional to design a phase-three randomized controlled study.

Figure 1. *Sample size with Two Stage Simon Design*



References

[1] Sin-Ho Jung. Statistical issues for design and analysis of single-arm multi-stage phase II cancer clinical trials. Contemporary Clinical Trials 42 (2015) 9–17

[2] Ousley R, Egan C, Dowling K, Cyna AM. Assessment of block height for satisfactory spinal anaesthesia for caesarean section. Anaesthesia. 2012, 67, 1356–1363; Kocarev M, Watkins E, McLure H, Columb M, Lyons G. Sensory testing of spinal anaesthesia for caesarean section: differential block and variability. Int J Obstet Anesth. 2010;19(3):261-5

[3] Simon R. Optimal two-stage designs for phase II clinical trials. Control Clin Trials. 1989;10(1):1-10.

[4] Guogen Shan, Hua Zhang and Tao Jiang. Minimax and admissible adaptive two-stage designs in phase II clinical trials. BMC Medical Research Methodology (2016).

[5] William QM, Gerald JH, Luis AE. Statistical Intervals: A Guide for Practitioners and Researchers. John Wiley & Sons, 2017. ISBN 1118595165, 9781118595169.

CARATTERIZZAZIONE DEL PROFILO PSICO-SOCIOLOGICO DEGLI STUDENTI IMMATRICOLATI AI CORSI DI LAUREA IN MEDICINA E CHIRURGIA PRESSO "SAPIENZA" UNIVERSITÀ DI ROMA. RISULTATI PRELIMINARI

Nofroni Italo[1], Vestri Anna Rita[1], Ferrara Carla[1], Cattaruzza Mariasofia[1], Liuccio Michaela[2], Carpi Matteo[3]

[1] *Dip. Sanità pubblica e Malattie infettive, Sapienza, Università di Roma*

[2] *Dip. di Comunicazione e Ricerca sociale, Sapienza, Università di Roma*

[3] *Dip. Neuroscienze Umane, Sapienza, Università di Roma*

Introduzione

La depressione è la causa principale di disabilità nel mondo e apporta il maggior contributo tra le cause del Global Burden of Disease. Si valuta che circa il 50% dei disturbi mentali abbia origine nell'adolescenza, ragion per cui "le organizzazioni che si occupano dei giovani devono essere attrezzate per riconoscere i sintomi e reagire tempestivamente". Viene inoltre sottolineato che i giovani adulti si confrontano con lo stress in termini di preoccupazione sulle performance scolastiche, l'ammissione agli studi universitari nonché la sicurezza di un lavoro e di un reddito. La letteratura internazionale che riguarda i giovani adulti si concentra spesso sugli studenti universitari, valutando i metodi d'ammissione alle diverse facoltà, con particolare attenzione per le facoltà di Medicina (Kreiter, 2013), sia valutandone le abilità cognitive (Lia, 2013) che la prevalenza di disturbi psicopatologici (Bore, 2016; Fiori Nastro, 2013).
È sembrato pertanto di estremo interesse effettuare un'indagine pressoché completa sugli immatricolati ai corsi di laurea in Medicina e Chirurgia presso Sapienza, Università di Roma.

Obiettivo

Questo studio prende in considerazione gli studenti universitari iscritti al primo anno dei Corsi di Laurea in Medicina e Chirurgia della Sapienza nell'AA 2017/18, per valutare, tramite quattro questionari, le caratteristiche di personalità, l'intelligenza emotiva e organizzativa e il funzionamento globale attraverso la valutazione di eventuali sintomi e patologie, anche sotto soglia. Si sono inoltre chieste informazioni sui molteplici comportamenti, anche a rischio, che caratterizzano il neo studente di una Facoltà medica.

Materiali e metodi

I quattro questionari di rilevazione dati sono stati somministrati in aula, in occasione delle lezioni prima della fine del primo semestre dai Docenti coinvolti nello studio. L'adesione da parte degli studenti alla compilazione dei questionari è stata comunque su base volontaria e incentivata tramite una presentazione dello studio e delle relative finalità da parte dei Docenti stessi. È stato garantito l'anonimato dei rispondenti, fornendo comunque loro possibilità di "riconoscere" in una seconda fase, tramite un codice personalizzato, le risposte da loro fornite ai vari items.
Coerentemente con gli obiettivi dello studio, sono stati adoperati tre strumenti *self-report*:
Questionario *Short Form 12* (**SF-12**; Ware, Kosinski, Keller 1996; Apolone et al., 2005), che indaga la qualità della vita percepita e fornisce due indici sintetici per il dominio della salute fisica (PCS-12) e per quello della salute mentale (MCS-12).
Questionaro *Psychological General Well-Being Index* (**PGWBI**; Dupuy, 1984; Grossi et al., 2002), che permette di ricavare un punteggio di benessere psicologico generale e sei punteggi specifici di ansia, depressione, benessere positivo, autocontrollo, salute generale e vitalità.

Questonario *Perceived Stress Scale* (**PSS**; Cohen & Williamson, 1988), che rileva lo stress psicologico percepito. Mentre per SF-12 un punteggio maggiore indica un maggiore livello di benessere, per la PSS al crescere del punteggio totale si associano livelli più elevati di stress percepito.
È stato inoltre somministrato un **questionario socio-demografico e comportamentale,** autoprodotto, con domande relative ai caratteri sociodemografici, alla motivazione per la scelta universitaria effettuata, nonché a numerose abitudini oggetto di interesse, quali:

- abitudini alimentari;
- abitudini voluttuarie (tabagismo, uso di alcool, uso di droghe, uso improprio di farmaci);
- sessualità (orientamento sessuale, frequenza di rapporti, conoscenza di metodi anticoncezionali e di protezione dalle malattie sessualmente trasmissibili);
- interessi personali (attività sportive, musica, lettura, attività artistiche, hobbies ecc.);
- opinioni personali su aspetti clinici (medicina alternativa, vaccini, eutanasia, aborti).

Sui dati raccolti è in corso l'analisi statistica sia descrittiva che inferenziale. Per valutare la presenza di associazione tra variabili di tipo categorico è stato utilizzato il test chiquadro. Per le variabili continue, in funzione del tipo di distribuzione dei dati, sono eseguiti test parametrici o non. Il livello di significatività è stabilito con α=0.05. Tutte le variabili significative all'analisi univariata saranno inserite in un modello di regressione logistica multivariata con procedura ENTER e/o STEPWISE di selezione delle variabili, al fine di individuare le covariate indipendenti come eventuali predittori della presenza di livelli elevati di stress o di comportamenti a rischio, eventualmente con “aggiustamento” delle variabili confondenti.

Risultati preliminari

Complessivamente, si sono resi disponibili all’indagine 407 studenti (circa il 62.8% degli immatricolati), dei quali 150 di sesso maschile (38.9%) e 249 di sesso femminile (61.2%), mentre 8 non hanno precisato il loro sesso. Purtroppo, a causa di risposte incomplete, i dati relativi ai questionari psicologici fanno riferimento a un campione ridotto a 369 studenti di età media 19.72 anni (137 maschi, 225 femmine e 7 casi di sesso non dichiarato).
La consistenza interna delle scale – indice empirico dell’attendibilità – si è dimostrata decisamente adeguata, con coefficienti alpha di Cronbach di 0.79 per il questionario SF-12, di 0.93 per il PGWBI e di 0.88 per il PSS, in linea con i riscontri ottenuti negli studi di validazione degli strumenti e con i principali dati segnalati in letteratura (Grossi et al., 2002; Cohen & Williamson, 1988; Nielsen et al., 2016). Per quel che riguarda le sei sottoscale del PGWBI, le medie e le deviazioni standard non differiscono notevolmente da quelle ottenute sul campione normativo rappresentativo della popolazione italiana cui fa riferimento il manuale del questionario (Grossi et al., 2005).
Per gli indici SF12, invece, la media di PCS-12 è – come atteso in ragione della giovane età del campione – superiore a quella della popolazione generale utilizzata per la validazione italiana dello strumento (Apolone et al., 2005), con una minore variabilità e un punteggio medio prossimo al 75° percentile del campione di validazione. Viceversa, per l’indice di salute mentale MCS-12, il punteggio medio degli studenti è minore della media della popolazione generale (e inferiore al 25° percentile).
I punteggi del PSS non possono essere confrontati con dati normativi italiani, anche se l’impiego di questo strumento su campioni di studenti di medicina in altri paesi europei ha rilevato risultati simili a quelli qui riportati (Zvauya et al., 2017).
In linea con le evidenze del significativo impatto del genere sulla salute mentale (Afifi, 2007) e su specifici domini di funzionamento, come ad esempio la risposta allo stress (Matud, 2004), sono state analizzate le differenze tra maschi e femmine nel campione oggetto di studio, anche alla luce dei dati contrastanti riscontrati su studenti di medicina (Gutrhie et al., 1998; Tyssen et al., 2001; Sugiura et al., 2005). I punteggi

medi ottenuti ai questionari sono stati confrontati con test t di Student per campioni indipendenti. Complessivamente, si sono riscontrate differenze statisticamente significative per quel che riguarda il punteggio totale del PSS (maggiore stress percepito per le femmine), l'indice MCS-12 (minore salute mentale generale per le femmine), PGWBI totale e sottoscale Ansia, Depressione, Benessere Positivo, Autocontrollo e Vitalità (punteggi inferiori per le femmine in tutti casi, compatibili con un minor grado di benessere e maggiori problematiche di ansia e depressione).

Bibliografia

[1] Afifi, M. (2007). Gender differences in mental health. *Singapore Medical Journal*, *48*(5), 385.

[2] Apolone, G., Mosconi, P., Quattrociocchi, L., Gianicolo, E. A. L., Groth, N., & Ware Jr, J. E. (2005). Questionario sullo stato di salute SF-12. Versione italiana. *Milano: Istituto di Ricerche Farmacologiche Mario Negri*.

[3] Bore M., Kelly B., Nair B. 2016 - Potential predictors of psychological distress and well-being in medical students: a cross-sectional pilot study. *Adv Med Educ Pract*. 2;7:125-35.

[4] Cohen S., & Williamson G.M. (1988). Perceived Stress Scale in a probability sample of the United States. In Spacapan S., & Oskamp, S. (Eds.), *The social psychology of health*. Newbury Park: Sage.

[5] Dupuy H.J. (1984). The Psychological General Well-being (PGWB) Index. In N. K. Wenger, M. E. Mattson, C. D. Furburg, & J. Elinson (Eds.), *Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies* (pp. 170-183). New York: Le Jacq Publishing.

[6] Fiori Nastro P., Armando M., Righetti V., Saba R., Dario C., Carnevali R., Birchwood M., Girardi P. 2013 - Emerging mental disorders in a community sample of young adults in Italy: help-seeking in a Generalist Community Mental Health setting. *Riv Psichiatr*. 48 (1):60-6.

[7] Grossi, E., Mosconi, P., Groth, N., Niero, M., & Apolone, G. (2002). PGWBI: Questionario Psychological General Well-Being Index, versione italiana. *Milano: Istituto di Ricerche Farmacologiche Mario Negri*.

[8] Guthrie, E., Black, D., Bagalkote, H., Shaw, C., Campbell, M., & Creed, F. (1998). Psychological stress and burnout in medical students: a five-year prospective longitudinal study. *Journal of the Royal Society of Medicine*, *91*(5), 237-243.

[9] Kreiter CD., Axelson RD. 2013 - A perspective on medical school admission research and practice over the last 25 years. *Teach Learn Med*. 25 Suppl 1: S50-6.

[10] Lia C., Cavaggioni G. 2013 - The selection process in the Faculty of Medicine: the usefulness of a psychological and aptitude assessment. *Clin Ter*. 164 (1):39-42.

[11] Matud, M. P. (2004). Gender differences in stress and coping styles. *Personality and individual differences*, *37*(7), 1401-1415.

[12] Nielsen, M. G., Ørnbøl, E., Vestergaard, M., Bech, P., Larsen, F. B., Lasgaard, M., & Christensen, K. S. (2016). The construct validity of the Perceived Stress Scale. *Journal of Psychosomatic Research*, *84*, 22-30.

[13] Sugiura, G., Shinada, K., & Kawaguchi, Y. (2005). Psychological well-being and perceptions of stress amongst Japanese dental students. *European Journal of Dental Education*, *9*(1), 17-25.

[14] Tyssen, R., Vaglum, P., Grønvold, N. T., & Ekeberg, Ø. (2001). Factors in medical school that predict postgraduate mental health problems in need of treatment. A nationwide and longitudinal study. *Medical Education*, *35*(2), 110-120.

[15] Ware Jr, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical care*, 220-233.

[16] Zvauya, R., Oyebode, F., Day, E. J., Thomas, C. P., & Jones, L. A. (2017). A comparison of stress levels, coping styles and psychological morbidity between graduate-entry and traditional undergraduate medical students during the first 2 years at a UK medical school. *BMC Research Notes*, *10*(1), 93.

TREATMENT PATTERNS AND CLINICAL OUTCOMES IN PATIENTS RECEIVING PALBOCICLIB FOR HORMONE RECEPTOR–POSITIVE, HUMAN EPIDERMAL GROWTH FACTOR RECEPTOR 2–NEGATIVE ADVANCED OR METASTATIC BREAST CANCER: A REAL-WORLD EXPERIENCE IN A SINGLE ITALIAN INSTITUTION.

Piezzo Michela[1,2], Milanese Alberto[2], Paccosi Sara[2], Simenon Vittorio[2], Tanturli Michele[2], Vestri Annarita[2], Riemma Maria[1], Cianniello Daniela[1], Caputo Roberta[1], Di Lauro Vincenzo[1], De Laurentiis Michele[1]

[1] *Istituto Nazionale Tumori IRCCS Fondazione "Giovanni Pascale", Dipartimento di Oncologia Clinica e Sperimentale di Senologia*

[2] *Sapienza Università di Roma, Dipartimento di Sanità Pubblica e Malattie Infettive*

Introduction

Breast cancer (BC) is the most common cancer and the most common cause of cancer-related death in women and it represents about 30% of all diagnosed tumours [1]. Despite the important advances in terms of prevention, diagnosis and availability of new treatments, metastatic breast cancer (mBC) remains an incurable disease and its evolution depends on several factors, such as site and extension of metastasis, histopathological characteristics and molecular profile of tumour. Hormone receptor positive (HR+) BC subtypes represent approximately 60-75% of all BC and respond well to endocrine therapy (ET), in both adjuvant and metastatic setting. However, the majority of these patients show de-novo resistance (primary) or develop acquired resistance (secondary) to ET, which requires the administration of sequential endocrine-based therapy, both as monotherapy and in combination with targeted therapy before switching to chemotherapy-based regimens [2]. Preclinical and clinical research focused on identification of new treatment options able to prolong or restore the endocrine sensitivity, delaying the use of chemotherapy and improving the survival and the quality of life of these patients. In this scenario, deregulation of cell cycle represented one of the most attractive therapeutical targets, since dysregulation of cyclin D-CDK4/6-pRb pathway is frequent in HR-positive breast cancers and represents a key mediator of endocrine resistance [3, 4]. Targeting CDK4/6 resulted in an efficient inhibition of this pathway in HR-positive breast cancers [5, 6]. Palbociclib was the first inhibitor of CDK4/6 receiving an accelerated approval by Food and Drug Administration (FDA) in February 2015, in combination with letrozole as initial therapy for postmenopausal women with HR+ mBC, based on results from PALOMA-1 and PALOMA-2 trial [7, 8]. More recently, based on data from PALOMA-3 trial, the indication was expanded to the treatment of HR-positive/HER2-negative advanced or metastatic breast cancer, in combination with fulvestrant for both postmenopausal and pre-perimenopausal women with disease progression following endocrine therapy [9, 10, 11]. Palbociclib is approved in Italy since December 2017 for the treatment of women with HR+, HER2-negative (HER2-) mBC in combination with letrozole, as initial ET-based therapy, or fulvestrant, for women with disease progression after ET [12]. The inclusion of CDK4/6 inhibitors in combination with endocrine therapy in international treatment guidelines, both as initial therapy and after disease progression following endocrine therapy, represents the most relevant advance in the management of HR+/HER2- mBC over the last years [13].

Objectives

Our aim is to evaluate treatment patterns and clinical outcomes of patients who received palbociclib in combination with letrozole or fulvestrant in any line of therapy for treatment of HR+/HER2- mBC in a real-world setting.

Methods

All patients treated at National Cancer Institute of Naples with palbociclib plus letrozolo (P+L) or palbociclib plus fulvestrant (P+F) between December 2017 and January 2019 were included in the survey. Medical charts were retrospectively reviewed in order to collect and analyse the information needed, including demographic and clinical characteristics, time of initial breast cancer diagnosis, cancer related treatments received before palbociclib based therapy, palbociclib treatment patterns, response to therapy, time to response, duration of treatment, previous lines of endocrine therapy or chemotherapy, number and sites of metastasis, follow-up data after treatment interruption. Tumor response and progression of disease was assessed according to local practice and not using RECIST evaluation criteria normally used in clinical trials. The main outcome of interest was the progression free survival (PFS), defined as time (months) from first administration of P+L or P+F and permanent discontinuation of treatment due to progressive disease, intolerable toxicity, death from any cause or discontinuation due to any reason. Patients were censored if they were still on treatment at June 20, 2019.
Eligible patients were age 18 years or older, had histologically confirmed HR+/HER2- advanced or metastatic breast cancer, had follow-up data for 3 months or more since the intiation of P+F or 6 months or more from the initiation of P+L.
All analyses were descriptive and exploratory in nature, and were conducted by R-3.6.0 statistical software. Survival outcomes were stratified by various levels, such as prior treatment for metastatic disease, prior response to endocrine therapy, sites of metastases (visceral disease, bone only disease, visceral plus bone disease), stage at diagnosis, dose reduction while on treatment with palbociclib, menopausal status. Results will be presented as frequencies, proportions, means and all other summary measures as appropriate. Event time measures were examined using the Kaplan-Meyer method. Finally the real-life survival curves for PFS were compared with curves extracted from published data, as described by Guyot et al in 2012 [14].

Results

A total of 192 patients were included in this analysis, of which 100 was treated with P+F and 92 was treated with P+L. Median follow-up was 9.7 months in P+F group and 9.2 in P+L group. Patients had a median age of 58 years and the 64% of them was in post-menopausal status. Overall median PFS after palbociclib based treatment initiation was 14.8 months (CI 95% 13.6, NA), while patients treated with P+F has a median PFS of 12.1 months (CI 95% 9.25 NA) and patients treated with P+L has not yet reached the median PFS. Event rates during the follow-up period were also calculated at 6 and 12 months after the initial administration of palbociclib based therapy, the percentage of patients alive was 80% and 34%, respectively. The comparison of PSF results were finally compared with results from PALOMA-2 and PALOMA-3 study. The PALOMA-2 study showed a median PFS of 25.3 months in palbociclib plus letrozole arm, in accordance with our data (median PFS not reached in P+L group), while the PALOMA-3 study showed a median PFS of 9.5 months, 3 months lower than PFS obtained from our data (12.1 months in P+F group).
Further results, including tables and figures, will be presented in September, in order to have longer follow-up data available.

Conclusions

Although palbociclib has been widely studied in the clinical trial setting, evidence for its use in real-world patients with mBC is scarce. In addition, it was studied as first or second line treatment for mBC, while in clinical practice more than 50% of patients heavily pretreated. Our data provide additional evidence to support the hypothesis that addition of CDK 4/6 inhibitor to standard endocrine therapy is the best available option to treat patients with HR+ mBC, even after failure of multiple lines of chemotherapy and endocrine therapy.

References

[1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018; 68: 7-30.

[2] Cardoso F, Costa A, Senkus E, et al. 3rd ESO–ESMO I nternational consensus guidelines for Advanced Breast Cancer (ABC 3). Breast. 2017 Feb; 31:244-259.

[3] Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast cancer. Nature 2012; 490:61-70

[4] Zardavas D, Baselga J, Piccart M. Emerging targeted agents in metastatic breast cancer. Nat Rev Clin Oncol 2013; 10:191-210

[5] O'Leary B, Finn RS, Turner NC. Treating cancer with selective CDK4/6 inhibitors. Nat Rev Clin Oncol 2016; 13: 417-30

[6] Finn RS, Dering J, Conklin D, et al. PD 0332991, a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines in vitro. Breast Cancer Res 2009; 11: R77

[7] Finn RS, Crown JP, Lang I, et al. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of estrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. Lancet Oncol 2015; 16:25e35

[8] Finn RS, Martin M, Rugo HS, et al. Palbociclib and Letrozole in Advanced Breast Cancer. N Engl J Med. 2016 Nov 17;375(20):1925-1936

[9] Turner NC, Ro J, André F, et al. Palbociclib in Hormone-Receptor-Positive Advanced Breast Cancer. N Engl J Med. 2015 Jun 1.

[10] Verma S, Bartlett CH, Schnell P, et al. Palbociclib in Combination With Fulvestrant in Women With Hormone Receptor-Positive/HER2-Negative Advanced Metastatic Breast Cancer: Detailed Safety Analysis From a Multicenter, Randomized, Placebo-Controlled, Phase III Study (PALOMA-3). Oncologist. 2016 Oct;21(10):1165-1175.

[11] Turner NC, Slamon DJ, Ro J, et al. Overall Survival with Palbociclib and Fulvestrant in Advanced Breast Cancer. N Engl J Med. 2018 Nov 15;379(20):1926-1936.

[12] GU Serie Generale n.298 del 22-12-2017

[13] Rugo HS, Rumble RB, Macrae E, et al. Endocrine Therapy for Hormone Receptor-Positive Metastatic Breast Cancer: American Society of Clinical Oncology Guideline. J Clin Oncol. 2016 Sep 1;34(25):3069-103.

[14] Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol. 2012 Feb 1;12:9.

PREVALENZA DI ASMA E BPCO IN DUE COMUNI DELLA LOMELLINA: UNO STUDIO TRASVERSALE

Puci Mariangela Valentina[1], Gnesi Marco[1], Ferraro Ottavia Eleonora[1], Gigli Berzolari Francesca[1], Monti Maria Cristina[1], Borrelli Paola[1], Oddone Enrico[2], Zottarelli Fabio[1], Montomoli Cristina[1], Villani Simona[1]

[1]Dipartimento di Sanità Pubblica, Medicina Sperimentale e Forense, U.O. di Biostatistica e Epidemiologia Clinica, Università degli Studi di Pavia

[2]Dipartimento di Sanità Pubblica, Medicina Sperimentale e Forense, sezione di Medicina del Lavoro "Salvatore Maugeri, Università degli Studi di Pavia.

Introduzione

Tra le malattie respiratorie l'asma rappresenta una delle patologie più diffuse al mondo con importanti ripercussioni non soltanto sulla salute dell'individuo ma anche sul sistema economico sanitario.
Nel mondo sono circa 300 milioni le persone che soffrono di asma [1], nonostante la prevalenza differisca da nazione a nazione, si stima che circa il 5% degli adulti e il 10% dei bambini ne siano affetti [2]. In Italia, la prevalenza di asma è inferiore rispetto a quella di altri paesi [3], tuttavia secondo i dati raccolti dalla Medicina generale, tale prevalenza nella popolazione di età >15 anni è pari al 6.1% (6.6% nelle donne, 5.5% negli uomini) [4] e al 6.6% nella popolazione di 20-44 anni [5].
L'asma è una patologia complessa, nella quale giocano un ruolo importante sia fattori genetici che ambientali. In merito a questo ultimo aspetto, in concomitanza con l'incremento della prevalenza dei disturbi respiratori (BPCO, rinite e sintomi simil-asmatici), diversi studi hanno focalizzato l'attenzione sulla relazione tra esposizione a inquinanti ambientali e disturbi respiratori in particolare nei soggetti residenti in aree industriali [6]. I dati qui presentati sono il risultato preliminare di uno studio attualmente in corso che mira a studiare lo stato di salute nella popolazione residente in due comuni della Lomellina sedi di una raffineria.

Obiettivi

Stimare la prevalenza di asma bronchiale e di BPCO nella popolazione adulta (20-64 anni) residente nei comuni di Sannazzaro de Burgondi e Ferrera Erbognone in provincia di Pavia.

Metodi

Studio trasversale condotto sui soggetti adulti (20-64 anni) residenti a Sannazzaro e Ferrera nell'anno 2019.
I dati sono stati raccolti tramite un questionario che è un adattamento di quello già utilizzato sia in uno studio internazionale che in uno nazionale (European Community Respiratory Health Survey, www.ecrhs.org e Gene Environment Interactions in Respiratory Diseases, www.geird.org). Il questionario raccoglie diverse informazioni quali: caratteristiche socio-demografiche (sesso, età, scolarità e professione); stile di vita (abitudine al fumo) e di natura clinica (asma, sintomi simil-asmatici, rinite allergica, bronchite cronica e BPCO).
La presenza di asma bronchiale e BPCO, principali end-points, è stata determinata a partire dalle risposte fornite al questionario di screening. È stato considerato come affetto da probabile asma il soggetto che ha dichiarato di essere asmatico, di aver avuto un attacco d'asma nell'ultimo anno e di essere in terapia farmacologica (assunzione di farmaci per la cura dell'asma al momento dell'indagine) [7].
È stato identificato come affetto da probabile BPCO il soggetto che ha dichiarato di avere tosse o di espellere catarro bronchiale per la maggior parte dell'anno, per almeno 3 mesi ogni anno.
È stato utilizzato il test Chi-quadro per valutare l'associazione tra variabili categoriche e il test t di Student per valutare la differenza tra variabili quantitative tra due gruppi indipendenti. La presenza di fattori associati

agli outcome principali (asma e BPCO) è stata studiata tramite l'implementazione di un modello logistico multivariabile. Poiché la raccolta dati è stata articolata su 3 diverse ondate di contatto, nel modello multivariabile sono state inserite come covariate l'ondata di risposta e la stagione. Il livello di significatività considerato è stato $p<0.05$. Le analisi sono state condotte usando STATA13.

Risultati

I soggetti che hanno risposto al questionario sono 1108 di cui il 75.2% residenti a Sannazzaro de' Burgondi e il 24.8% residenti a Ferrera Erbognone. Il 53.1% dei soggetti è di genere femminile con un'età media di 48.0±12.3 anni. Il 25.8% dei rispondenti era un fumatore al momento della compilazione del questionario, mentre il 25.7% un ex fumatore che aveva fumato per almeno un anno nel corso della vita, la restante maggioranza ha dichiarato di non aver mai fumato. Tra i partecipanti all'indagine, la prevalenza di asma nel campione è pari all'11.0% (119 soggetti), senza differenze significative tra maschi e femmine: i maschi asmatici sono il 10.6% contro l'11.3% delle femmine (χ^2=0.134, p=0.715). In merito alla BPCO la prevalenza osservata è pari al 25.1% (273 soggetti), ma a differenza dell'asma, tale prevalenza è significativamente superiore nei maschi rispetto alle femmine (29.0% vs 21.6%, χ^2=7.772, p=0.005). Infine per quanto concerne i sintomi simil-asmatici, il 19.13% (212) dei partecipanti ha dichiarato di averne accusato almeno uno nell'ultimo anno. Anche in quest'ultimo caso le differenze di genere non sono risultate significative: tra le donne il 10.6% contro il 18.2% degli uomini (χ2=0.990, p=0.320). La Tabella 1 mostra i risultati ottenuti dai modelli di regressione logistica. In riferimento all'asma si osserva che, a parità di tutte le altre variabili, l'unico fattore significativamente associato all'essere un probabile caso di asma è l'avere fumato per almeno un anno nel corso della vita. L'aumento dell'età, seppur lievemente, sembra invece diminuire significativamente la prevalenza di asma (circa 2.2% per ogni incremento di età di un anno). In merito all'ondata di contatto in cui il partecipante ha compilato il questionario, sembra che tra coloro che hanno risposto nella seconda ondata vi fosse una maggiore probabilità di essere asmatici. Anche per la BPCO l'unico fattore significativamente associato è l'essere fumatore, con una prevalenza di BPCO maggiore di due volte e mezzo nei fumatori rispetto a chi non ha mai fumato. Nel caso della BPCO si riscontra anche un effetto protettivo del genere femminile rispetto a quello maschile, con una prevalenza di BPCO ridotta di circa il 25% tra le donne ma tale effetto è al limite della significatività statistica.

Tabella 1. *Fattori associati all'asma e alla BPCO - modello di regressione logistica multivariabile*

Asma		
Variabili	**OR (95% CI)**	**p value**
Età (+1 anno)	0.976 (0.961-0.992)	0.003
Sesso (F vs. M)	1.171 (0.781-1.754)	0.445
Comune (FE vs. SdB)	1.150 (0.696-1.898)	0.586
Fumo (vs. Non fumatore)		
Ex fumatore	1.229 (0.721-2.010)	0.448
Fumatore	2.028 (1.275-3.226)	0.003
Stagione (vs. Primavera)		
Estate	0.858 (0.439-1.674)	0.653
Autunno	0.919 (0.369-2.291)	0.857
Inverno	0.713 (0.203-2.503)	0.598
Ondata di contatto (vs. 1^)		
2^	2.382 (1.054-5.385)	0.037
3^	1.127 (0.595-2.137)	0.713

BPCO		
Variabili	**OR (95% CI)**	**p value**
Età (+1 anno)	0.996 (0.984-1.007)	0.450
Sesso (F vs. M)	0.743 (0.555-0.996)	0.047
Comune (FE vs. SdB)	1.342 (0.931-1.935)	0.115
Fumo (vs. Non fumatore)		
Ex fumatore	1.175 (0.804-1.716)	0.404
Fumatore	2.542 (1.806-3.578)	<0.001
Stagione (vs. Primavera)		
Estate	1.073 (0.662-1.740)	0.775
Autunno	0.733 (0.366-1.467)	0.380
Inverno	1.101 (0.489-2.478)	0.816
Ondata di contatto (vs. 1^)		
2^	1.350 (0.669-2.722)	0.402
3^	1.082 (0.690-1.698)	0.731

Conclusioni

La prevalenza di asma sembra superiore rispetto a quella trovata da altri studi condotti in modo simile in Italia [8] supportando però quelle evidenze che la danno nuovamente in aumento [9].
I nostri risultati confermano quanto già riscontrato in letteratura: il fumo rappresenta il fattore di rischio più importante nello sviluppo di BPCO [10], sebbene la presenza di BPCO sia stata definita sulla base delle risposte date ad un questionario auto compilato. Per ridurre la prevalenza di tali patologie è importante agire su diversi fronti, non soltanto terapeutici ma anche di prevenzione.
Questi risultati suggeriscono l'importanza di attuare interventi mirati a ridurre i principali fattori di rischio. Pertanto, l'implementazione di programmi di promozione della salute rappresenta la strategia cardine su cui far leva per promuovere stili di vita corretti sia per il singolo che per la collettività. Attualmente è in corso l'approfondimento della relazione con l'esposizione ambientale.

Bibliografia

[1] Bahadori, K., Doyle-Waters, M.M., Marra, C. et al. (2009). Economic burden of asthma: a systematic review. BMC Pulmonary Medicine, 19, 9-24

[2] LINE GUIDA GINA ITALIANE 2016-2017 reperibile su: http://ginasma.it/wp-content/uploads/materiali/2017/GINA_adattamento_ita_2017_doc.pdf

[3] Sestini, P., De Sario, M., Bugiani, M. et al; il Gruppo Collaborativo SIDRIA-2. (2005). Frequency of asthma and allergies in Italian children and adolescents: results from SIDRIA-2. Epidemiologia e Prevenzione, 29(2) suppl:24-31

[4] Cazzola, M., Puxeddu, E., Bettoncelli, G. et al. (2011). The prevalence of asthma and COPD in Italy: a practice-based study. Respiratory Medicine, 105:386-391

[5] De Marco, R., Cappa, V., Accordini, S. et al (2012); the GEIRD study group. Trends in the prevalence of asthma and allergic rhinitis in Italy between 1991 and 2010. European Respiratory Journal, 39:883-892

[8] Verlato G., Corsico A., Villani S. et al. Is the prevalence of adult asthma and allergic rhinitis still increasing? Results of an Italian study. J Allergy Clin Immunol 2003; 111:1232-8.

[9] De Marco R1, Cappa V, Accordini S et al. Trends in the prevalence of asthma and allergic rhinitis in Italy between 1991 and 2010. Eur Respir J. 2012 Apr;39(4).

[10] De Marco R, Accordini S, Marcon A, et al. Risk factors for chronic obstructive pulmonary disease in a European cohort of young adults. Am J Respir Crit Care Med. 2011 Apr 1;183(7):891-7.



sismec

X CONGRESSO NAZIONALE SISMEC

Società Italiana di Statistica Medica ed Epidemiologia Clinica

Nuovi disegni nella ricerca clinica: la sfida della complessità tra etica e salute