

First BIAS National Congress
Florence 29/30 April 2009



Case Study: **A comparison between LOCF and ML**

Giada Rizzi
Senior Statistician
Chiesi Farmaceutici



Outline

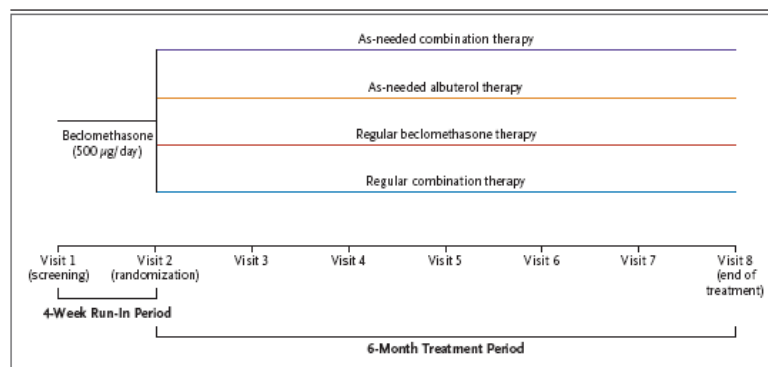
- Overview of the study
- Submission for publication
- General considerations on LOCF and ML
- Presentation of results using LOCF and ML
- Conclusions – Lessons Learnt



The study

- A multinational, multicentre, double-blind, randomised, four-arm parallel group study of 24-week duration on the therapeutic efficacy and safety of beclomethasone (BDP) dipropionate 250 mg/albuterol 100 mg in the treatment of patients with mild persistent asthma.
- Protocol dated 2001

The study



Study Objectives

The study was planned to assess:

- the equivalence of "as needed" BDP/albuterol (combo) versus:
 - (i) regular BDP b.i.d. plus "as needed" albuterol
 - (ii) regular BDP/albuterol b.i.d. plus "as needed" albuterol
- and
- the superiority of "as needed" BDP/albuterol versus "as needed" albuterol.

The primary efficacy variable was pre-dose morning Peak Expiratory Flow (PEF) at study end (mean over the last two weeks). PEF was recorded daily by patients at home.



Handling of missing data

"Missing data of patients who discontinue the study will be replaced with the LOCF approach" (protocol - 2001)

"Regarding the primary variable, a minimum of ten measurements are required in each two-week period. In case of less than ten measurements available in the two-week period, the variable will be considered as missing for that two-week period. ...missing data will be replaced according to the LOCF method, which will be applied only to post-baseline data. If a subject has only baseline value, no replacement will be done and the subject will be excluded from the LOCF analysis for that particular variable." (SAP - 2004).



Results

The administration of BDP/albuterol on an "as needed" basis was equivalent to BDP 500 µg/day plus "as needed" albuterol and to BDP/albuterol 200/500 µg/day plus "as needed" albuterol and was superior to "as needed" albuterol alone.

Submission for publication First round of review

The paper was submitted for publication in 2006.

***"The use of LOCF is not acceptable.* It is well known to *inflate the sample size and spuriously increase the precision* of estimates of treatment effect. Instead you should use statistical methods for maximum likelihood analysis of incomplete data as implemented in SAS PROC MIXED. *The ML method... produces valid estimates of the standard errors of the differences between treatment groups when the MAR assumption holds*".**

Our reply

As per reviewers' comments, we performed an additional analysis of the primary variable, using the ML estimation analysis.

We included the ML analysis in a Supplementary Appendix along with a brief explanation of the reason why we added this analysis.

Second round of review

"It seems to me that if the LOCF approach is felt to lead to spurious results, then the ML results should be the main ones..."

"I feel strongly that LOCF is not the most rigorous way to analyze your data... ***By using LOCF, you impute data from earlier periods and then act as if these were actual observations. This can introduce bias but, more importantly, inflates the apparent sample size and thus spuriously increases precision...*** the primary analysis should use ML. ***I understand that LOCF analysis is popular in drug trials but it is not state of the art...*** it is not satisfactory to present only the LOCF analysis."

Third round of review

"It is ***more important that the analysis be correct than that it have been decided on a priori...***"

"Certainly one should not shift away from a priori hypothesis... I am not convinced of the need to place the central focus on a ***suboptimal analysis [LOCF]*** simply because...originally planned...My suggestion would be to put the ML tables and results in the main manuscript."

"I understand your concerns about the specification of the LOCF analysis in the protocol, as well as the support for this approach by regulatory bodies. Thus, I recommend... inclusion of ...both the LOCF and ML versions of the primary analyses."

Our last reply

In the end we included both the analyses using LOCF and ML in the main body of the paper.

"Although as per our original statistical plan we used the LOCF method to handle missing data, in the final analysis, the ML method of analysis was also used to test the hypothesis of superiority or equivalence with regard to primary ...efficacy variable."

Publication



"Rescue Use of Beclomethasone and Albuterol
in a Single inhaler for Mild Asthma"

N Engl J Med 2007; 356: 2040-52



The NEW ENGLAND
JOURNAL of MEDICINE



Amount of missing data

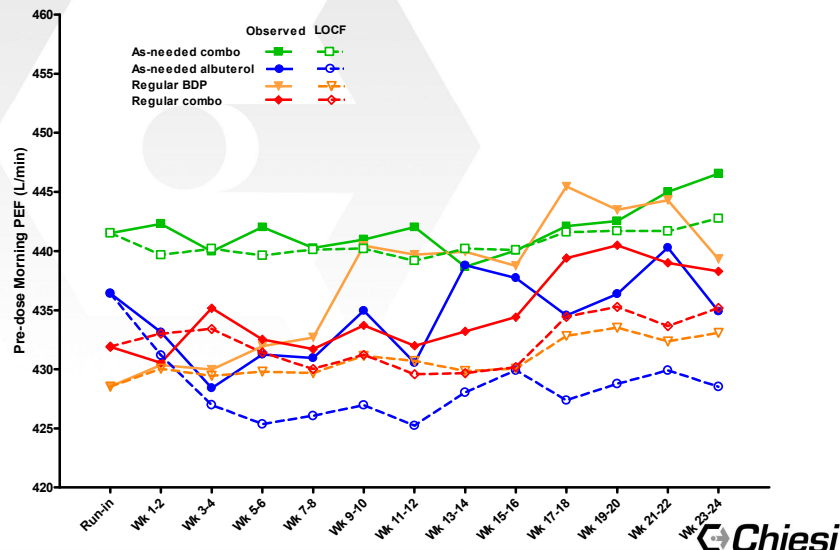
- Randomised N=466
- Completers N=393 (73 dropouts)

For the analysis:

- Modified ITT population (all randomised who took at least one dose of study drug and with at least a post-baseline assessment) N=455
- Primary efficacy variable: baseline N=448
study end N=377



Means of observed data and LOCF



First BIAS National Congress - 29/30 Apr 2009

15

Why did we use LOCF/why LOCF is popular?

- Very easy to implement and we did not have any experience with other methods (actually this is not a good reason).
- Points to Consider on Missing Data; Nov 2001 ("one widely used method is LOCF"). Now, "Recommendation for the revision of the points to consider on missing data"; Dec 2007 ("The misconception that LOCF represents a necessary and sufficient approach to missing data should be dispelled").
- Up to that time, we had never faced any problems with Regulatory Authorities using LOCF.

Chiesi

First BIAS National Congress - 29/30 Apr 2009

16

Were the reviewers right? Criticism to LOCF

There is an extensive literature on weakness of LOCF and other simple methods to deal with missing data:

- LOCF is not based on statistical principles. "It imputed a single value...the subsequent analysis gives the imputed responses the same status as actual observed responses" (5). As a consequence, LOCF artificially increases information and precision (7).
- As a matter of fact it is an analysis of last observed value. Thus, we should consider the scientific relevance of the analysis using last observed value (in many settings, this analysis is unlikely to be meaningful) (4, 5).

Were the reviewers right? Criticism to LOCF

- LOCF is based on unclear, strong assumptions unlikely to be plausible.

It goes far beyond MAR assumptions (i.e. conditional on the independent variables in the analysis model and the observed measurements of the variable being analysed - patient history, the probability of missingness does not depend on the unobserved measurements of the variable being analysed).

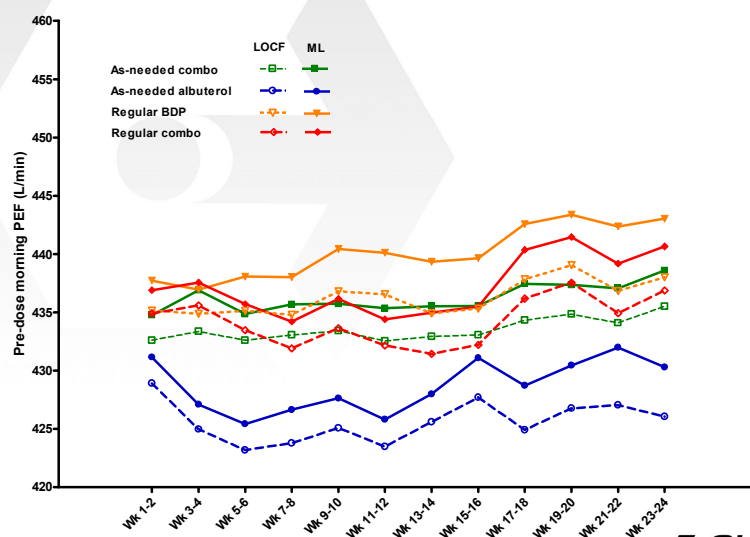
It "is not valid even under the strong MCAR condition" (4, 5, 7).

- LOCF can be either conservative or liberal depending on data (7) - "LOCF typically produces bias of which the direction and magnitude depend on the true but unknown treatment effect" (4).

Were the reviewers right? ML-based methods and MAR

- ML methods are valid under MAR assumptions (4, 5).
- MAR is often plausible in longitudinal confirmatory clinical studies (6).
- Even if it took quite a long time, there is now general agreement that analyses of longitudinal trials should move away from simple methods (as LOCF) in favour of more principled methods (as ML-based methods) (6).

Adjusted means with LOCF and ML



Results with LOCF and ML

Adjusted means (SEs) for morning PEF (L/min) at the end of study – Modified ITT population

	LOCF	ML
As-needed combo	435.5 (3.12)	438.6 (2.80)
As-needed albuterol	426.0 (3.21)	430.3 (2.84)
Regular BDP	438.0 (3.41)	442.4 (3.01)
Regular combo	436.9 (3.32)	440.7 (2.93)



Results with LOCF and ML

**Morning PEF (L/min) at the end of study – Modified ITT population
Comparisons between groups**

	LOCF	ML
As-needed combo vs. as-needed albuterol	9.47 (0.83 to 18.11) p=0.032	8.31 (0.58 to 16.04) p=0.035
As-needed combo vs. regular BDP	-2.49 (-11.40 to 6.42)	-4.44 (-12.39 to 3.52)
As-needed combo vs. regular combo	-1.36 (-10.13 to 7.42)	-2.05 (-9.89 to 5.78)

Data are differences between adjusted means (95% CI)



Comparison of results

Results from LOCF and ML analyses were very similar (all objectives were still met). This does not mean that LOCF is a proper method for dealing with missing data.

As a reviewer commented: "Because an appreciable fraction of the participants were not seen at the six-month visit, this agreement between LOCF and ML] was not a sure thing. It is reassuring and adds strength to the findings."

A possible reason for the agreement could be in the study design which made the assumption of constant patient profile less implausible compared to other settings.

We were lucky!

Conclusions – Lessons learnt

LOCF is not a good/sensible method for dealing with missing data.

ML-based methods produce valid estimation under MAR assumptions and should be used for the primary analysis (longitudinal confirmatory trials).

Sensitivity analyses should be performed focusing on different assumptions (MNAR) rather than on different methods (5).

References

- 1) Papi et al. "Rescue Use of Beclomethasone and Albuterol in a Single inhaler for Mild Asthma" N Engl J Med 2007; 356: 2040-52
- 2) Committee for Proprietary Medicinal Products: "Points to Consider on Missing data"; Nov 2001
- 3) Committee for Proprietary Medicinal Products: "Recommendation for revision on PtC on Missing data"; Dec 2007
- 4) Molenberghs et al. "Analyzing incomplete longitudinal clinical trial data"; Biostatistics (2004), 5, 3, pp.445-464
- 5) J. Carpenter, M. Kenward "Missing data in randomised controlled trials – a practical guide"
- 6) Mallinckrodt, Kenward "Conceptual Considerations Regarding Choice of Endpoints, Hypotheses, and Analyses In Longitudinal Clinical Trials" in press
- 7) Kenward, Molenberghs "LOCF: A Crystal Ball?" in press



THANKS

Any questions?

Giada Rizzi
g.rizzi@chiesigroup.com

