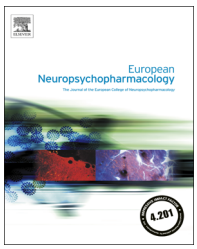
European Neuropsychopharmacology (2014) **24**, 357–368

www.elsevier.com/locate/euroneuro



Combining efficacy and completion rates with no data imputation: A composite approach with greater sensitivity for the statistical evaluation of active comparisons in antipsychotic trials

Jonathan Rabinowitz[a,*], Nomi Werbeloff[a], Ivo Caers[b], Francine S. Mandel[c], Judith Jaeger[d], Virginia Stauffer[e], François Menard[f], Bruce J. Kinon[e], Shitij Kapur[g]

[a]*Bar Ilan University, Ramat Gan, Israel*
[b]*Janssen Research & Development, Beerse, Belgium*
[c]*Pfizer, 235 East 42nd Street, New York, NY 10017, USA*
[d]*Astra-Zenecca, Wilmington, DE, USA*
[e]*Eli Lilly, 893 South Delaware Street, Indianapolis, IN 46285, USA*
[f]*Lundbeck SAS, 37-45, Quai du Président Roosevelt, 92445 Paris, France*
[g]*Institute of Psychiatry, Kings College London, De Crespigny Park, London SE5 8AF, UK*

Received 8 October 2013; accepted 17 November 2013

KEYWORDS
Clinical trials;
Efficacy;
Design;
Methodology

Abstract
Outcomes in RCT's of antipsychotic medications are often examined using last observation carried forward (LOCF) and mixed effect models (MMRM), these ignore meaning of non-completion and thus rely on questionable assumptions. We tested an approach that combines into a single statistic, the drug effect in those who *complete* trial and *proportion* of patients in each treatment group who complete trial. This approach offers a conceptually and clinically meaningful endpoint. Composite approach was compared to LOCF (ANCOVA) and MMRM in 59 industry sponsored RCT's. For within study comparisons we computed effect size (*z*-score) and *p* values for (a) rates of completion, (b) symptom change for complete cases, which were combined into composite statistic, and (c) symptom change for all cases using last observation forward (LOCF). In the 30 active comparator studies, composite approach detected larger differences in effect size than LOCF (ES=.05) and MMRM (ES=.076). In 10 of the 49 *comparisons* composite lead to significant differences ($p \leq .05$) where LOCF and MMRM did not. In 3 comparisons LOCF was significant, in 2 MMRM lead to significant differences whereas composite did not. In placebo controlled trials, there was no meaningful difference in effect

*Corresponding author. Tel.: +972 54 464 3889.
E-mail address: jonathan.rabinowitz@biu.ac.il (J. Rabinowitz).

0924-977X/$ - see front matter
http://dx.doi.org/10.1016/j.euroneuro.2013.11.010

size between composite and LOCF and MMRM when comparing placebo to active treatment, however composite detected greater differences than other approaches when comparing between active treatments. Composite was more sensitive to effects of experimental treatment vs. active controls (but not placebo) than LOCF and MMRM thereby increasing study power while answering a more relevant question.

1. Introduction

Dropout or discontinuation is a major cause of missing data in clinical trials generally, and trials of antipsychotic medications, in particular. It is an important outcome as it may reflect a lack of drug tolerability, lack of efficacy, adverse effects or lack of compliance. It creates uncertainty in interpreting study results. It is not uncommon for dropout rates of antipsychotic medication trials to exceed 50% (Martin et al., 2006; Rabinowitz et al., 2009; Wahlbeck et al., 2001). In our NewMeds repository completion rate in the 30 active controlled studies (n=12,846, treatment duration 4–104 weeks) was only 48.7% and in the 29 placebo controlled trials (n=9174, treatment duration 4–52 weeks) the completion rate was only 53.9% (after removing two trials longer than 8 weeks for comparability). Recently, there has been increased recognition of the problem of missing data in clinical trials by regulatory authorities (O'Neill and Temple, 2012) and the limitations of conventional ways of accounting for it.

Dropout leads to missing data that varies as to the extent to which it affects modeling and analysis. The literature distinguishes between three mechanisms of missing data (Little and Rubin, 1987). First, *missing completely at random* (MCAR); this refers to a situation where missingness does not depend on either the observed or unobserved data. A possible example is data lost because some patient records were destroyed in a flood. MCAR can be handled in the analysis using standard approaches such as mixed models or LOCF. Nevertheless MCAR leads to loss of power due to diminished sample size. Second, *missing at random* (MAR) occurs if the missing data depends on variables that are observed during the trial and not on unobserved data. The data is MCAR if, for example, the probability of dropout is unrelated to any of the other variables of potential interest and relevance (e.g. the rate of dropout is unrelated to starting severity, illness type, age, gender etc.). If the probability of dropout varies by a given variable, say gender of the subject, but since gender is known one can examine the differences within a gender and hence control for it. If after controlling for gender (i.e. within men and women separately) the dropouts are unrelated to any further variable the data are MAR. Third, *missing not at random* (MNAR) occurs if the missingness depends on unobserved data. An example could be a patient who was improving and then was lost to follow-up because of a relapse after the last observed visit and was admitted to a different hospital. In this case the observed data could not predict the missing data. The unobserved data contained information not foreseen by the observed data (Mallinckrodt et al., 2003). MNAR cannot be corrected without explicitly specifying a model for the missing data mechanism, which by definition, cannot be tested.

MCAR and MAR are termed *ignorable non-response* since the first requires no special attention when analyzing the data and the second can be controlled for in the analysis. MNAR is termed *non-ignorable non-response* since it cannot be ignored. It cannot be ignored as it is informative, for example dropout due to lack of efficacy. In-fact MAR and MNAR are also sometimes referred to as "informative" as the data that is missing is informative as it relates to study variables.

Missing data in clinical trials of antipsychotic medication because of dropout are problematic since they are rarely MCAR and it is generally difficult to determine if they are MAR or MNAR. Historically, a standard approach used in clinical trials is the last observation carried forward (LOCF). LOCF uses the last completed observation while on treatment to estimate a hypothetical last visit value. This is problematic since it assumes that the data are MCAR and that symptoms would have remained stable and constant, with no within-subject variation after dropout. This leads to inflation of Type I error rates, since the estimated standard error of test statistics is biased downward until the end of the trial. Some recent trials (Duan et al., 2006; Lieberman et al., 2005; Lieberman et al., 2003) have applied a mixed-effects model (Mallinckrodt et al., 2003) which is thought to provide more accurate estimates of treatment than LOCF. LOCF analysis can lead to substantial biases in estimating treatment effects and can greatly inflate Type I error rates of the statistical tests, whereas MMRM analysis on the available data leads to estimates with smaller bias, and controls Type I error rates at a nominal level in the presence of missing completely at random (MCAR) or missing at random (MAR) (Siddiqui et al., 2009). These estimates are based on data available at each given time point. Mixed-effect models work if data is MCAR or MAR, however if the data is MNAR then inferences based on these methods will probably not be valid.

The above highlights that when using standard approaches, the mechanism of dealing with missing data is of critical importance. However, an alternative and newer approach has been proposed to address the dropout problem (Shih and Quan, 1997) which can be applied *regardless of the missing data mechanism*. This approach termed the composite approach was developed by Shih and Quan (Shih and Quan, 1997). It combines two hypotheses stating that more patients will complete the trial on the better drug and that patients who complete the trial will improve more on the better drug. Accordingly, this is termed the composite approach. Specifically, this approach (Shih and Quan, 1997) combines the p value of the difference in completion rates between drug treatments and the p value obtained in comparing the difference in treatment outcomes of complete cases. The approach gives a single p value that reflects both outcomes. If the result is statistically significant it means that the groups differ on the combined hypothesis. Thus when the null hypothesis is rejected, the conclusion is that the chance of