Patient Preference and Adherence

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

A closer look at the baseline-observation-carried-forward (BOCF)

This article was published in the following Dove Press journal:
Patient Preference and Adherence
16 January 2010
Number of times this article has been viewed

Hong Liu-Seifert
Shuyu Zhang
Deborah D'Souza
Vladimir Skljarevski

Lilly Research Laboratories, Indianapolis, IN, USA

Purpose: The baseline-observation-carried-forward (BOCF) approach is one method to handle missing data from early treatment discontinuation. We examined modifications of this approach, taking into consideration treatment-related and nontreatment-related reasons for discontinuation.

Methods: Two duloxetine chronic pain trials (placebo-controlled) were used to examine the impact of different analytical methods on study outcome. Reasons for discontinuation were categorized as treatment-related and nontreatment-related. Missing data in the primary efficacy outcome were handled using five statistical methods: mixed-model repeated measures (MMRM), last-observation-carried-forward (LOCF), BOCF, modified BOCF (mBOCF, discontinuation due to treatment-related reasons, ie, adverse events [AEs] or lack of efficacy), and aeBOCF (discontinuation due to AEs only).

Results: Duloxetine was superior to placebo on mean change from baseline in Brief Pain Inventory average pain rating, using MMRM (study 1, $P = 0.004$; study 2, $P < 0.001$), LOCF (study 1, $P = 0.019$; study 2, $P < 0.001$), BOCF (study 1, $P = 0.019$; study 2, $P = 0.013$), and mBOCF (study 1, $P = 0.041$; study 2, $P = 0.005$). Using aeBOCF, duloxetine was superior to placebo in study 2 ($P = 0.005$) and numerically better in study 1 ($P = 0.075$).

Conclusion: Due to the different assumptions made by various methods regarding accounting for missing data, the analytical methods chosen may influence the interpretation of study results. Consideration should be given to the effect of actual treatment outcomes from patients. Employing different statistical approaches such as sensitivity analyses may help to assess the robustness of the study results and provide a more accurate reflection of the treatment outcome.

Keywords: discontinuation, treatment-related, pain, statistical

Introduction

A clinical trial is conducted with the intent to treat and evaluate all patients enrolled in the study. One of the most common problems encountered during clinical trials is the evaluation of data from patients who are unable to complete the full schedule of the clinical trial or otherwise drop out of the study. There are many potential reasons why patients may drop out of clinical trials, including poor compliance, inconvenience, schedule conflicts, protocol violations, adverse events, death, early recovery, or other nontreatment-related causes which are often out of the investigator's control. Nonetheless, any reason for dropping out from the trial results in data lost from these patients (drop-outs) and this could lead to a potential bias in the final study outcome. It is therefore important to accommodate drop-outs to appropriately analyze the outcome of the clinical trial.

Correspondence: Hong Liu-Seifert
Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285 USA
Tel +1 317-433-0662
Fax +1 317-277-7839
Email liu-seifert_hong@lilly.com

Various statistical methods have been used to handle missing data in clinical trials. One of these methods includes the last-observation-carried-forward (LOCF) approach, in which the last observed nonmissing value is used in place of the missing endpoint. However, this method can have drawbacks because it provides biased estimates of treatment effects and biased tests of the null hypothesis associated with no treatment effect.[1] In addition, if there is equal drop-out in both the active and placebo arms, LOCF is liable to underestimate the treatment effect, and in cases of unequal drop-outs, the bias can be much larger in either direction.[2] In addition, with the LOCF method, one has to assume that subjects' measurement remains at the same level from the moment of dropout onward to the end of the trial. Another method, the mixed-model repeated-measures (MMRM) analyses, utilizes likelihood-based estimation, subject-specific effects and correlations between the repeated measurements.[3] This method is considered to be more reliable when conducting a primary analysis, and is sometimes preferred over the simple imputation approaches using the last, or baseline-observation-carried-forward (BOCF) methods.[4] The BOCF method is another approach used to handle patient drop-outs.[5] The difference between the BOCF and LOCF method is that BOCF uses the baseline observation, whereas LOCF uses the last observation in place of the missing endpoint. The BOCF method requires that the patient remain active in the trial in order to be evaluated for a response. In this method, if the patient drops out from the trial, the baseline observation is treated as the final response from the patient regardless of the reason the patient dropped out or the scores at the time of withdrawal. Regulatory agencies often suggest this method, that may seem to be more conservative, be used when evaluating clinical trials, where it is commonly assumed that a patient withdrew from the trial because of lack of benefit or treatment-emergent adverse events. However, there are various other reasons why patients withdraw from trials, as mentioned previously. Hence, when evaluating the outcome of a clinical trial, it would be prudent to take into consideration treatment-related and nontreatment-related reasons for dropout.

Using data from 2 clinical trials of duloxetine for the treatment of chronic pain, ie, chronic lower back pain (CLBP) (study 1),[6] and chronic pain due to osteoarthritis (OA) of the knee (study 2),[7] we investigated the underlying reasons for patient discontinuation from the trials and their roles in the BOCF approach, and propose examining modifications of the BOCF approach to take into consideration treatment-related and nontreatment-related reasons for discontinuation.

Methods

This research was carried out in two 13-week, randomized, double-blind, placebo-controlled trials. Study 1 included male and female patients (duloxetine N = 115, placebo N = 121) who had a clinical diagnosis of CLBP. Study 2 included male and female patients ≥40 years (duloxetine N = 128, placebo N = 128) with OA of the knee. The patient characteristics are comparable between the two studies. Detailed patient demographics and other information about the patient characteristics for each of the individual studies have been previously presented.[6,7]

The primary study objective of these clinical studies was to assess the efficacy of duloxetine, as compared with placebo, on the reduction of pain severity, as measured by the Brief Pain Inventory (BPI) 24-hour average pain rating.

Statistical analyses

Primary efficacy analysis in both studies was analysis of mean change from baseline in BPI average pain using an MMRM approach.

Both LOCF and BOCF were prespecified in the study protocols as additional analytic approaches. For patients who completed the treatment phase, the BOCF endpoint was defined as the last nonmissing observation, and for patients who discontinued early, the BOCF endpoint was defined as the baseline value.

For patients who discontinued trials early, drop-out reasons were classified as either treatment-related or non-treatment-related reasons. Treatment-related reasons included "adverse events" and "lack of efficacy." Nontreatment-related reasons included "entry criteria not met," "protocol violation," "lost to follow-up," "subject decision," (eg, work conflict, lack of transportation, change of location, or unwillingness to fill out questionnaires) and "physician decision (eg, investigator sites closing or patients deemed unreliable)."

As patients who discontinued due to non-treatment-related reasons provide useful information in assessing the treatment effect, a modified BOCF (mBOCF) endpoint was defined as follows 1) for patients who completed the treatment phase, the BOCF endpoint was defined as the last nonmissing observation, or 2) for patients who discontinued early due to treatment-related reasons (ie, adverse events [AEs] or lack of efficacy [LOE]), the BOCF endpoint was defined as the baseline value, and 3) for patients who

discontinued early due to nontreatment-related reasons (ie, reasons other than AEs or LOE), the BOCF endpoint was defined as the last nonmissing observation.

In the mBOCF approach, the change from baseline to the last observation (ie, the LOCF endpoint) was used for patients who discontinued due to nontreatment-related reasons and for completers, while a change of zero was used for patients who discontinued due to treatment-related reasons.

An additional BOCF approach, ie, aeBOCF was used where a change of zero was utilized for patients who discontinued due to adverse events, while the change from baseline to the last observation (LOCF endpoint) was used for patients who discontinued due to any other reasons as well as for those who completed the studies.

For the mean change analyses using LOCF, BOCF, mBOCF, and aeBOCF approaches, an analysis of covariance (ANCOVA) model including change from baseline-to-endpoint baseline value, treatment, investigator, and treatment-by-investigator was used. Type III sum-of-squares for the least-squares mean (LS mean) was used to assess treatment difference.

In addition to mean change analysis, categorical analyses of response rate (defined as ≥30% reduction in BPI average pain, a change of ≥30% decrease is considered 'moderately important'[8]) were also conducted using the LOCF, BOCF, mBOCF, and aeBOCF approaches. Fisher's exact test was used to assess the treatment difference.

To assess the impact of different drop-out reasons on the patients' efficacy outcomes, change from baseline to LOCF endpoint in BPI average pain was analyzed for the following disposition categories: 1) completions of 13-week treatment, 2) discontinuations due to adverse events, 3) discontinuations due to lack of efficacy, and 4) discontinuations due to reasons other than AEs or LOE.

Effect size was calculated using the treatment difference/(standard error x the square root of N).

Patients with baseline and at least one postbaseline observation were included in all analyses. All tests were 2-sided at the 0.05 significance level. The term 'significant' indicates statistical significance throughout the manuscript.

Results

Patient disposition is shown in Table 1. Compared with the placebo-treated group, the duloxetine-treated group had significantly more discontinuations due to adverse events in both study 1 ($P < 0.05$) and study 2 ($P < 0.01$). In addition, in study 2, there was a significant difference ($P < 0.01$) between the duloxetine-treated (72.7%) and placebo-treated (86.7%) groups in the number of patients who completed the study. In both studies, there were no differences between treatment groups for discontinuations due to subject decision, protocol violation, physician decision, lost to follow-up, lack of efficacy, and entry criteria not met.

In study 1, duloxetine was superior on the primary efficacy measure of mean change in the BPI average pain (Table 2), using the MMRM (–2.32), BOCF (–1.86), LOCF (–2.09), and mBOCF (–1.91), compared with placebo (MMRM [–1.50, $P = 0.004$], BOCF [–1.25, $P = 0.019$], LOCF [–1.45, $P = 0.019$], and mBOCF [–1.35, $P = 0.041$]). However, using the aeBOCF approach, there was no significant difference between duloxetine [–1.94] and placebo [–1.46, $P = 0.075$]. In study 2, duloxetine was superior in the

Table 1 Patient disposition

	Study 1		Study 2	
	Placebo (N = 121) n (%)	Duloxetine 60/120 mg/day (N = 115) n (%)	Placebo (N = 128) n (%)	Duloxetine 60/120 mg/day (N = 128) n (%)
Completers of the 13-week treatment	98 (81.0)	84 (73.0)	111 (86.7)	93 (72.7)**
Discontinuation for any reason	23 (19.0)	31 (27.0)	17 (13.3)	35 (27.3)**
Adverse event	7 (5.8)	16 (13.9)*	7 (5.5)	24 (18.8)**
Subject decision	10 (8.3)	11 (9.6)	2 (1.6)	4 (3.1)
Protocol violation	2 (1.7)	2 (1.7)	2 (1.6)	3 (2.3)
Physician decision	2 (1.7)	1 (0.9)	0 (0.0)	2 (1.6)
Lost to follow-up	1 (0.8)	1 (0.9)	0 (0.0)	1 (0.8)
Lack of efficacy	1 (0.8)	0 (0.0)	5 (3.9)	1 (0.8)
Entry criteria not met	0 (0.0)	0 (0.0)	1 (0.8)	0 (0.0)

Notes: *$P < 0.05$ versus placebo; **$P < 0.01$ versus placebo.
Abbreviations: n, number of patients in the specified category; N, total number of randomized patients.

Table 2 Mean change analysis of Brief Pain Inventory average pain for all randomized patients in the 13-week treatment phase

Study	Analysis	Treatment group[a]	LS mean change (SE)	Effect size	*P* value[b]
Study 1[c]	MMRM[d]	DLX 60/120 mg QD	–2.32 (0.22)	0.36	0.004
		Placebo	–1.50 (0.21)		
	BOCF	DLX 60/120 mg QD	–1.86 (0.20)	0.28	0.019
		Placebo	–1.25 (0.20)		
	LOCF	DLX 60/120 mg QD	–2.09 (0.21)	0.28	0.019
		Placebo	–1.45 (0.21)		
	mBOCF	DLX 60/120 mg QD	–1.91 (0.21)	0.25	0.041
		Placebo	–1.35 (0.21)		
	aeBOCF	DLX 60/120 mg QD	–1.94 (0.21)	0.21	0.075
		Placebo	–1.46 (0.21)		
Study 2[e]	MMRM[d]	DLX 60/120 mg QD	–2.72 (0.20)	0.41	<0.001
		Placebo	–1.88 (0.18)		
	BOCF	DLX 60/120 mg QD	–2.23 (0.20)	0.28	0.013
		Placebo	–1.63 (0.19)		
	LOCF	DLX 60/120 mg QD	–2.51 (0.20)	0.39	<0.001
		Placebo	–1.72 (0.18)		
	mBOCF	DLX 60/120 mg QD	–2.29 (0.20)	0.32	0.005
		Placebo	–1.61 (0.19)		
	aeBOCF	DLX 60/120 mg QD	–2.29 (0.20)	0.31	0.005
		Placebo	–1.62 (0.19)		

[a]Study 1: N (DLX 60/120 QD) = 109, N (placebo) = 115.
Study 2: N (DLX 60/120 QD) = 121, N (placebo) = 127.
[b]*P* value comparison with placebo.
[c]Baseline mean (standard deviation): DLX 60/120 mg QD = 5.91 (1.61), placebo = 5.93 (1.67).
[d]Primary efficacy analysis in study 1.
[e]Baseline mean (standard deviation): DLX 60/120 mg QD = 6.09 (1.38), placebo = 6.16 (1.26).
Abbreviations: aeBOCF, discontinuation due to adverse events only; BOCF, baseline observation carried forward; DLX, duloxetine; LOCF, last observation carried forward; LS Mean, least-squares mean; mBOCF, modified BOCF; MMRM, mixed-model repeated measures; QD, once daily; SE, standard error.

BPI average pain, using the MMRM (–2.72), BOCF (–2.23), LOCF (–2.51), mBOCF (–2.29), and aeBOCF (–2.29), compared with placebo (MMRM [–1.88, $P < 0.001$], BOCF [–1.63, $P = 0.013$], LOCF [–1.72, $P < 0.001$], mBOCF [–1.61, $P = 0.005$], and aeBOCF [–1.62, $P = 0.005$]). Effect size information is also provided.

There were no differences between the duloxetine and placebo groups in the percentage of responders in study 1 (BOCF [45.9% versus 33.0%, $P = 0.056$], LOCF [53.2% versus 40.0%, $P = 0.060$], mBOCF [47.7% versus 37.4%, $P = 0.137$], and aeBOCF [48.6% versus 40.0%, $P = 0.226$]). In study 2, results of the 30% response rates (Table 3) showed a significantly greater percentage of responders in the duloxetine group than in the placebo group, using the BOCF (57% versus 42.5%, $P = 0.031$), LOCF (65.3% versus 44.1%, $P < 0.001$), mBOCF (59.5% versus 42.5%, $P = 0.008$), and aeBOCF (59.5% versus 42.5%, $P = 0.008$) approaches.

Regardless of the disposition reasons, the treatment difference (not statistically significant) in the BPI average pain between the duloxetine-treated and placebo-treated groups occurred in the same direction as in completers (Table 4)

Table 3 Thirty percent response rate to Brief Pain Inventory average pain

Study	Analysis	Treatment group[a]	Responder[b] n (%)	*P* value[c]
Study 1	BOCF	DLX 60/120 mg QD	50 (45.9)	0.056
		Placebo	38 (33.0)	
	LOCF	DLX 60/120 mg QD	58 (53.2)	0.060
		Placebo	46 (40.0)	
	mBOCF	DLX 60/120 mg QD	52 (47.7)	0.137
		Placebo	43 (37.4)	
	aeBOCF	DLX 60/120 mg QD	53 (48.6)	0.226
		Placebo	46 (40.0)	
Study 2	BOCF	DLX 60/120 mg QD	69 (57.0)	0.031
		Placebo	54 (42.5)	
	LOCF	DLX 60/120 mg QD	79 (65.3)	<0.001
		Placebo	56 (44.1)	
	mBOCF	DLX 60/120 mg QD	72 (59.5)	0.008
		Placebo	54 (42.5)	
	aeBOCF	DLX 60/120 mg QD	72 (59.5)	0.008
		Placebo	54 (42.5)	

[a]Study 1: N (DLX 60/120 mg QD) = 109, N (placebo) = 115.
Study 2: N (DLX 60/120 mg QD) = 121, N (placebo) = 127.
[b]Response was defined as at least a 30% reduction in BPI average pain.
[c]*P* value comparison with placebo.
Abbreviations: aeBOCF, discontinuation due to adverse events only; BOCF, baseline observation carried forward; BPI, Brief Pain Inventory; DLX, duloxetine; LOCF, last observation carried forward; mBOCF, modified BOCF; MMRM, mixed-model repeated measures; N, number of patients in the specified category; QD, once daily.

Table 4 Mean change analysis of BPI average pain (LOCF approach) for all randomized patients – by disposition status

	Study 1				Study 2			
	Placebo		Duloxetine 60/120 mg/day		Placebo		Duloxetine 60/120 mg/day	
	n	LS mean change (SE)	n	LS mean change (SE)	n	LS mean change (SE)	n	LS mean change (SE)
Completers	98	–1.51 (0.20)	84	–2.47 (0.22)***	111	–2.02 (0.17)	92	–3.04 (0.19)***
Discontinuation due to								
Adverse events	4	0.85 (1.46)	13	–0.65 (1.06)	7	–0.88 (0.72)	20	–0.93 (0.45)
Lack of efficacy	4	–0.79 (1.44)	3	–0.79 (1.30)	6	–0.36 (0.23)	1	–1.62 (0.54)
Other reasons	9	–1.48 (0.95)	9	–0.49 (0.93)	3	0.07 (1.08)	8	–1.64 (0.69)

***$P \leq 0.001$ versus placebo.

Abbreviations: BPI, Brief Pain Inventory; LOCF, last observation carried forward; LS, least squares; n, number of patients in the specified category; SE, standard error.

except for discontinuation due to reasons other than AEs or LOE in study 1.

Discussion

In both studies 1 and 2, duloxetine was superior to placebo on the primary efficacy measure (ie, change in BPI average pain) using the MMRM, BOCF, LOCF, and mBOCF approaches. Using aeBOCF, duloxetine did not separate from placebo in study 1, but did so in study 2. In study 2, the pattern seen in response rates (ie, ≥30% reduction in BPI average pain) was similar to that in the mean change analysis of BPI average pain ratings, with BOCF, LOCF, mBOCF, and aeBOCF approaches all demonstrating superiority of duloxetine over placebo. However, in study 1, none of the analytical methods demonstrated statistically significant differences between duloxetine and placebo in the 30% response rate of BPI average pain.

In patients who discontinued the trials, the drug-placebo treatment differences were generally in the same direction as in completers with the exception of discontinuation due to reasons other than AEs or LOE in study 1. However, none of these differences were statistically significant. While a consistent treatment difference was observed in completers, and a general trend was also seen among drop-outs, a definitive characterization of the treatment effect in each of the discontinuation categories was limited by the small sample sizes.

The BOCF approach does not consider changes observed in patients who discontinued due to reasons unrelated to treatment. A previous publication of a multicenter, multi-study database reported factors associated with early study discontinuation[9,10] and demonstrated that some of the variables associated with high loss to follow-up were age, female sex, African-American race, no previous enrollment in a study, and geography of sites other than the central United States. Additionally, in large multicenter trials, it is not uncommon to see a higher rate of discontinuation due to nontreatment-related events in some countries compared with others, and this may be due to cultural differences among populations worldwide.

With the BOCF approach, the group of patients that are not taken into consideration because of discontinuation due to nontreatment-related reasons may actually have benefited from the studied treatment and may also benefit in a clinical practice setting, hence these patients should not be discounted in the analyses. In addition, even though the BOCF approach is sometimes perceived to be more conservative, this is not always the case, as demonstrated by the results from study 1 which showed a lower *P* value for the BOCF method ($P = 0.019$) versus the mBOCF ($P = 0.041$) and the aeBOCF ($P = 0.075$) methods.

The mBOCF approach is designed to take into consideration the reasons for discontinuation by treating patients who discontinue due to treatment-related and nontreatment-related reasons differently. In the mBOCF approach, the change in mean pain scores from baseline to the last observation (LOCF endpoint) is used for patients who discontinue due to nontreatment-related reasons, while the baseline score (BOCF endpoint) is used for patients who discontinue due to treatment-related reasons. For patients who discontinued due to nontreatment-related reasons, their response may still reflect the true treatment response. Therefore, in the mBOCF approach, actual treatment outcomes from these patients were included in the assessment of the treatment difference between duloxetine and placebo. The aeBOCF approach was designed to take into consideration the reasons for discontinuation by treating patients who discontinue due to

treatment-related (adverse events) reasons only. The rationale for using aeBOCF is that if a drug is intolerable and leads to discontinuation, then it is considered not effective. By the similar argument, if a treatment is discontinued due to lack of efficacy, it should also be considered not effective.

Because patients who discontinue early from the clinical trials due to nontreatment-related reasons may make a reasonable contribution when assessing treatment differences, consideration should be given to including these patients in the assessment of treatment effects, as in the mBOCF approach examined in these analyses. This would prevent biasing the outcome of a clinical trial due to patient drop-out for nontreatment-related reasons. It is a common and real challenge to handle early discontinuation in longitudinal clinical trials. It is therefore important to not rely solely on one single approach as the only statistical method to draw inferences. Different methods and sensitivity analyses should be utilized to determine the robustness of the conclusions.

Acknowledgments/disclosures

This work was sponsored by Eli Lilly and Company and Boehringer Ingelheim GmbH.

Hong Liu-Seifert, Shuyu Zhang, Deborah D'Souza, and Vladimir Skljarevski are employees and stockholders of Eli Lilly and Company.

References

1. Liu G, Gould AL. Comparison of alternative strategies for analysis of longitudinal trials with dropouts. *J Biopharm Stat*. 2002;12(2):207–226.
2. Lane P. Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. *Pharm Stat*. 2008;7(2):93–106.
3. Mallinckrodt CH, Sanger TM, Dubé S, et al. Assessing and interpreting treatment effects in longitudinal clinical trials with missing data. *Biol Psychiatry*. 2003;53(8):754–760.
4. Mallinckrodt CH, Lane PW, Schnell D, Peng Y, Mancuso JP. Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials. *Drug Inf J*. 2008;42(4):303–319.
5. Shao J, Jordan DC, Pritchett YL. Baseline observation carry forward: reasoning, properties, and practical issues. *J Biopharm Stat*. 2009;19(4):672–684.
6. Skljarevski V, Desaiah D, Liu-Seifert H, et al. Efficacy and safety of duloxetine in patients with chronic low back pain. *Spine*. 2009. In press.
7. Chappell AS, Desaiah D, Liu-Seifert H, et al. Duloxetine 60 to 120 mg once daily versus placebo in the treatment of patients with osteoarthritis knee pain. [Poster] Presented at the American Academy of Pain Medicine (AAPM) – 25th Annual Meeting, Honolulu, HI, USA; Jan 28–31, 2009.
8. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain*. 2008;9(2):105–121.
9. Andersen JW, Van Der Horst C, Gorski H, Fass RJ. Conference on Retroviruses and Opportunistic Infections. The association of site with early study discontinuation on HIV clinical studies in the protease inhibitor era: DACS200. Program Abstr 8th Conf Retrovir Oppor Infect Conf Retrovir Oppor Infect 8th 2001 Chic Ill; 2001 Feb 4–8;8:188 (abstract no. 486).
10. Anderson JW, Fass R, van der Horst C. Factors associated with early study discontinuation in AACTG studies, DACS 200. *Contemp Clin Trials*. 2007;28(5):583–592.