CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

022516Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	NDA 22-516
Drug Name:	Cymbalta
Indication(s):	For the management of chronic pain
Applicant:	Eli Lilly and Company
Date(s):	Submitted: May 15, 2009
	PDUFA: March 15, 2010
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Yongman Kim, Ph.D.
Concurring Reviewers:	Dionne Price, Ph.D.
	Thomas Permutt, Ph.D.
Medical Division:	Division of Anesthesia, Analgesia, and Rheumatology Products
Clinical Team:	Anjelina Pokrovnichka, M.D.
	Ellen Fields, M.D.
Project Manager:	Ayanna Augustus, Ph.D.

Keywords: Clinical studies, missing data, sensitivity analyses

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

The applicant has evaluated the use of duloxetine 60 – 120 mg for the treatment of chronic pain. Based on my review of the data, I find that patients receiving duloxetine experienced greater pain reduction compared to patients receiving placebo. The treatment effect was evident in a single trial conducted in osteoarthritis (OA) patients as well as two trials conducted in patients with chronic low back pain (CLBP). Due to concerns about hepatotoxicity regarding the 120 mg dose, additional analyses focusing on the 60 mg dose were conducted. I conclude that the 60 mg dose was effective in reducing pain for the 12-week trial duration in a single CLBP study. Supportive evidence of this effect was also derived from my analysis of pain reductions up to 7 week and applicant's post-hoc analysis of 13-week pain reductions treating non-responders at Week 7 as treatment failures regardless of randomized treatment group in a second CLBP study. However, I conclude that there was no sufficient data to support the efficacy of duloxetine 60 mg in treating OA because the successful OA study with combined dose did not plan to provide evidence of effects of the 60 mg dose separated from the 120 mg dose although my and applicant's post-hoc analyses on effects of the 60 mg dose demonstrated significance in the study.

1.2 Brief Overview of Clinical Studies

Four studies (HMEP, HMEN, HMEO, and HMFG) were submitted at the time of filing and one study (HMGC) was submitted via the 120 day safety update.

Study F1J-MC-HMEP (HMEP hereafter) was a 13-week, double-blind, placebo-controlled, international, multi-center trial investigating the safety and efficacy of duloxetine in patients with osteoarthritis knee pain. In the study, 231 patients were randomized to duloxetine 60 mg (n = 111) or placebo (n = 120). At Week 7, patients initially randomized to duloxetine 60 mg were re-randomized to either duloxetine 60 mg or duloxetine 120 mg. The primary efficacy variable was the change from baseline to Week 13 in the weekly mean of the 24-hour average pain. Secondary efficacy measures included Patient Global Impression of Improvement (PGI-I) and Western Ontario and McMaster Universities (WOMAC) physical function subscale.

Study F1J-MC-HMEN (HMEN hereafter) was a 13-week, double-blind, placebo-controlled, international, multi-center trial investigating the safety and efficacy of duloxetine in patients with chronic low back pain. In the study, 236 patients were randomized to duloxetine 60 mg (n = 115) or placebo (n = 121). At Week 7, patients who were randomized to duloxetine 60 mg and did not meet response criterion defined as at least 30% reduction in pain scores had their dose increased to 120 mg. The primary efficacy outcome was the pain severity as measured by the BPI 24-hour average pain scores from baseline to Week 13. Secondary efficacy measures included Clinical Global

Impression of Severity (CGI-Severity) and Roland-Morris Disability Questionnaire (RMDQ-24).

Study F1J-MC-HMEO (HMEO hereafter) was a 13-week, double-blind, placebo-controlled, international, multi-center trial investigating the safety and efficacy of duloxetine in patients with chronic low back pain. In the study, 404 patients were randomized to duloxetine 20 mg (n = 59), duloxetine 60 mg (n = 116), duloxetine 120 mg (n = 112), or placebo (n = 117). The primary efficacy variable was the change from baseline to Week 13 in the weekly mean of the 24-hour average pain. Secondary efficacy measures included Patient Global Impression of Improvement (PGI-I) and Roland-Morris Disability Questionnaire (RMDQ-24).

Study F1J-MC-HMFG (HMFG hereafter) was a 13-week, double-blind, placebo-controlled, international, multi-center trial investigating the safety and efficacy of duloxetine in patients with osteoarthritis knee pain. In the study, 256 patients were randomized to duloxetine 60 mg (n = 128) or placebo (n = 128). At Week 7, patients who were randomized to duloxetine 60 mg and did not meet response criterion defined as at least 30% reduction in pain scores had their dose increased to 120 mg. The primary efficacy outcome was the pain severity as measured by the BPI 24-hour average pain scores from baseline to Week 13. Secondary efficacy measures included the patient reported outcomes such as Patient Global Impression of Improvement (PGI-I) and Western Ontario and McMaster Universities (WOMAC) physical function subscale.

Study HMGC was a 12-week, double-blind, placebo-controlled, international, multi-center trial investigating the safety and efficacy of duloxetine in patients with chronic low back pain. In the study, 401 patients were randomized to duloxetine 60 mg (n = 198) or placebo (n = 203). The primary efficacy outcome was the pain severity as measured by the BPI 24-hour average pain scores from baseline to Week 12. Secondary efficacy measures included the patient reported outcomes such as Patient Global Impression of Improvement (PGI-I) and Roland-Morris Disability Questionnaire (RMDQ-24).

1.3 Statistical Issues and Findings

Study HMEP was a flexible dose trial conducted in OA patients. The study demonstrated a significant effect of duloxetine 60 mg-120 mg compared to placebo in the pre-specified mixed-effect repeated measures analysis on the evaluations of pain severity as measured by the weekly mean of 24-hour average pain scores from the baseline to Week 13. However, a conservative analysis employing the baseline observation carried forward imputation strategy for missing data did not demonstrate the statistical significance.

Study HMEN was a flexible dose trial conducted in CLBP patients. The study demonstrated a significant effect of duloxetine 60 mg-120 mg compared to placebo from the evaluations of BPI pain severity from the baseline to Week 13 in a conservative analysis using a baseline observation carried forward imputation strategy for missing data.

Study HMEO was a fixed dose trial conducted in CLBP patients. Based on the applicant's analysis, the study failed to demonstrate a significant effect in for the duloxetine dose of 20 mg, 60mg, or 120 mg compared to placebo from the evaluations of the weekly mean of pain severity from the baseline to Week 13 in a conservative analysis employing the baseline observation carried forward imputation strategy for missing data.

Study HMFG was a flexible dose trial conducted in OA patients. When evaluating the data using the baseline observation carried forward imputation strategy, the study demonstrated a significant effect of duloxetine 60 mg-120 mg compared to placebo Study HMGC was a fixed dose trial conducted in CLBP patients. The study demonstrated a significant effect of duloxetine 60 mg compared to placebo in a conservative analysis employing the baseline observation carried forward imputation strategy for missing data.

There were several issues in the statistical analyses of the data. In chronic pain trials, patients who withdraw before the end of the study should be treated as non-responders, and no benefit should be assigned based on the pain scores before dropout. However, the applicant proposed the mixed effect repeated measures (MMRM) analysis as the primary analysis method in all studies. The method uses pain data from patients who withdraw before the study ends. Also in order for the MMRM to be valid, missing at random (MAR) should be assumed as the mechanism generating missing data. However in chronic pain trials, missing data is often informative and therefore the MAR assumption is not supported.

Second, because of the safety concern of hepatotoxicity regarding duloxetine 120 mg, the clinical team focused on the efficacy of duloxetine 60 mg. However with the exception of Study HMGC, the studies were designed to evaluate the combined the 60 mg and 120 mg doses of duloxetine. HMEP re-randomized patients initially randomized to duloxetine 60 mg to duloxetine 60 mg or 120 mg at Week 7. HMEP did not demonstrate a significant difference when I compared patients treated with duloxetine 60 mg over 13 weeks with patients treated with placebo. HMEN and HMFG increased the dose of duloxetine from 60 mg to 120 mg for patients who were randomized to duloxetine 60 mg but did not show at least a 30% reduction of pain from baseline at Week 7. In analyses submitted in response to information requests, the applicant treated patients who were not showing at least a 30% pain reduction from baseline at Week 7 as failures regardless of randomized treatment group, and the baseline scores were imputed to those patients. The analyses demonstrated a significant difference between duloxetine 60 mg and placebo. I additionally conducted analyses of the change from baseline up to Week 7. In my analyses, HMEN and HMFG demonstrated a significant difference between duloxetine 60 mg and placebo.

Although not statistical, another issue that is noteworthy is the fact that the applicant seeks a more broad chronic pain indication. The requirement for a broad chronic pain indication is a current topic under consideration within the Agency as well as the clinical community.

2. INTRODUCTION

2.1 Overview

2.1.1 Drug class and regulatory history

Cymbalta (duloxetine) is approved in the United States for the treatment of major depressive disorder (MDD), generalized anxiety disorder (GAD), diabetic peripheral neuropathic pain (DPNP), and Fibromyalgia (FM). The current NDA is submitted in support of an indication for chronic pain. In a pre-IND meeting held in 2005, the clinical development plan for chronic low back pain was discussed and statistical advice on missing data handling and multiple testing on secondary endpoints was provided. In a teleconference held in 2006, the development plan for OA and CLBP studies was discussed. Conservative imputation methods for missing data and continuous responder analyses were recommended. In 2008, the applicant submitted NDA 22-333 for a broad chronic pain indication with data from studies HMEO, HMEP, and HMEN. However, after interaction between the applicant and the Division regarding insufficient data to support the indication, the applicant withdrew the NDA. In 2009, the applicant resubmitted an updated NDA (N22-516) with a new study, HMFG, and open-label extension period data from Study HMEN. During the review cycle of the NDA, along with 120-day safety update, the applicant submitted data from Study HMGC with a fixed dose of duloxetine 60 mg.

2.1.2 Proposed Indication

The proposed indication is for the treatment of chronic pain including management of diabetic peripheral neuropathic pain (DPNP), fibromyalgia (FM), chronic pain due to osteoarthritis (OA) and chronic low back pain (CLBP).

2.2 Data Sources

NDA 22-516 was submitted on May 15, 2009. Data are located in the electronic document room (EDR) of the Center for Drug Evaluation and Research. The electronic SAS data sets were also provided in the EDR using the following path:

\\CDSESUB1\EVSPROD\NDA022516

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

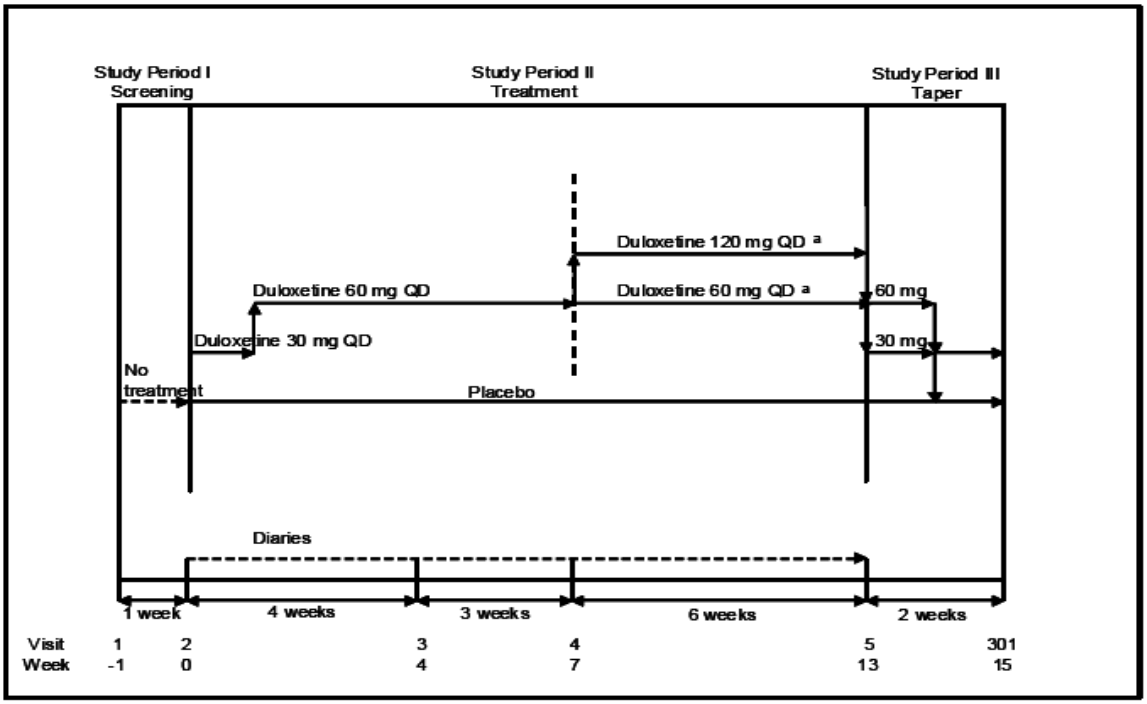
My review will focus on studies HMEP, HMEN, HMFG, and HMGC. I will not review HMEO because it was declared as a failed study by the applicant.

3.1.1 HMEP

3.1.1.1 Study Design and Endpoints

Study HMEP was a 13-week, double-blind, placebo-controlled, international, multi-center trial investigating the safety and efficacy of duloxetine (DLX hereafter) in patients with OA knee pain. In HMEP, 231 eligible patients were randomized in a 1:1 ratio to DLX 60 mg QD (n = 111) or placebo (n = 120) stratified by non-steroidal anti-inflammatory drug (NSAID) use at 29 centers in 3 countries worldwide. Patients were required to have a mean score of 4 or greater on the 24-hour average pain score measured on the 11-point Likert scale. Figure 1 presents a schematic of the study design.

Figure 1 Schematic of Study Design: HMEP



a Duloxetine arm re-randomized to 60 QD or 120 QD.

At Week 7, patients who were initially randomized to duloxetine 60 mg and had stayed in the trial were re-randomized to either duloxetine 60 mg (n=46) or duloxetine 120 mg (n=43).

Use of some short-acting analgesics, such as acetaminophen and codeine, was allowed for rescue from an OA knee pain flare after screening.

The primary efficacy endpoint was the weekly mean of the 24-hour average pain rating from patient diaries. The endpoint was measured on the Likert scale, an ordinal scale with scores from 0 (no pain) to 10 (worst possible pain).

The secondary endpoints proposed for possible inclusion in the label were:

- Patient Global Impression of Improvement (PGI-I)
- Western Ontario and McMaster Universities (WOMAC) physical function subscale.

Other secondary efficacy variables included the following:

- Brief Pain Inventory (BPI)-Severity and Interference
- Clinical Global Impression of Severity (CGI-S)
- The Hospital Anxiety and Depression Scale (HADS).

3.1.1.2 Disposition and Demographics

Approximately 25% of the patients discontinued before the end of study (Table 1). However, more patients from the DLX group discontinued compared to the placebo group. Thirty-one percent of DLX patients discontinued while 20% of placebo patients discontinued. As expected, the majority of DLX dropouts were due to adverse events. Fourteen percent of DLX patients discontinued due to adverse events. However, unexpectedly, the majority of placebo dropouts were not due to lack of efficacy, but due to subject decision. Six percent of placebo patients discontinued due to adverse events and 2% of placebo patients discontinued due to lack of efficacy.

Table 1 Subject Disposition: HMEP

	Number of Patients	
	Placebo	**DLX 60-120mg**
Randomized	120 (100%)	111 (100%)
ITT	120	111
mITT	119	108
Completed	96 (80%)	77 (69%)
Reasons for dropout		
AE	7 (6%)	15 (14%)
LOE	3 (2%)	2 (2%)

Subject Decision	9 (8%)	8 (7%)
Other	5 (3%)	9 (8%)

Patient demographics are presented by treatment groups in the appendix (Table 23). There were no noticeable imbalances among treatment groups with respect to demographic variables of age, race, sex, and weight.

Table 23 also shows baseline values for the efficacy variable of BPI average pain score by treatment groups. Distributions of the efficacy variables at baseline were comparable between treatment groups.

3.1.1.3 Statistical Methodologies

The primary analysis used a mixed-model repeated measures (MMRM) model on the change from baseline to Week 13 of the weekly mean of the 24-hour average pain scores. The model included terms for treatment, NSAID use, site, week and treatment-by-week interactions, and baseline pain score and baseline-by-week interaction. The contrast at Week 13 comparing treatments was used to test if DLX was superior to placebo. In addition, an analysis of covariance (ANCOVA) model was used to compare treatments. The ANCOVA model included terms for treatment, NSAID use, site, and baseline pain score as a covariate.

To assess the impact of missing data on the ANCOVA analysis, the analysis was conducted with last observation carried forward (LOCF), baseline observation carried forward (BOCF), and modified BOCF (mBOCF) imputation strategies. In the mBOCF approach, a BOCF strategy was used to impute missing data from dropouts due to lack of efficacy (LOE) or adverse event (AE) and an LOCF strategy was used to impute missing data from dropouts due to other reasons. As a sensitivity analysis, I conducted a continuous responder analysis treating dropouts as non-responders.

The primary analysis was conducted on the modified intent-to-treat (mITT) population defined as all patients who were randomized and had baseline scores and at least 1 post-baseline observation. I conducted the primary analysis on the ITT population defined as all patients who were randomized.

For the analysis of the Patient Global Impression of Improvement (PGI-I) and the Western Ontario and McMaster Universities (WOMAC) physical function subscale, an ANCOVA model with terms for treatment, NSAID use, site, and baseline score as covariate was used.

In order to adjust for multiple testing on these secondary endpoints, a serial gate-keeper multiple testing method was used, i.e., PGI-I and WOMAC physical function subscale were tested sequentially only if the primary endpoint was statistically significant.

3.1.1.4 Results and Conclusions

A greater treatment effect was achieved by patients receiving DLX 60-120mg as compared to those receiving placebo (Table 2).

Table 2 Applicant's Primary Efficacy Analysis: HMEP (mITT)

LS Mean Change (SE) from Baseline to Week 13 in weekly mean of 24-hour average pain	**Placebo (N=119)**	**DLX60-120mg (N=108)**	**P-value**
MMRM*	-2.1 (0.16)	-2.9 (0.17)	**<0.001**
ANCOVA/BOCF**	-1.8 (0.19)	-2.2 (0.20)	0.086

*P-value calculated from MMRM model with terms for treatment, week, treatment*week, site, NSAID use, baseline, week*baseline.
**P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.
Note: mITT population excluded patients with no post-baseline observations.

Because I could not reproduce the applicant's primary analysis, I conducted the same ANCOVA analysis with BOCF on the mITT population. My analysis gave similar conclusions as the applicant's analysis (Table 3).

Table 3 Reviewer's Primary Efficacy Analysis: HMEP (mITT)

LS Mean Change (SE) from Baseline to Week 13 in weekly mean of 24-hour average pain	**Placebo (N=120)**	**DLX60-120mg (N=111)**	**P-value**
ANCOVA/BOCF*	-1.8 (0.19)	-2.0 (0.20)	0.338

*P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.
Note: mITT population excluded patients with no post-baseline observations.

The applicant's primary analysis excluded four patients who had no post-baseline observations. I conducted the analyses on the ITT population including those patients. My BOCF analysis and continuous responder analysis on the ITT population did not demonstrate a statistically significant difference although my mBOCF analysis demonstrated a statistically significant difference (Table 4 & Figure 2). This is partly due to the fact that more than half of the dropouts from DLX were not attributed to clinical

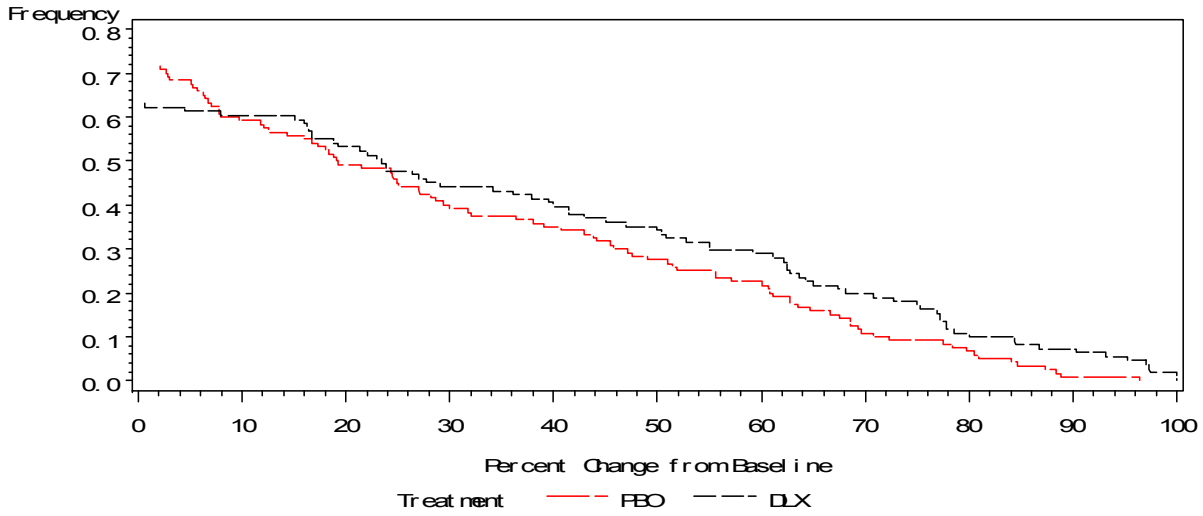
reasons such as LOE or AE and, therefore, the mBOCF used LOCF for those dropouts, which led to imputations of good scores to the majority of DLX dropouts. Due to this concern with the mBOCF analysis and the failure in a conservative BOCF analysis, I considered the study be a failure.

Table 4 Reviewer's Primary Efficacy Analysis: HMEP (ITT)

LS Mean Change (SE) from Baseline to Week 13 in weekly mean of 24-hour average pain	**Placebo (N=120)**	**DLX60-120mg (N=111)**	**P-value**
ANCOVA/BOCF*	-1.7 (0.19)	-2.0 (0.19)	0.412
ANCOVA/mBOCF*	-1.9 (0.13)	-2.5 (0.14)	**0.002**

*P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.

Figure 2 Reviewer's Continuous Responder Analysis on Primary Efficacy Variable: HMEP (ITT)



Note: P-value of 0.884 is generated by van der Waerden test.

The secondary efficacy analyses on PGI-I and WOMAC physical function appeared to demonstrate a statistically significant difference (Tables 5 & 6). However, since the primary analysis failed to demonstrate statistical significance in my opinion, the results of the analyses of PGI-I and WOMAC physical function cannot be included in the label.

Table 5 Applicant's Analysis of Patient Global Impression: HMEP

Patient Global Impression - Improvement	**Placebo (N=114)**	**DLX60-120mg (N=104)**
LS Means (SE)	2.9 (0.12)	2.4 (0.12)
p-value vs. Placebo*		**0.001**

*P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.

Table 6 Applicant's Analysis of WOMAC Physical Function: HMEP

WOMAC Physical Function Change from Baseline	**Placebo (N=117)**	**DLX60-120mg (N=107)**
LS Means (SE)	-3.2 (0.35)	-4.6 (0.35)
p-value vs. Placebo*		**0.003**

*P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.

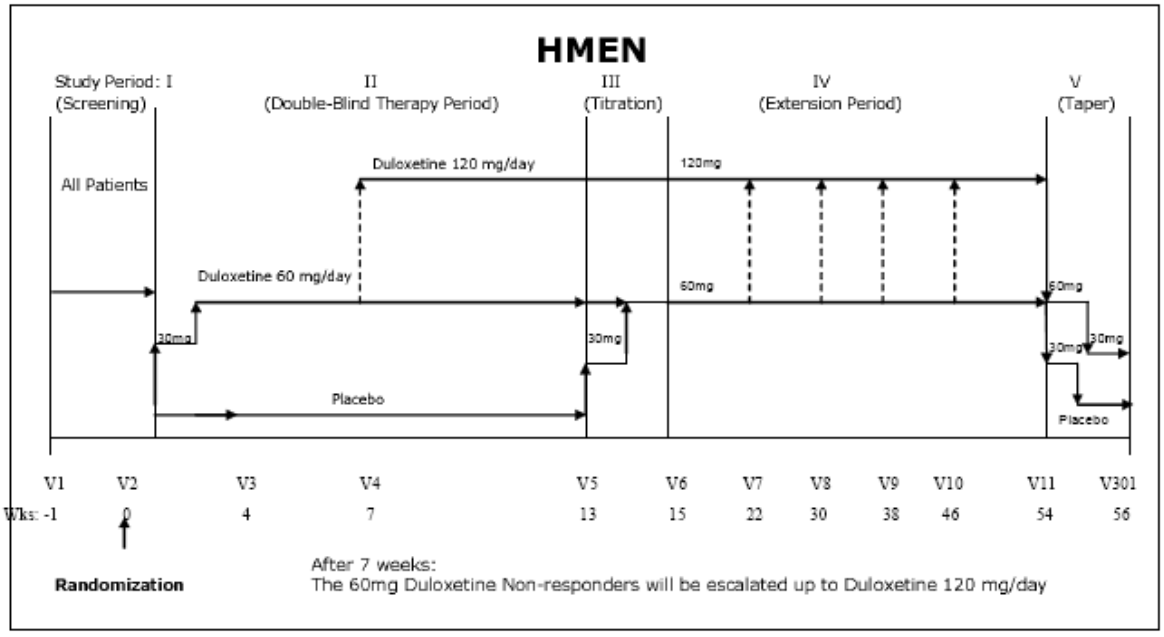
In summary, the applicant's MMRM analysis conducted on the mITT population demonstrated a statistically significant difference between treatments. However, the MMRM analysis is not acceptable because dropouts resulting from a 'bad' outcome such as intolerability of study drug may be artificially be assigned some benefit from treatment. In addition, the analysis is based on an untenable MAR assumption. The mITT analysis population is not acceptable because the mITT excludes patients without post-baseline observations, which could lead to a biased conclusion of a treatment effect. The conservative approaches in handling missing data and in the definition of the analysis set are preferred since these approaches result in conclusions not biased towards favoring study drug. The BOCF analysis and continuous responder analysis conducted on the ITT population which used a conservative method in handling missing data did not demonstrate a statistically significant difference.

3.1.2 HMEN

3.1.2.1 Study Design and Endpoints

Study HMEN was a 13-week, double-blind, placebo-controlled, international, multi-center trial investigating the safety and efficacy of duloxetine in patients with CLBP. In HMEN, 236 eligible patients were randomized in a 1:1 ratio to DLX 60 mg QD (n = 115) or placebo (n = 121) stratified by non-steroidal anti-inflammatory drug (NSAID) use at 19 centers in 5 countries worldwide. Patients were required to have a mean score of 4 or greater on the 24-hour average pain score (0-10). Figure 3 presents a schematic of the study design.

Figure 3 Schematic of Study Design: HMEN



At week 7, patients who did not meet response criteria, defined as at least 30% reduction in weekly mean of the BPI average score compared to baseline, had their dose increased to 120 mg QD.

The primary efficacy measure was the BPI average pain score. The BPI average pain score is a self-reported scale that measures the severity of pain over last 24-hours. The scores range from 0 (no pain) to 10 (pain as severe as you can imagine).

The secondary endpoints proposed for possible inclusion in the label were:

- Patient Global Impression of Improvement (PGI-I).
- Roland-Morris Disability Questionnaire (RMDQ-24) total score.

The other secondary efficacy variables included the following:

- Patient Global Impression of Improvement (PGI-I))
- Clinical Global Impression of Severity (CGI-S)

- Beck Depression Inventory-II (BDI-II)
- The Hospital Anxiety and Depression Scale (HADS)
- Athens Insomnia Scale (AIS).

3.1.2.2 Disposition and Demographics

Approximately 23% of the patients discontinued before the end of study (Table 7). However, more patients from the DLX group discontinued compared to placebo group. Twenty-seven percent of DLX patients discontinued while 19% of placebo patients discontinued. The majority of DLX dropouts were due to adverse events. Fourteen percent of DLX patients discontinued due to adverse events. However, unexpectedly, the majority of placebo dropouts were not due to lack of efficacy, but due to subject decision. Six percent of placebo patients discontinued due to adverse events and 1% of placebo patients discontinued due to lack of efficacy.

Table 7 Subject Disposition: HMEN

	Number of Patients	
	Placebo	**DLX 60-120mg**
Randomized	121 (100%)	115 (100%)
ITT	121	115
mITT	115	109
Completed	98 (81%)	84 (73%)
Reasons for dropout		
AE	7 (6%)	16 (14%)
LOE	1 (1%)	0 (0%)
Subject Decision	10 (8%)	11 (10%)
Other	5 (4%)	4 (3%)

Patient demographics are presented by treatment groups in the appendix (Table 23). There were no noticeable imbalances among treatment groups with respect to demographic variables of age, race, sex, and weight.

Table 23 also shows baseline values for the efficacy variable of BPI average pain score by treatment groups. Distributions of the efficacy variable at baseline were comparable between treatment groups.

3.1.2.3 Statistical Methodologies

Statistical methods are identical to those in HMEP. The primary analysis used the MMRM with same terms as in HMEP. Similarly, ANCOVA model was used to compare treatments on the primary endpoint. The ANCOVA model includes same terms as in HMEP.

To assess impact of missing data on the primary analysis, the same ANCOVA analysis was conducted with LOCF, BOCF, and mBOCF imputation strategies. As a sensitivity analysis, I conducted the same continuous responder analysis as in HMEP.

The primary analysis was conducted on the mITT population. I conducted the primary analysis on the ITT population.

For the analysis of secondary endpoints of Patient Global Impression of Improvement (PGI-I) and Roland-Morris Disability Questionnaire (RMDQ-24) total score, the same ANCOVA model as in HMEP was used.

In order to adjust for multiple testing on the secondary endpoints, a serial gate-keeper multiple testing method was used, i.e., PGI-I and RMDQ-24 total score were tested sequentially only if the primary endpoint was statistically significant.

3.1.2.4 Results and Conclusions

A greater treatment effect was achieved by patients receiving DLX 60-120mg as compared to those receiving placebo (Table 8).

Table 8 Applicant's Primary Efficacy Analysis: HMEN (mITT)

LS Mean Change (SE) from Baseline to Week 13 in BPI 24-hour average pain	**Placebo (N=115)**	**DLX60-120mg (N=109)**	**P-value**
MMRM*	-1.5 (0.21)	-2.3 (0.22)	**0.004**
ANCOVA/BOCF**	-1.3 (0.20)	-1.9 (0.20)	**0.019**
ANCOVA/mBOCF**	-1.4 (0.21)	-1.9 (0.21)	**0.041**

*P-value calculated from MMRM model with terms for treatment, week, treatment*week, site, NSAID use, baseline, week*baseline.
**P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.
Note: mITT population excluded patients with no post-baseline observations.

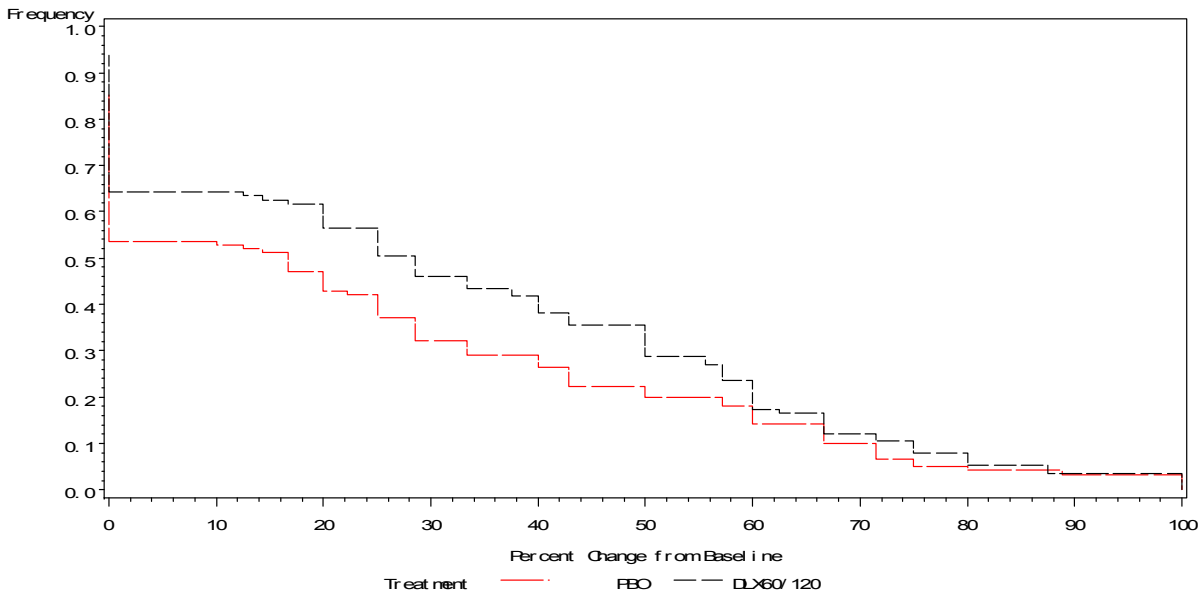
The applicant's primary analysis excluded 12 patients who had no post-baseline observations. I conducted the same analysis on the ITT set including those patients. My BOCF and mBOCF analyses and continuous responder analysis also demonstrated a statistically significant difference (Table 9 & Figure 4).

Table 9 Reviewer's Primary Efficacy Analysis: HMEN (ITT)

LS Mean Change (SE) from Baseline to Week 13 in BPI 24-hour average pain	**Placebo (N=121)**	**DLX60-120mg (N=115)**	**P-value**
ANCOVA/BOCF*	-1.2 (0.19)	-1.9 (0.19)	**0.009**
ANCOVA/mBOCF*	-1.2 (0.20)	-1.8 (0.20)	**0.020**

*P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.

Figure 4 Reviewer's Continuous Responder Analysis on Primary Efficacy Variable: HMEN (ITT)



Note: P-value of **0.018** is generated by van der Waerden test.

The secondary efficacy analyses on PGI-I and RMDQ-24 total score demonstrated a statistically significant difference (Tables 10 & 11).

Table 10 Applicant's Analysis of Patient Global Impression: HMEN

Patient Global Impression - Improvement	**Placebo (N=115)**	**DLX60-120mg (N=107)**
LS Means (SE)	3.2 (0.13)	2.8 (0.13)
p-value vs. Placebo*		**0.014**

*P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.

Table 11 Applicant's Analysis of RMDQ-24 Total Score: HMEN

RMDQ-24 total score Change from Baseline	**Placebo (N=105)**	**DLX60-120mg (N=99)**
LS Means (SE)	-1.9 (0.50)	-3.6 (0.51)
p-value vs. Placebo*		**0.009**

*P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.

In summary, the conservative analyses on the primary endpoint provided evidence of a treatment effect of duloxetine. The secondary efficacy analyses also demonstrated a statistically significant difference.

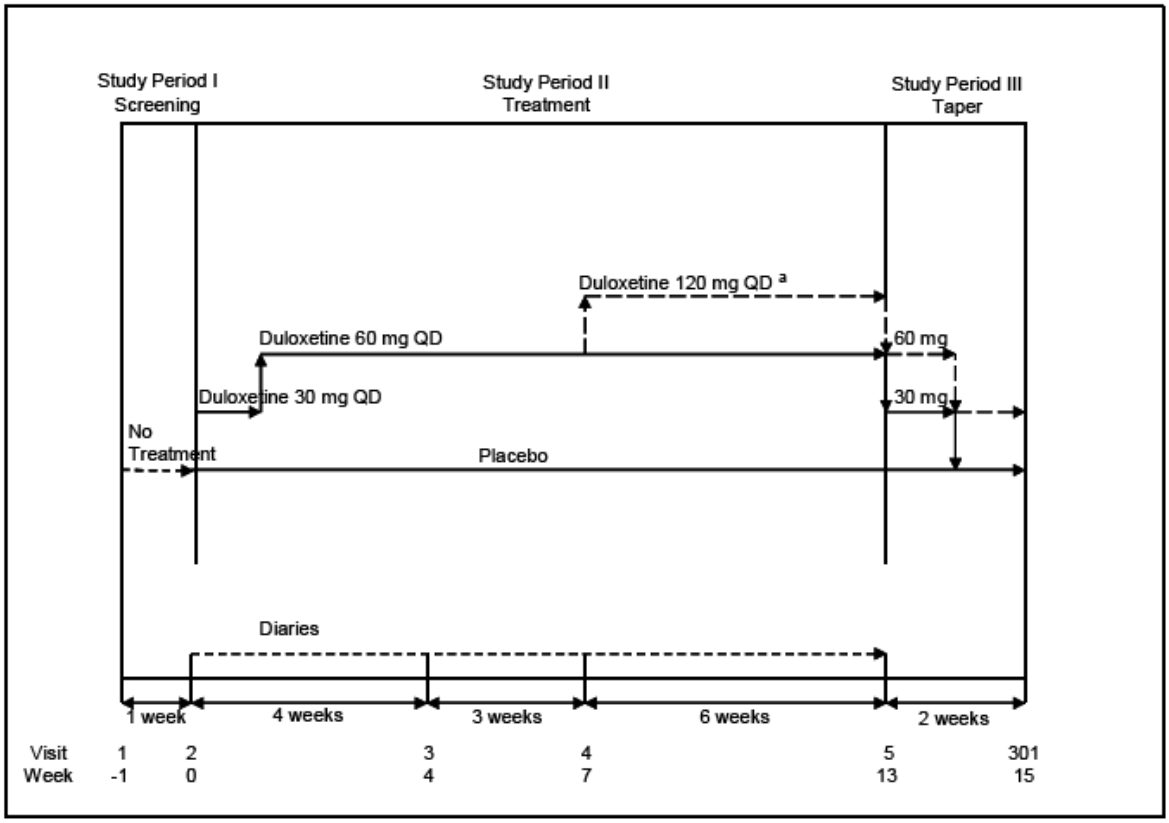
3.1.3 HMFG

3.1.3.1 Study Design and Endpoints

Study HMFG was a 13-week, double-blind, placebo-controlled, international, multi-center trial investigating the safety and efficacy of duloxetine in patients with OA knee pain. In

HMFG, 256 eligible patients were randomized in a 1:1 ratio to DLX 60 mg QD (n = 128) or placebo (n = 128) stratified by non-steroidal anti-inflammatory drug (NSAID) use at 21 centers in 5 countries worldwide. Patients were required to have a mean score of 4 or greater on the 24-hour average pain score (0-10). Figure 5 presents a schematic of the study design.

Figure 5 Schematic of Study Design: HMFG



a Non-responders at Visit 4 will titrate up to 120 mg QD

Similar to Study HMEN at Week 7, patients who did not meet response criteria, defined as at least 30% reduction in weekly mean of the BPI average score compared to baseline, had their dose increased to 120 mg QD.

Use of some short-acting analgesics, such as acetaminophen and codeine, was allowed for rescue from an OA knee pain flare after screening.

The primary efficacy measure was the BPI average pain score.

The secondary endpoints proposed for possible inclusion in the label were:

- Patient Global Impression of Improvement (PGI-I)
- Western Ontario and McMaster Universities (WOMAC) physical function subscale.

The other secondary efficacy variables included the following:

- Clinical Global Impression of Severity (CGI-S)
- The Hospital Anxiety and Depression Scale (HADS)
- Brief Pain Inventory (BPI)-Severity and Interference
- Percentage of patients who responded to treatment (with response defined as a 30% or 50% reduction of the BPI average pain score)
- BDI-II score
- HADS-A score.

3.1.3.2 Disposition and Demographics

Approximately 20% of the patients discontinued before the end of study (Table 12). However, more patients from the DLX group discontinued compared to placebo group. Twenty-seven percent of DLX patients discontinued while 13% of placebo patients discontinued. As expected, majority of DLX dropouts were due to adverse events. Nineteen percent of DLX patients discontinued due to adverse events. However, unexpectedly, the majority of placebo dropouts were also due to adverse events. Five percent of placebo patients discontinued due to adverse events and 4% of placebo patients discontinued due to lack of efficacy.

Table 12 Subject Disposition: HMFG

	Number of Patients	
	Placebo	**DLX 60-120mg**
Randomized	128 (100%)	128 (100%)
ITT	128	128
mITT	127	121
Completed	111 (87%)	93 (73%)
Reasons for dropout		
AE	7 (5%)	24 (19%)
LOE	5 (4%)	1 (1%)
Subject Decision	2 (2%)	4 (3%)
Other	3 (2%)	6 (4%)

Patient demographics are presented by treatment groups in the appendix (Table 23). There were no noticeable imbalances among treatment groups with respect to demographic variables of age, race, sex, and weight.

Table 23 also shows baseline values for the efficacy variable of BPI average pain score by treatment groups. Distributions of the efficacy variable at baseline were comparable between treatment groups.

3.1.3.3 Statistical Methodologies

The statistical methods used in HMFG were identical to those in HMEP.

3.1.3.4 Results and Conclusions

A greater treatment effect was achieved by patients receiving DLX 60-120mg as compared to those receiving placebo (Table 13).

Table 13 Applicant's Primary Efficacy Analysis: HMFG (mITT)

LS Mean Change (SE) from Baseline to Week 13 in BPI 24-hour average pain	**Placebo (N=127)**	**DLX60-120mg (N=121)**	**P-value**
MMRM*	-1.9 (0.18)	-2.7 (0.20)	**<0.001**
ANCOVA/BOCF**	-1.6 (0.19)	-2.2 (0.20)	**0.013**
ANCOVA/mBOCF**	-1.6 (0.19)	-2.3 (0.20)	**0.005**

*P-value calculated from MMRM model with terms for treatment, week, treatment*week, site, NSAID use, baseline, week*baseline.
**P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.
Note: mITT population excluded patients with no post-baseline observations.

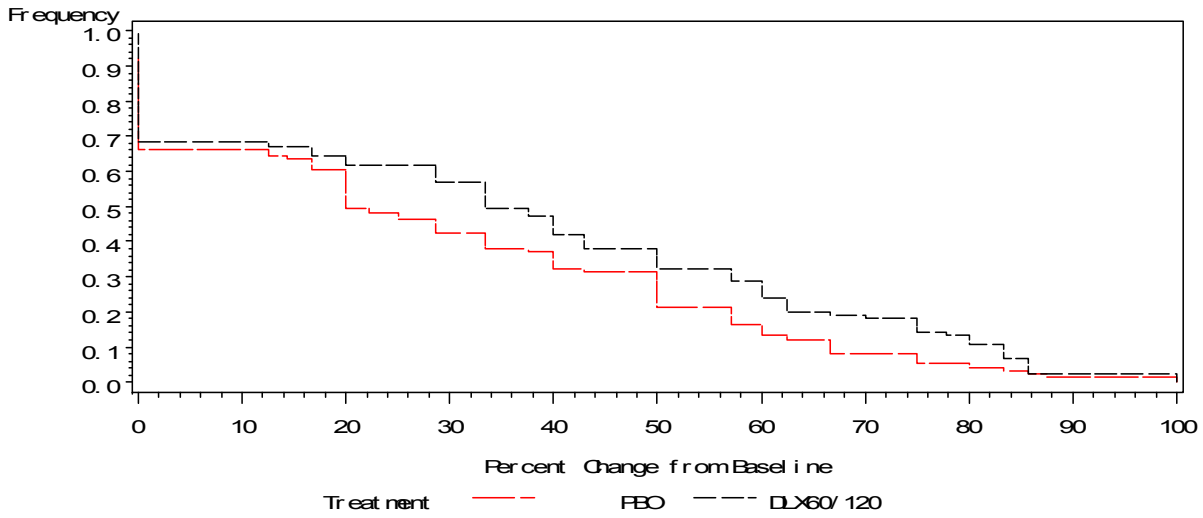
The applicant's primary analysis excluded eight patients who had no post-baseline observations. I conducted the same analysis on the ITT population including those patients. My BOCF and mBOCF analyses and continuous responder analysis on the ITT set also demonstrated a statistically significant difference. (Table 14 & Figure 6).

Table 14 Reviewer's Primary Efficacy Analysis: HMFG (ITT)

LS Mean Change (SE) from Baseline to Week 13 in BPI 24-hour average pain	Placebo (N=128)	DLX60-120mg (N=128)	P-value
ANCOVA/BOCF*	-1.6 (0.19)	-2.2 (0.20)	**0.013**
ANCOVA/mBOCF*	-1.6 (0.18)	-2.3 (0.19)	**0.005**

*P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.

Figure 6 Reviewer's Continuous Responder Analysis on Primary Efficacy Variable: HMFG (ITT)



Note: P-value of **0.016** is generated by van der Waerden test.

The secondary efficacy analysis on PGI-I failed to demonstrate a statistically significant difference (Tables 15 & 16). Since the analysis on PGI-I failed, the sequential test procedure stopped and the next analysis on WOMAC physical function should not be considered.

Table 15 Applicant's Analysis of Patient Global Impression: HMFG

Patient Global Impression - Improvement	**Placebo (N=127)**	**DLX60-120mg (N=123)**
LS Means (SE)	3.1 (0.12)	2.9 (0.12)
p-value vs. Placebo*		0.164

*P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.

Table 16 Applicant's Analysis of WOMAC Physical Function: HMFG

WOMAC Physical Function Change from Baseline	**Placebo (N=126)**	**DLX60-120mg (N=118)**
LS Means (SE)	-9.4 (1.08)	-12.69 (1.15)
p-value vs. Placebo*		**(0.016)**

*P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.
Note: Because the first gate-keeper analysis failed, the p-value reported here is nominal.

In summary, the conservative analyses on the primary endpoint provided evidence of a treatment effect of duloxetine. The secondary efficacy analysis on PGI-I did not demonstrate a statistically significant difference. Since the analysis on PGI-I failed, the sequential test procedure stopped and the analysis on WOMAC physical function was not be considered.

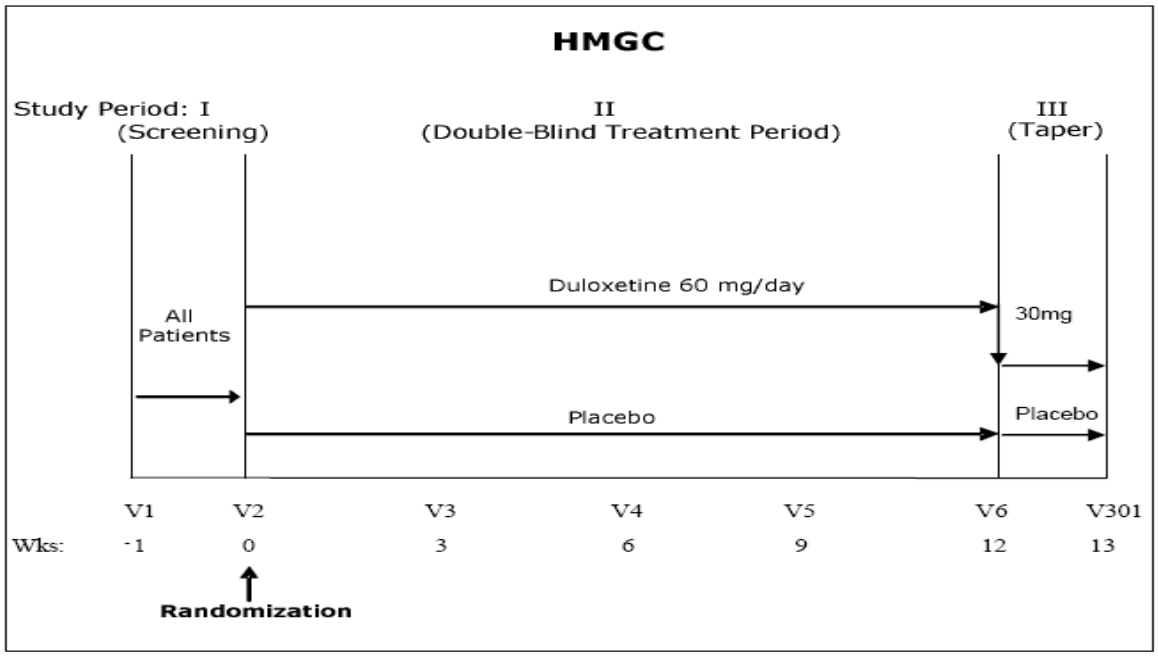
3.1.4 HMGC

3.1.4.1 Study Design and Endpoints

Study HMGC was a 12-week, double-blind, placebo-controlled, international, multi-center trial investigating the safety and efficacy of duloxetine in patients with CLBP. In HMGC, 401 eligible patients were randomized in a 1:1 ratio to DLX 60 mg QD (n = 198) or placebo

(n = 203) stratified by non-steroidal anti-inflammatory drug (NSAID) use at 27 centers in 6 countries worldwide. Patients were required to have a rating of 4 or greater on the BPI average pain score. Figure 7 presents a schematic of the study design.

Figure 7 Schematic of Study Design: HMGC



Use of some short-acting analgesics was allowed for management of breakthrough CLBP (rescue therapy) after screening.

Similar to Studies HMEN and HMFG, the primary efficacy measure was the BPI average pain score.

The secondary endpoints proposed for possible inclusion in the label were:

- Patient Global Impression of Improvement (PGI-I).
- Roland-Morris Disability Questionnaire (RMDQ-24) total score.

The other secondary efficacy variables included the following:

- Clinical Global Impression of Severity (CGI-S)
- Profile of Mood States –Brief Form (POMS-Brief Form)
- 36-item Short-Form Health Survey (SF-36)
- European Quality of Life Questionnaire- 5 Dimension (EQ-5D)
- Work Productivity and Activity Impairment Instrument (WPAI).

3.1.4.2 Disposition and Demographics

Approximately 24% of the patients discontinued before the end of study (Table 17). However, more patients from the DLX group discontinued compared to placebo group. Twenty-six percent of DLX patients discontinued while 23% of placebo patients discontinued. As expected, the majority of DLX dropouts were due to adverse events. Fifteen percent of DLX patients discontinued due to adverse events. However, unexpectedly, the majority of placebo dropouts were not due to lack of efficacy, but due to subject decision. Five percent of placebo patients discontinued due to adverse events and 4% of placebo patients discontinued due to lack of efficacy.

Table 17 Subject Disposition: HMGC

	Number of Patients	
	Placebo	**DLX 60mg**
Randomized	203 (100%)	198 (100%)
ITT	203	198
mITT	199	195
Completed	156 (77%)	147 (74%)
Reasons for dropout		
AE	11 (5%)	30 (15%)
LOE	9 (4%)	1 (1%)
Subject Decision	13 (7%)	8 (4%)
Protocol Violation	5 (3%)	6 (3%)
Other	9 (4%)	6 (3%)

Patient demographics are presented by treatment groups in the appendix (Table 23). There were no noticeable imbalances among treatment groups with respect to demographic variables of age, race, sex, and weight.

Table 23 also shows baseline values for the efficacy variable of BPI average pain score by treatment groups. Distributions of the efficacy variable at baseline were comparable between treatment groups.

3.1.4.3 Statistical Methodologies

The statistical methods were identical to those in HMEN.

3.1.4.4 Results and Conclusions

A greater treatment effect was achieved by patients receiving DLX 60mg as compared to those receiving placebo (Table 18).

Table 18 Applicant's Primary Efficacy Analysis: HMGC (mITT)

LS Mean Change (SE) from Baseline to Week 12 in BPI 24-hour average pain	**Placebo (N=203)**	**DLX60mg (N=198)**	**P-value**
MMRM*	-1.9 (0.15)	-2.5 (0.16)	**0.001**
ANCOVA/BOCF**	-1.4 (0.15)	-1.9 (0.15)	**0.004**
ANCOVA/mBOCF**	-1.6 (0.15)	-2.1 (0.15)	**0.004**

*P-value calculated from MMRM model with terms for treatment, week, treatment*week, site, baseline, week*baseline.
**P-value calculated from ANCOVA model with terms for treatment, site, and baseline score as covariate.
Note: mITT population excluded patients with no post-baseline observations.

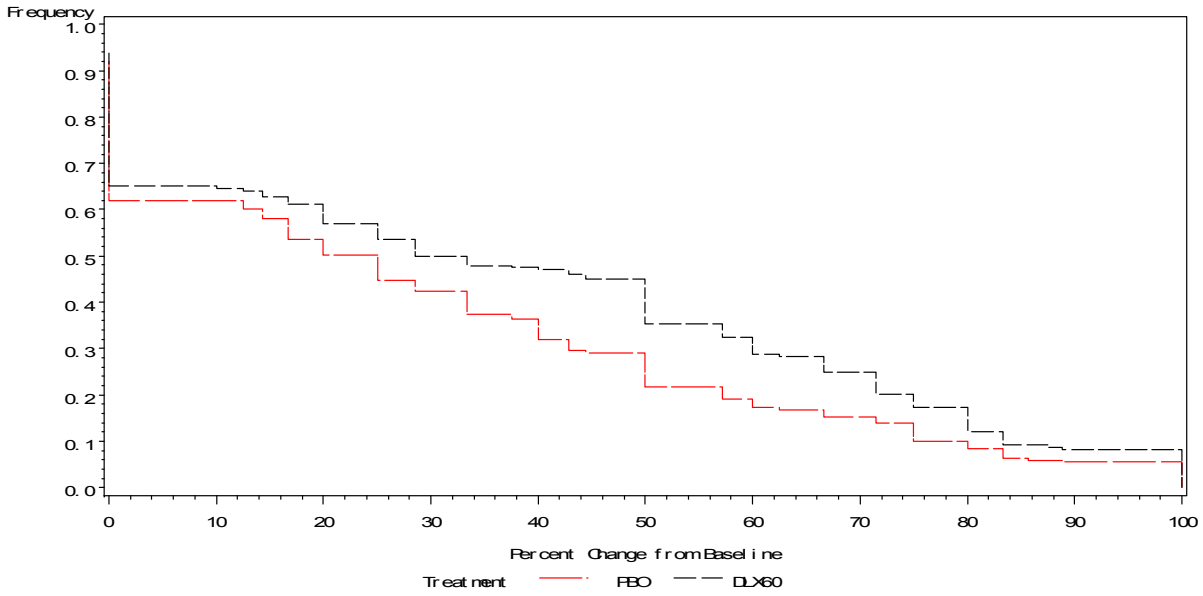
The applicant's primary analysis excluded seven patients who had no post-baseline observations. I conducted the same analysis on the ITT set including those patients. My BOCF and mBOCF analyses and continuous responder analysis on the ITT set also demonstrated a statistically significant difference (Table 19 & Figure 8).

Table 19 Reviewer's Primary Efficacy Analysis: HMGC (ITT)

LS Mean Change (SE) from Baseline to Week 12 in BPI 24-hour average pain	**Placebo (N=203)**	**DLX60mg (N=198)**	**P-value**
ANCOVA/BOCF*	-1.5 (0.15)	-2.0 (0.15)	**0.004**
ANCOVA/mBOCF*	-1.8 (0.18)	-2.6 (0.18)	**<0.001**

*P-value calculated from ANCOVA model with terms for treatment, site, and baseline score as covariate.

Figure 8 Reviewer's Continuous Responder Analysis on Primary Efficacy Variable: HMGC (ITT)



Note: P-value of **0.024** is generated by van der Waerden test.

The secondary efficacy analysis on PGI-I demonstrated a statistically significant difference while the secondary efficacy analysis on RMDQ-24 failed to demonstrate a statistically significant difference (Tables 20 & 21).

Table 20 Applicant's Analysis of Patient Global Impression: HMGC

Patient Global Impression - Improvement	**Placebo (N=199)**	**DLX60mg (N=194)**
LS Means (SE)	3.2 (0.09)	2.9 (0.09)
p-value vs. Placebo*		**0.011**

*P-value calculated from ANCOVA model with terms for treatment, site, and baseline score as covariate.

Table 21 Applicant's Analysis of RMDQ-24 Total Score: HMGC

RMDQ-24 total score Change from Baseline	Placebo (N=179)	DLX60mg (N=178)
LS Means (SE)	-2.2 (0.32)	-2.7 (0.31)
p-value vs. Placebo*		0.255

*P-value calculated from ANCOVA model with terms for treatment, site, and baseline score as covariate.

In summary, the BOCF analysis and continuous responder analyses conducted on the ITT population demonstrated a statistically significant difference. The secondary efficacy analysis on PGI-I demonstrated a statistically significant difference, while the analysis on RMDQ-24 failed.

3.1.5 Clinical Concern regarding Safety of Duloxetine 120mg Dose

Because there was a safety concern of hepatotoxicity on the DLX 120 mg dose and the dose was not approved for other pain indications such as DPNP and FM, the clinical review team posed the question, 'Is DLX 60mg effective?' However, studies included at the time of the NDA submission did not compare DLX 60mg with placebo directly. All studies employed study designs comparing combined doses of DLX 60mg and 120mg with placebo. Study HMGC submitted with the 120-day safety update employed a study design comparing DLX 60mg directly with placebo.

To investigate the effectiveness of DLX 60mg dose in studies HMEP, HMEN, and HMFG, post-hoc analyses were conducted. First, the pain changes at Week 13 were compared between the DLX 60mg group and the placebo group from studies HMEP, HMEN and HMFG. When comparing DLX 60mg with placebo in HMEP, only patients re-randomized to DLX 60mg and patients initially randomized to placebo were compared. The analysis did not demonstrate a significant difference. When comparing DLX 60mg with placebo in HMEN and HMFG, the non-responders at Week 7 were treated as failures regardless of randomized treatment group. The results of the analyses demonstrated a significant difference between DLX 60mg and placebo at Week 13 (Table 27). Second, I compared DLX 60mg with placebo in terms of BPI average pain change from baseline to Week 7 in studies HMEN and HMFG. A statistically significant difference between DLX 60mg and placebo was demonstrated.

In conclusion, OA study HMEP failed to demonstrate a difference in pain when comparing DLX 60 mg with placebo. The OA study HMFG yielded significant differences between DLX 60 mg and placebo both in 7-week and 13-week analyses. Two CLBP studies, HMEN and HMGC, demonstrated statistically significant differences when comparing the DLX 60 mg dose to placebo. The evidence of efficacy of the DLX 60 mg dose was apparent in both 7-week and 13-week analyses in the two CLBP studies.

3.2 Evaluation of Safety

The evaluation of safety was conducted by the clinical reviewer, Anjelina Pokrovnichka, M.D.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

I explored the heterogeneity of the treatment effect across age group, race group, and sex by inclusion of interaction terms in the ANCOVA model. In the analyses of primary efficacy variables, there were no statistically significant interactions between treatment and age group ('<55 yr.' or '≥55 yr.' for HMEN and HMGC and '<65 yr.' or '≥65 yr.' for HMFG), sex, or race group ('Caucasian' or 'Other'). I conducted subgroup analyses for studies HMEN, HMFG, and HMGC and results can be found in appendix (Tables 24 –26).

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

5.1.1 Statistical Issues

There were several issues in the statistical analyses of the data. In chronic pain trials, patients who withdraw before the end of the study should be treated as non-responders, and no benefit should be assigned based on the pain scores before dropout. However, the applicant proposed the mixed effect repeated measures (MMRM) analysis as the primary analysis method in all studies. The method uses pain data from patients who withdraw before the study ends. Also in order for the MMRM to be valid, missing at random (MAR) should be assumed as the mechanism generating missing data. However in chronic pain trials, missing data is often informative and therefore the MAR assumption is not supported.

Second, because of the safety concern of hepatotoxicity regarding duloxetine 120 mg, the clinical team focused on the efficacy of duloxetine 60 mg. However with the exception of Study HMGC, the studies were designed to evaluate the combined the 60 mg and 120 mg

doses of duloxetine. HMEP re-randomized patients initially randomized to duloxetine 60 mg to duloxetine 60 mg or 120 mg at Week 7. HMEP did not demonstrate a significant difference when I compared patients treated with duloxetine 60 mg over 13 weeks with patients treated with placebo. HMEN and HMFG increased the dose of duloxetine from 60 mg to 120 mg for patients who were randomized to duloxetine 60 mg but did not show at least a 30% reduction of pain from baseline at Week 7. In analyses submitted in response to information requests, the applicant treated patients who were not showing at least a 30% pain reduction from baseline at Week 7 as failures regardless of randomized treatment group, and the baseline scores were imputed to those patients. The analyses demonstrated a significant difference between duloxetine 60 mg and placebo. I additionally conducted analyses of the change from baseline up to Week 7. In my analyses, HMEN and HMFG demonstrated a significant difference between duloxetine 60 mg and placebo.

Although not statistical, another issue that is noteworthy is the fact that the applicant seeks a more broad chronic pain indication. The requirement for a broad chronic pain indication is a current topic under consideration within the Agency as well as the clinical community.

5.1.2 Collective Evidence

In reviewing the collective evidence from the applicant's primary and sensitivity analyses as well as my additional analyses, I conclude that the data from studies HMEN and HMGC provide evidence of the efficacy of duloxetine 60 mg once daily for treating chronic low back pain. Data from two OA studies, HMEP and HMFG, are not sufficient to provide evidence of the efficacy of duloxetine 60 mg once daily because the studies with combined dose did not plan to provide evidence of effects of the 60 mg dose separated from the 120 mg dose although my and applicant's post-hoc analyses on effects of the 60mg dose demonstrated significance in HMFG. The following table summarizes results from the four studies.

Table 22 Summary of Primary Efficacy Analyses

	HMEN (CLBP)		HMEP (OA)	HMFG (OA)		HMGC (CLBP)
MMRM	P<0.05		P<0.05	P<0.05		
ANCOVA/BOCF	P<0.05		NS	P<0.05		
ANCOVA/mBOCF	P<0.05		P<0.05	P<0.05		
CRA/vdW*	P<0.05		NS	P<0.05		
Focusing on DLX60mg**		**Focusing on DLX60mg up to week 7**			**Focusing on DLX60mg up to week 7**	
MMRM	P<0.05		NS			P<0.05
ANCOVA/BOCF	P<0.05	P<0.05	NS	P<0.05	P<0.05	P<0.05
ANCOVA/mBOCF	P<0.05		NS	P<0.05		P<0.05
CRA/vdW*						P<0.05

*vdW stands for van der Waerden test comparing two cumulative responder curves.
** Post-hoc analyses on duloxetine 60 mg dose in HMEN, HMFG, and HMGC were conducted on ITT population and analysis in HMEP was conducted on the sub-population with patients re-randomized to duloxetine 60mg and placebo patients. Patients who did not show at least 30% pain reduction at Week 7 regardless of randomized treatment group were treated as failures.

5.2 Conclusions and Recommendations

The applicant has evaluated the use of duloxetine 60 – 120 mg for the treatment of chronic pain. Based on my review of the data, I find that patients receiving duloxetine experienced greater pain reduction compared to patients receiving placebo. The treatment effect was evident in a single trial conducted in osteoarthritis (OA) patients as well as two trials conducted in patients with chronic low back pain (CLBP). Due to concerns about hepatotoxicity regarding the 120 mg dose, additional analyses focusing on the 60 mg dose were conducted. I conclude that the 60 mg dose was effective in reducing pain for the 12-week trial duration in a single CLBP study. Supportive evidence of this effect was also derived from my analysis of pain reductions up to 7 week and applicant's post-hoc analysis of 13-week pain reductions treating non-responders at Week 7 as treatment failures regardless of randomized treatment group in a second CLBP study. However, I conclude that there was no sufficient data to support the efficacy of duloxetine 60 mg in treating OA because the successful OA study with combined dose did not plan to provide evidence of effects of the 60 mg dose separated from the 120 mg dose although my and applicant's post-hoc analyses on effects of the 60 mg dose demonstrated significance in the study.

5.3 Review of Clinical Studies of Proposed Label

The following is the portion of the Clinical Study section from the proposed label with the results of OA and CLBP studies data analyses. I have included several comments throughout Section 14.3. The same comments apply to Section 14.4.

(b) (4)

Reviewer comment 1: I recommend rounding of the mean baseline pain score.

(b) (4)

Reviewer comment 2: The claims in the preceding paragraph are all based on secondary efficacy outcome variables. Unless these variables provide information deemed necessary by the clinical review team, I recommend the claims be deleted.

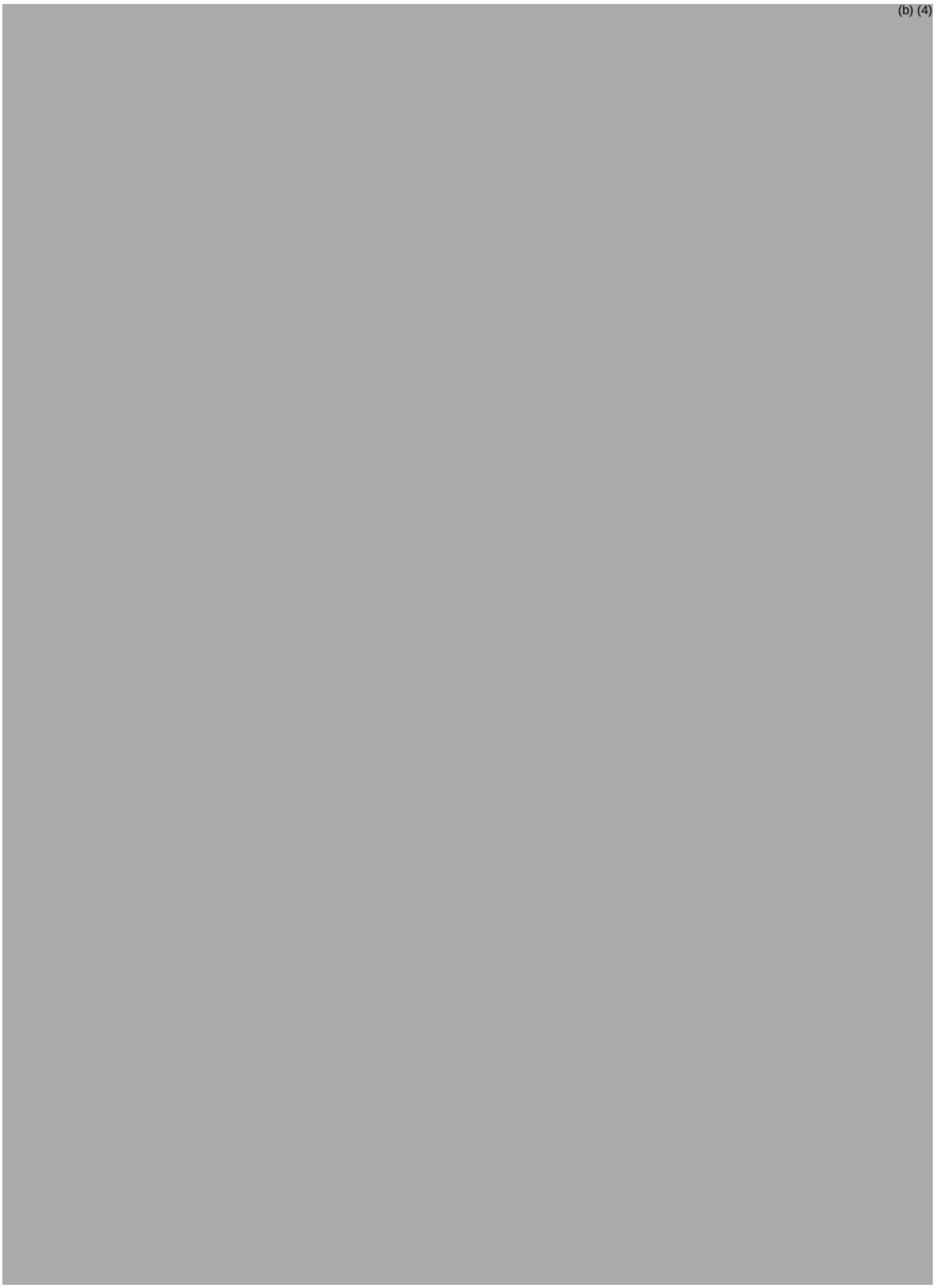
(b) (4)

Reviewer comment 3 I am unaware of the analysis used to support this statement. The statement should not be allowed if it is based on an analysis of means.

1 Page of Draft Labeling has been Withheld in Full as B4 (CCI/TS) Immediately Following this Page

Reviewer comment 4 The time course graph may be misleading as it conveys information based on means instead of individual responses.

(b) (4)



2 Pages of Draft Labeling have been Withheld in Full as b4 (CCI/TS) immediately following this page.

APPENDIX

Table 23 Patient Demographic and Baseline Characteristics

Study HMEP:

	Placebo (n=120)	DLX 60-120mg (n=111)
Gender n (%)		
Female	81 (68%)	70 (63%)
Male	39 (32%)	41 (37%)
Race n (%)		
African	6 (5%)	6 (5%)
Caucasian	100 (83%)	94 (85%)
East Asian	3 (3%)	0 (0%)
Other	11 (9%)	11 (10%)
Age (years)		
Median	63	62
Range	44 – 87	40 – 82
Weight (kg)		
Median	84	84
Range	53 – 127	50 – 129
BPI Average Pain		
Median	6	6
Range	2 – 10	3 – 10

Study HMEN:

	Placebo (n=121)	DLX 60-120mg (n=115)
Gender n (%)		
Female	73 (60%)	71 (62%)
Male	48 (40%)	44 (38%)
Race n (%)		
African	6 (5%)	6 (5%)
Caucasian	91 (75%)	85 (74%)
East Asian	2 (2%)	0 (0%)
Other	22 (18%)	24 (21%)
Age (years)		

Median	50	50
Range	21 – 80	20 – 85
Weight (kg)		
Median	73	76
Range	42 – 115	45 – 120
BPI Average Pain		
Median	6	6
Range	2 – 10	2 – 10

Study HMFG:

	Placebo (n=128)	**DLX 60-120mg (n=128)**
Gender n (%)		
Female	107 (84%)	89 (70%)
Male	21 (16%)	39 (30%)
Race n (%)		
African	3 (2%)	0 (0%)
Caucasian	124 (97%)	126 (98%)
East Asian	1 (1%)	0 (0%)
Other	0 (0%)	2 (2%)
Age (years)		
Median	62	63
Range	40 – 79	44 – 79
Weight (kg)		
Median	80	80
Range	55 – 110	50 – 116
BPI Average Pain		
Median	6	6
Range	4 – 9	3 – 10

Study HMGC:

	Placebo (n=203)	DLX 60mg (n=198)
Gender n (%)		
Female	128 (63%)	118 (60%)
Male	75 (37%)	80 (40%)
Race n (%)		
African	5 (2%)	5 (3%)
Caucasian	193 (95%)	189 (96%)
Hispanic	4 (2%)	4 (2%)
Other	1 (1%)	0 (0%)
Age (years)		
Median	55	57
Range	19 – 89	19 – 80
Weight (kg)		
Median	79	77
Range	47 – 114	44 – 128
BPI Average Pain		
Median	6	6
Range	4 – 10	4 – 10

Table 24 Subgroup Analyses on Primary Efficacy Endpoint: HMEN

LS Mean Change (SE) from Baseline to Week 12 in BPI 24-hour average pain*	Placebo (N=121)	DLX60-120mg (N=115)
Caucasian	-1.2 (0.22)	-1.9 (0.23)
Non-Caucasian	-0.9 (0.51)	-1.4 (0.59)
Age <55	-1.2 (0.26)	-2.1 (0.29)
Age >=55	-1.1 (0.34)	-1.3 (0.33)
Female	-1.3 (0.25)	-1.8 (0.27)
Male	-1.2 (0.31)	-2.2 (0.35)

*LSMeans calculated from ANCOVA/BOCF model with terms for treatment, site, NSAID use, and baseline score as covariate.

Table 25 Subgroup Analyses on Primary Efficacy Endpoint: HMFG

LS Mean Change (SE) from Baseline to Week 12 in BPI 24-hour average pain*	Placebo (N=128)	DLX60-120mg (N=128)
Caucasian	-1.7 (0.19)	-2.2 (0.20)
Non-Caucasian	NE	NE
Age <65	-1.5 (0.24)	-2.0 (0.27)
Age >=65	-1.7 (0.32)	-2.4 (0.33)
Female	-1.6 (0.22)	-2.3 (0.26)
Male	-1.7 (0.50)	-2.0 (0.43)

Note: NE stands for non-estimable.
*LSMeans calculated from ANCOVA/BOCF model with terms for treatment, site, NSAID use, and baseline score as covariate.

Table 26 Subgroup Analyses on Primary Efficacy Endpoint: HMGC

LS Mean Change (SE) from Baseline to Week 12 in BPI 24-hour average pain*	Placebo (N=203)	DLX60mg (N=198)
Caucasian	-1.5 (0.15)	-2.0 (0.15)
Non-Caucasian	-0.9 (0.80)	-3.8 (0.93)
Age <55	-1.5 (0.21)	-2.0 (0.22)
Age >=55	-1.4 (0.23)	-2.0 (0.23)
Female	-1.3 (0.21)	-2.0 (0.22)
Male	-1.4 (0.26)	-2.0 (0.25)

*LSMeans calculated from ANCOVA/BOCF model with terms for treatment, site, and baseline score as covariate.

Table 27 Applicant's Post-Hoc Analyses Investigating Duloxetine 60 mg versus Placebo

Study	Analysis Method	Treatment	LSMean Change (SE)	p-value
HMEN	BOCF	DLX60	-1.5 (0.19)	0.009
		Placebo	-0.9 (0.18)	
HMFG	BOCF	DLX60	-1.8 (0.20)	0.012
		Placebo	-1.2 (0.19)	
HMEN	BOCF	DLX60	-1.8 (0.21)	0.005
		Placebo	-1.1 (0.20)	
HMFG	BOCF	DLX60	-2.1 (0.20)	0.002
		Placebo	-1.3 (0.18)	

*LSMeans on BPI 24-hour average pain calculated from ANCOVA model with terms for treatment, site, and baseline score as covariate. (Table 6.1 in response to information request submitted 8/14/2009)

SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: Yongman Kim, Ph.D.
Mathematical Statistician
Date: February 10, 2010

Concurring Reviewer: Dionne Price, Ph.D.
Team Leader

cc:

DAARP/Ayanna Augustus
DAARP/Anjelina Pokrovnichka, M.D.
DAARP/Ellen Fields, M.D.
HFD-715/Yongman Kim, Ph.D.
HFD-715/Dionne Price, Ph.D.
HFD-715/Thomas Permutt, Ph.D.
HFD-700/Edward Nevius, Ph.D.
HFD-700/Ram Tiwari, Ph.D.
HFD-700/Lillian Patrician

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-22516	ORIG-1	ELI LILLY AND CO	CYMBALTA

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

YONGMAN KIM
02/17/2010

DIONNE L PRICE
02/19/2010
See my review

THOMAS J PERMUTT
02/23/2010
I concur with Dr. Price's secondary review.

DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
OFFICE OF BIOSTATISTICS

Statistical Review and Evaluation

SECONDARY REVIEW

NDA:	22-516
Name of drug:	Cymbalta (duloxetine)
Applicant:	Eli Lilly and Company
Indication:	Management of chronic pain
Dates:	Submitted May 15, 2009; PDUFA March 15, 2010
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II
Statistical reviewer:	Yongman Kim, Ph.D.
Secondary reviewer:	Dionne Price, Ph.D.
Medical Division:	Division of Anesthesia, Analgesia, and Rheumatology Products
Clinical Team:	Anjelina Pokrovnichka, M.D.
	Ellen Fields, M.D.
Project Manager:	Ayanna Augustus, Ph.D.

Keywords: clinical studies, NDA review, missing data

1 BACKGROUND

Eli Lilly and Company proposes Cymbalta (duloxetine) for the treatment of chronic pain including management of diabetic neuropathic pain, fibromyalgia, and chronic pain due to osteoarthritis (OA) and chronic low back pain (CLBP). The current supplemental application includes studies conducted in the OA and CLBP populations. The primary statistical review was conducted by Dr. Yongman Kim. Dr. Kim's assessment of the efficacy of Cymbalta focused on four studies, two studies conducted in the OA population and two studies conducted in the CLBP population. Dr. Kim concluded that the clinical studies demonstrated the efficacy of Cymbalta 60 – 120 mg. There was concern regarding the hepatotoxicity of the 120 mg dose; therefore, the review team investigated the effect of the 60 mg dose. The following excerpt from Dr. Kim's review summarizes his conclusions regarding the 60 mg dose.

> I conclude that the 60 mg dose was effective in reducing pain for the 12-week trial duration in a single CLBP study. Supportive evidence of this effect was also derived from my analysis of pain reductions up to 7 week and applicant's post-hoc analysis of 13-week pain reductions treating non-responders at Week 7 as treatment failures regardless of randomized treatment group in a second CLBP study. However, I conclude that there was no sufficient data to support the efficacy of duloxetine 60 mg in treating OA because the successful OA study with combined dose did not plan to provide evidence of effects of the 60 mg dose separated from the 120 mg dose although my and applicant's post-hoc analyses on effects of the 60 mg dose demonstrated significance in the study.

I agree that the efficacy of Cymbalta 60 – 120 mg has been demonstrated. However, my overall conclusions regarding the efficacy of the 60 mg dose slightly differ from those of Dr. Kim.

2 REVIEW

The applicant submitted five studies. Of the five studies, only two were designed to evaluate the 60 mg dose of Cymbalta. Study HMGC was a fixed-dose study evaluating the efficacy and safety of Cymbalta 60 mg in chronic low back pain patients. The study successfully demonstrated the efficacy of Cymbalta 60 mg. Study HMEO was also a fixed-dose study evaluating several doses of Cymbalta including 60 mg in patients with chronic low back pain. The study failed to demonstrate the efficacy of Cymbalta 60 mg. The results and conclusions from these two studies are not in question. The applicant submitted three studies that evaluated flexible doses of Cymbalta 60 mg and 120 mg. I am in general agreement with Dr. Kim's assessment of the pre-specified primary analyses of these studies which evaluated Cymbalta 60 mg – 120 mg. Consequently, my review will primarily focus on analyses conducted to elucidate the effect of the 60 mg dose. The

reader is referred to the statistical review of Dr. Kim for complete details regarding all studies.

2.1 STUDIES HMEN AND HMFG

Studies HMEN and HMFG were 13-week randomized, double-blind, placebo-controlled studies conducted in patients with chronic low back pain and osteoarthritis, respectively. Patients were initially randomized to Cymbalta 60 mg or placebo. At Week 7, the dose of Cymbalta was escalated to 120 mg for patients who were not experiencing at least a 30% reduction in pain. The primary analysis for both studies evaluated the difference in mean pain at Week 13 for patients receiving Cymbalta compared to patients receiving placebo. A mixed-effects model repeated measures analysis was employed. The analysis used available data for patients withdrawing prior to the end of the study. This analysis was deemed inappropriate by Dr. Kim. Specifically in chronic pain trials, 12 weeks of treatment is a surrogate for years of treatment; therefore, a drug is considered ineffective if patients cannot continue it for the study duration. This clinical logic motivates the need for analyses which assign little or no benefit to patients withdrawing prior to completing the study. Thus, Dr. Kim appropriately focused on an analysis of covariance model using conservative imputation strategies. The division recognized the concern regarding the 120 mg dose early in the review cycle and consequently requested additional analyses to evaluate the effect of 60 mg in a July 17, 2009 filing communication. In response, the applicant conducted analyses whereby all non-responders at Week 7, regardless of the treatment assignment, were treated as discontinuations due to lack of efficacy. Lilly conducted the analyses using the baseline observation carried forward (BOCF) and modified BOCF strategies. The BOCF strategy imputed the baseline score for all discontinuations. The modified BOCF strategy used BOCF for dropouts due to adverse events and the last observation carried forward for all other dropouts. The results of the applicant's analyses are depicted in Table 1.

Table 1: Mean Change Analysis of BPI Average Pain Rating (Duloxetine 60 mg QD versus placebo)

Study	Analysis Method	Treatment	LSMean Change (SE)	P -value
HMEN	BOCF	DLX60 QD	− 1.53 (0.19)	0.009
		Placebo	− 0.90 (0.18)	
HMFG	BOCF	DLX60 QD	− 1.76 (0.20)	0.012
		Placebo	− 1.17 (0.19)	
HMEN	Modified BOCF	DLX60 QD	− 1.83 (0.21)	0.005
		Placebo	− 1.07 (0.20)	
HMFG	Modified BOCF	DLX60 QD	− 2.05 (0.20)	0.002
		Placebo	− 1.33 (0.18)	

Abbreviations: BOCF = baseline-observation-carried-forward; DLX 60 QD = duloxetine 60 mg once daily; LSMean = least-squares mean; SE = standard error.
Source: Applicant's Table 6 - Response to Filing Communication

Although the analyses submitted in response to the division's request were reasonable, studies HMEN and HMFG were not designed for a 13-week analysis of the 60 mg dose only. Thus to further assess the effect, Dr. Kim conducted an analysis of data collected up to Week 7. Since all patients randomized to Cymbalta received the 60 mg dose prior to Week 7, this analysis appeared to be more consistent with the design of the study. His analyses suggested evidence of a treatment effect up to Week 7 for both studies.

2.2 STUDY HMEP

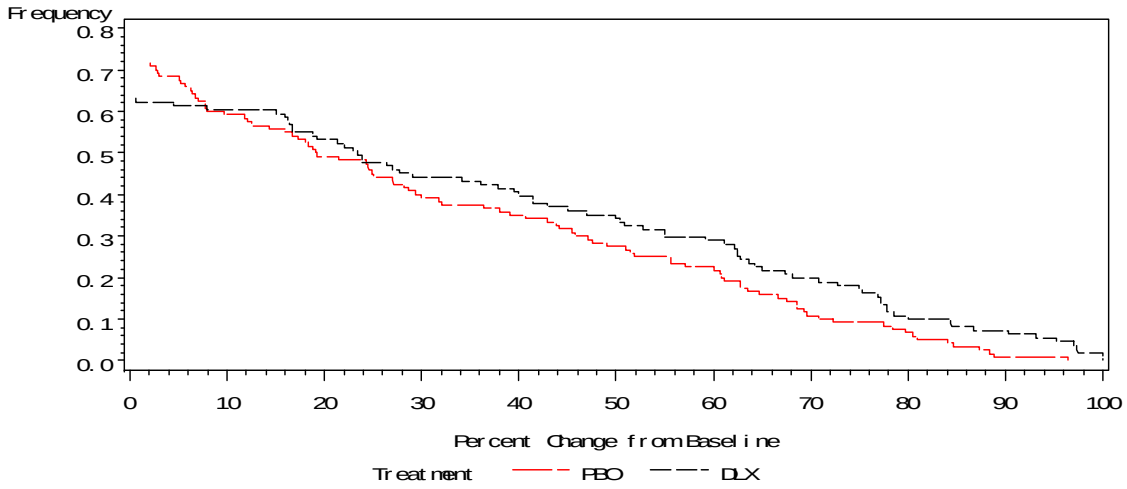
Study HMEP was a randomized, double-blind, placebo-controlled study conducted in osteoarthritis patients. Patients were randomized to Cymbalta 60 mg or placebo. At Week 7, patients in the Cymbalta arm were re-randomized to Cymbalta 60 mg or Cymbalta 120 mg. The primary analysis mimicked that of studies HMEN and HMFG and compared two groups, Cymbalta-treated patients to placebo-treated patients. Dr. Kim performed analyses using the BOCF and mBOCF imputation strategies. His results were sensitive to the procedure used to handle missing data. The BOCF strategy resulted in a mean difference between groups that was not statistically significant. In contrast, the mBOCF strategy yielded results that were significantly different. Dr. Kim examined the reasons for dropouts (see Table 1) and concluded that the mBOCF strategy assigned good scores to patients that dropped out for reasons other than lack of efficacy or adverse events. The assignment of good scores in this scenario is concerning since reasons for dropouts reported as "other" and "subject decision" may have masked adverse events. Dr. Kim gained further insight into the treatment effect by examining the continuous responder curves (see Figure 1). A small degree of separation between the placebo and Cymbalta curves was noted.

Table 1: Subject Disposition

	Placebo n (%)	**DLX 60-120mg n (%)**
Randomized	120 (100%)	111 (100%)
ITT	120	111
mITT	119	108
Completed	96 (80%)	77 (69%)
Reasons for dropout		
AE	7 (6%)	15 (14%)
LOE	3 (2%)	2 (2%)
Subject Decision	9 (8%)	8 (7%)
Other	5 (3%)	9 (8%)

Source: Dr. Kim's Review, Table 1

Figure 1: Cumulative Improvement in Pain from Baseline



Source: Dr. Kim's Review, Figure 2

Dr. Kim additionally evaluated the 60 mg dose. Dr. Kim compared the placebo patients (n=113) to the group of patients that received 60 mg (n=52) for the duration of the study. Of note, in the analysis there was approximately a 2:1 ratio of placebo patients to patients receiving 60 mg throughout the study. This analysis only included a subset of the intent-to-treat population by design and was not powered for the comparison. The difference was not statistically significant.

3 CONCLUSIONS AND RECOMMENDATIONS

Dr. Kim concluded that Cymbalta 60 – 120 mg provided pain reduction for patients in the OA and CLBP populations. His conclusions pertaining to the CLBP population were derived from two studies, namely HMEN and HMGC. In the OA population, his finding was based on evidence from Study HMFG. He concluded that evidence of an effect was not apparent in Study HMEP since the results were not robust to the procedure used for handling missing data. I concur with Dr. Kim's concern that the mBOCF strategy potentially assigned good scores to patients who discontinued for reasons such as "other" and "subject decision". However, some support of the analgesic effect demonstrated in Study HMFG may be gained from Study HMEP via both the magnitude of the effect and the continuous responder curves. The magnitude of the mean reduction in pain intensity for the Cymbalta-treated patients in Study HMEP was similar to that of Study HMFG. In addition, a separation in the responder curves was apparent. The separation was small, but such small differences are common in studies conducted in OA patients.

Table 2 provides a summary of the results of the analyses of Cymbalta 60 mg.

Table 2: Summary of Efficacy Analyses of Cymbalta (duloxetine) 60 mg QD

Study	Treatment	LS Mean Change (SE)	p-value
CLBP Studies			
HMEN	DLX60 QD	-1.5 (0.19)	0.009
	Placebo	-0.9 (0.18)	
HMGC	DLX60 QD	-2.0 (0.15)	0.004
	Placebo	-1.5 (0.15)	
HMEO	DLX60 QD	-1.9 (0.20)	0.228
	Placebo	-1.5 (0.19)	
OA Studies			
HMEP	DLX60 QD	-2.1 (0.20)	0.591
	Placebo	-1.9 (0.20)	
HMFG	DLX60 QD	-1.8 (0.20)	0.012
	Placebo	-1.2 (0.19)	

Results produced using an ANCOVA model and BOCF imputation strategy.
Sources: Applicant's Table 6 – Filing Communication; Clinical Study Report Table HMEO.11.8; Dr. Kim's review, Table 19; Presentation by Dr. Kim (see Appendix)

Dr. Kim concluded that a fixed dose of Cymbalta 60 mg was effective in reducing pain in the chronic low back pain population. This conclusion was based on his review of Study

HMGC which evaluated the 60 mg fixed dose for 12 weeks. He additionally derived supportive evidence of the effect from his analysis as well as the applicant's analysis of Study HMEN. In contrast, Dr. Kim was unable to conclude that the dose was effective in the OA population since a fixed-dose study was not conducted.

I concur with Dr. Kim's assessment that Study HMFG conducted in the OA population was not designed to evaluate Cymbalta 60 mg. However, I find the additional analyses of Study HMFG provide evidence of the effect of Cymbalta 60 mg. There may be concern that the effect was not apparent in Study HMEP. However, the design of Study HMEP differed from that of Study HMFG in that patients randomized to Cymbalta were re-randomized at Week 7. This design only allowed for a comparison of a subset of the intent-to-treat population and lacked power to detect a statistically significant difference.

In conclusion, I find that Cymbalta reduces pain in patients with osteoarthritis or chronic low back pain. Moreover, there is evidence to specifically support the efficacy of 60 mg. The applicant seeks a broad chronic pain indication. While evidence exists for label claims in the populations studied, the requirements for a broader chronic pain indication is currently being discussed within the FDA as well as among external thought-leaders in the pain community.

4 APPENDIX

In Table 2 of this review, results from Study HMEP are presented. The results were provided by Dr. Yongman Kim and presented at an internal meeting. The table presented by Dr. Kim is provided below.

LS Mean Change (SE) from Baseline to Week 13 in weekly mean of 24-hour average pain	**Placebo (N=113)**	**DLX60mg (N=52)**	**P-value**
MMRM*	-2.1 (0.15)	-2.5 (0.25)	0.153
ANCOVA/BOCF**	-1.9 (0.20)	-2.1 (0.20)	0.591

*P-value calculated from MMRM model with terms for treatment, week, treatment*week, site, NSAID use, baseline, week*baseline.
**P-value calculated from ANCOVA model with terms for treatment, site, NSAID use, and baseline score as covariate.

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-22516	ORIG-1	ELI LILLY AND CO	CYMBALTA

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

DIONNE L PRICE
02/22/2010

THOMAS J PERMUTT
02/23/2010
concur