The Cost-Effectiveness of Duloxetine in Chronic Low Back Pain: A US Private Payer Perspective

Ronald C. Wielage, MPH1,*, Megha Bansal, MA1, J. Scott Andrews, PharmD2, Madelaine M. Wohlreich, MD2, Robert W. Klein, MS1, Michael Happich, PhD3

1Medical Decision Modeling Inc., Indianapolis, IN, USA; 2Eli Lilly and Company, Indianapolis, IN, USA; 3Lilly Deutschland GmbH, Bad Homburg, Germany

ABSTRACT

Objective: To assess the cost-effectiveness of duloxetine in the treatment of chronic low back pain (CLBP) from a US private payer perspective. Methods: A cost-utility analysis was undertaken for duloxetine and seven oral post–first-line comparators, including nonsteroidal anti-inflammatory drugs (NSAIDs), weak and strong opioids, and an anticonvulsant. We created a Markov model on the basis of the National Institute for Health and Clinical Excellence model documented in its 2008 osteoarthritis clinical guidelines. Health states included treatment, death, and 12 states associated with serious adverse events (AEs). We estimated treatment-specific utilities by carrying out a meta-analysis of pain scores from CLBP clinical trials and developing a transfer-to-utility equation using duloxetine CLBP patient-level data. Probabilities of AEs were taken from the National Institute for Health and Clinical Excellence model or estimated from osteoarthritis clinical trials by using a novel maximum-likelihood simulation technique. Costs were gathered from Red Book, Agency for Healthcare Research and Quality's Healthcare Cost and Utilization Project database, the literature, and, for a limited number of inputs, expert opinion. The model performed one-way and probabilistic sensitivity analyses and generated incremental cost-effectiveness ratios (ICERs) and cost acceptability curves. Results: The model estimated an ICER of $59,473 for duloxetine over naproxen. ICERs under $30,000 were estimated for duloxetine over non-NSAIDs, with duloxetine dominating all strong opioids. In subpopulations at a higher risk of NSAID-related AEs, the ICER over naproxen was $33,105 or lower. Conclusions: Duloxetine appears to be a cost-effective post–first-line treatment for CLBP compared with all but generic NSAIDs. In subpopulations at risk of NSAID-related AEs, it is particularly cost-effective.

Keywords: chronic low back pain, cost-effectiveness, cost-utility analysis, duloxetine, pharmacoeconomic model.

Copyright © 2013, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Introduction

Low back pain (LBP) is the second most common cause of disability in the United States, exceeded only by arthritis and rheumatism [1]. During a 3-month period, 28.1% of US adults experience a day or more of LBP [2]. An estimated 70% to 80% of the population will experience LBP in their lifetimes [3] of whom 10% will progress to chronic LBP (CLBP). Although studies estimating the economic burden to the US economy have varied in methodology, they agree that the cost of LBP is large: $12 to $90 billion is incurred annually in direct costs, with indirect costs perhaps three times higher [4]. Few studies differentiate the cost of CLBP versus nonchronic or acute LBP.

CLBP has been variously defined but typically is described as LBP that is present longer than 3 months [5,6]. In North America, the prevalence of CLBP is estimated at 9% to 10.2% and appears to be increasing, up from 3.9% in 1992 [7,8]. Approximately 74.4% of the CLBP population suffers moderate to severe pain [8]. CLBP is often a mixed pain syndrome with nociceptive, neuropathic, and hyperalgesic components [9].

Few clinical trials of oral treatments for CLBP have been conducted. A 2011 review of pharmacological treatments for CLBP found only four studies of nonsteroidal anti-inflammatory drugs (NSAIDs), five of antidepressants, and eight of opioids; moreover, the review found that the quality of the evidence as of publication was low [10]. A 2009 review of pharmacotherapy for chronic pain reported no treatments with good-quality evidence of substantial benefit in LBP [11].

Cost-effectiveness analyses (CEAs) have been recognized as a pressing need [12]. Even so, few have been conducted [13]. The National Institute for Health and Clinical Excellence (NICE) reported that no economic models could be located for NSAIDs, opioids, or antidepressants for its 2009 guideline for LBP [14].

Duloxetine, a selective serotonin and norepinephrine reuptake inhibitor, has demonstrated analgesic effects in CLBP in three 13-week randomized controlled trials (RCTs) [15–17]. Long-term efficacy has been demonstrated in an open label extension trial of 41 weeks [18]. A pooled analysis of duloxetine RCTs completed through 2008 reported that most treatment emergent adverse events (AEs) tended to be mild to moderate in severity.
and transitory in nature [19]. In November 2010, the US Food and Drug Administration approved duloxetine for chronic musculoskeletal pain [20]. By using the relative wealth of data available for duloxetine, we parameterized a pharmacoeconomic model to compare the cost-effectiveness of duloxetine and seven oral post–first-line comparators in CLBP treatment, including NSAIDs, weak and strong opioids, and an anticonvulsant.

**Methods**

The authors developed a semi-Markov model from a US private payer perspective for oral treatments of CLBP in a post–first-line (post-acetaminophen [APAP]) place in therapy. We modeled duloxetine and comparators representing commonly used drug classes in the US market: a nonselective NSAID and a COX-2 inhibitor, strong opioids, weak opioid/monoamine reuptake inhibitors, an anticonvulsant as well as a combination product. Specifically, the model included naproxen, celecoxib, oxycodeone extended release (oxycodeone), tapentadol extended release (tapentadol), tramadol immediate release (tramadol), pregabalin, and oxycodone/APAP as comparators to duloxetine (Table 1).

CLBP and osteoarthritis (OA) are both chronic musculoskeletal conditions that are commonly treated with NSAIDs and opioids. In both conditions, these oral treatments are not disease modifying, but provide symptomatic improvement in pain associated with the condition. Therefore, in the absence of CEAs in CLBP, we referenced the 2008 OA economic model published by NICE as a framework. Appendix D of the NICE OA guidelines documents the model, with additional documentation of treatment-specific utilities in Appendix C [33]. The NICE LBP clinical guidance refers to the OA guidelines concerning treatment with NSAIDs [14].

**Model Overview**

The model is a discrete-state, time-dependent semi-Markov model with changing probabilities as the cohort ages. Treatment efficacy, AE profile, and discontinuation as well as concomitant proton pump inhibitor (PPI) usage are the clinical dimensions modeled. Economic inputs include drug costs and medical utilization for the management of AEs, titration, and discontinuation. The model includes two types of AEs: persistent and transient. Persistent AEs disrupt treatment, increase costs, and have a permanent effect on mortality and health-related quality of life (HRQoL). Transient AEs temporarily increase costs and lower HRQoL but have no permanent effect. Treatment-specific utilities represent treatment efficacy in the model. AE profiles are modeled with 3-month probabilities of incurring persistent or transient events. We modeled aging by using age-dependent relative risks of persistent AEs, age-specific general population utility weights, and increasing mortality. The model calculates cycle-specific utilities from the interaction between utilities/ utility weights representing treatment efficacy, age, and AEs. Model inputs are parameterized from the NICE model, meta-analysis, the literature, and, for a small number of inputs, expert opinion.

**Structure**

We used a lifetime time horizon with 3-month cycles to the maximum length of treatment and annual cycles thereafter. This allows the model to accumulate the long-term effects of NSAID-related AEs. Health states include treatment, death, and 12 during- and post-persistent-AE states. In the treatment state, the patient experiences the increased HRQoL due to treatment, reductions to HRQoL due to transient AEs, and changes in HRQoL due to discontinuation and switch to a post-discontinuation basket of treatments (PDBT). The PDBT is composed of all comparators weighted by market share (days prescribed). Costs are incurred in the treatment state for treatment drugs, management of transient AEs, and medical services related to titration and discontinuation. Upon the end of the treatment period, the portion of the cohort still receiving each comparator discontinues and switches to the PDBT. In the base case, treatment is for the lesser of 1 year, until discontinuation, or until occurrence of a persistent AE and is followed by treatment from the PDBT until death.

Patients transition to other health states upon death or any of six persistent AEs. These health state transitions may take place during the original treatment period or after switch to the PDBT. The patient enters a 3-month during-AE health state followed by a post-AE state in which the patient continues until death. During these states, HRQoL, excess mortality, and cost are assessed as appropriate to the AE. The age of the cohort during each cycle determines the appropriate background mortality. Figure 1 is a simplified depiction of the model structure.

**Adverse Events**

Persistent AEs include cardiovascular (CV) and gastrointestinal (GI) AEs associated with NSAID treatment as well as fracture, an AE associated with opioid and anticonvulsant comparators (Table 2). Transient AEs included in the model occur at substantially different rates among the comparators, potentially having economic and HRQoL impacts (Table 3).

**Transition Probabilities**

We used the 3-month CV and GI AE probabilities from the NICE model for naproxen and celecoxib, and assumed equivalence to no treatment for other comparators [33]. The probabilities of fracture were derived principally from odds ratios calculated by Vestergaard and colleagues and applied to rates of fracture in the general population [35,36,46,52,62–65]. We examined duloxetine CLBP clinical trial reports for rates of fracture; fractures occurred at or below the rates found in the control arms [34,66,67]. Age-dependent relative risks from the literature were then applied to these probabilities [46,57,58].

The age-dependent probability of background mortality at each cycle is calculated from a US life table [68], while excess mortality associated with each persistent AE was derived from a variety of sources in the literature [33,46,51,53–56,69].

**Transient AE and Discontinuation Probabilities**

We conducted meta-analyses for most transient AEs and for discontinuation by using CLBP RCTs for duloxetine and OA RCTs for NSAIDs and opioids, as more OA RCTs were available. A 12-week minimum duration of treatment was among the inclusion criteria. Three-month probabilities of dyspepsia for naproxen and celecoxib were taken from the NICE model, and were assumed equivalent to no treatment for other comparators [33].

We used conventional techniques for the AE meta-analysis when possible. A maximum-likelihood simulation technique was used when an AE rate fell below the reporting threshold for one or more RCTs of a treatment. This technique assumed that all RCTs for that treatment experienced the AE rate within the same binomial distribution truncated by the publication reporting thresholds. In the case that no publications for a treatment reported a rate for an AE, a rate was assumed equal to that of another medication in the same class. AE rates for tapentadol were taken from the Nucynta package insert.

A meta-analysis of discontinuation rates from the OA RCTs above was used to calculate discontinuation probabilities in the first 3-month cycle of treatment for NSAIDs and opioids. Data were pooled from two RCTs in neuropathic pain for pregabalin
## Table 1 – Treatment characteristics.

<table>
<thead>
<tr>
<th>Comparator</th>
<th>Duloxetine</th>
<th>Celecoxib</th>
<th>Naproxen</th>
<th>Pregabalin</th>
<th>Oxycodone/APAP</th>
<th>Oxycodone ER</th>
<th>Tapentadol ER</th>
<th>Tramadol IR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dosing</td>
<td>60–120 mg</td>
<td>200 mg QD</td>
<td>500 mg BID</td>
<td>300 mg BID</td>
<td>7.5/35–15/650 Qsh</td>
<td>10–30 mg BID</td>
<td>300–600 mg QD</td>
<td>200–300 mg QD</td>
</tr>
<tr>
<td>Utility</td>
<td>0.7541*</td>
<td>0.7688*</td>
<td>0.7688*</td>
<td>0.7282† [21]</td>
<td>0.7628*</td>
<td>0.7628*</td>
<td>0.7603*</td>
<td>0.7587*</td>
</tr>
<tr>
<td>Discon—in first 3 mo (%)</td>
<td>27.6*</td>
<td>23.8†</td>
<td>30.0‡</td>
<td>35.0 [22,23]</td>
<td>58.9*</td>
<td>58.9‡</td>
<td>44.0 [24]</td>
<td>48.5‡</td>
</tr>
<tr>
<td>Discon—subsequent 3 mo (%)</td>
<td>1.9</td>
<td>4.7</td>
<td>5.7</td>
<td>4.5</td>
<td>13.3</td>
<td>13.3</td>
<td>8.3</td>
<td>25.4</td>
</tr>
<tr>
<td>Share of PDT (%) [28]</td>
<td>5.0</td>
<td>9.2</td>
<td>17.1</td>
<td>8.6</td>
<td>19.2</td>
<td>6.0</td>
<td>1.0</td>
<td>33.9</td>
</tr>
<tr>
<td><strong>Treatment costs ($)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial 3-mo drug cost [29,30]</td>
<td>576.41</td>
<td>371.09</td>
<td>162.41</td>
<td>439.83</td>
<td>154.55</td>
<td>589.04</td>
<td>1,229.27</td>
<td>262.63</td>
</tr>
<tr>
<td>Initial 3-mo physician cost [31,32]</td>
<td>167.50</td>
<td>0.00</td>
<td>0.00</td>
<td>192.84</td>
<td>184.06</td>
<td>287.65</td>
<td>169.95</td>
<td>153.82</td>
</tr>
<tr>
<td>Cost—subsequent 3 mo [29]</td>
<td>590.23</td>
<td>371.09</td>
<td>162.41</td>
<td>474.28</td>
<td>188.20</td>
<td>667.51</td>
<td>1,340.30</td>
<td>309.89</td>
</tr>
<tr>
<td>Discon drug cost [29,30]</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>94.62</td>
<td>27.57</td>
<td>190.34</td>
<td>632.28</td>
<td>44.01</td>
</tr>
<tr>
<td>Discon provider cost [31,32]</td>
<td>94.80</td>
<td>0.00</td>
<td>0.00</td>
<td>106.03</td>
<td>222.28</td>
<td>183.89</td>
<td>92.62</td>
<td>117.47</td>
</tr>
<tr>
<td><strong>3-mo persistent AE probabilities (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptomatic ulcer [33]</td>
<td>0.04*</td>
<td>0.09</td>
<td>0.28</td>
<td>0.04†</td>
<td>0.04*</td>
<td>0.04*</td>
<td>0.04*</td>
<td>0.04*</td>
</tr>
<tr>
<td>Complicated GI bleed [33]</td>
<td>0.02*</td>
<td>0.05</td>
<td>0.07</td>
<td>0.02†</td>
<td>0.02*</td>
<td>0.02*</td>
<td>0.02*</td>
<td>0.02*</td>
</tr>
<tr>
<td>Myocardial infarction [33]</td>
<td>0.06*</td>
<td>0.15</td>
<td>0.06</td>
<td>0.06†</td>
<td>0.06*</td>
<td>0.06*</td>
<td>0.06*</td>
<td>0.06*</td>
</tr>
<tr>
<td>Stroke [33]</td>
<td>0.03*</td>
<td>0.03</td>
<td>0.08</td>
<td>0.03†</td>
<td>0.03*</td>
<td>0.03*</td>
<td>0.03*</td>
<td>0.03*</td>
</tr>
<tr>
<td>Heart failure [33]</td>
<td>0.01*</td>
<td>0.04</td>
<td>0.09</td>
<td>0.01†</td>
<td>0.01*</td>
<td>0.01*</td>
<td>0.01*</td>
<td>0.01*</td>
</tr>
<tr>
<td>Fracture</td>
<td>0.40 [34]</td>
<td>0.40 [35]</td>
<td>0.45 [35]</td>
<td>0.66 [36]</td>
<td>0.59*</td>
<td>0.59 [35]</td>
<td>0.89*</td>
<td>0.89 [35]</td>
</tr>
<tr>
<td><strong>3-mo transient AE probabilities (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyspepsia</td>
<td>7.52* [33]</td>
<td>14.96 [33]</td>
<td>14.96 [33]</td>
<td>7.52* [33]</td>
<td>7.52* [33]</td>
<td>7.52* [33]</td>
<td>7.52* [33]</td>
<td>7.52* [33]</td>
</tr>
<tr>
<td>Nausea</td>
<td>8.30*</td>
<td>2.80**</td>
<td>5.00**</td>
<td>7.90 [37]</td>
<td>37.20*</td>
<td>37.20**</td>
<td>21.00 [38]</td>
<td>19.10*</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>5.70*</td>
<td>2.80**</td>
<td>4.10**</td>
<td>3.90 [37]</td>
<td>5.90*</td>
<td>5.90**</td>
<td>5.00 [38]</td>
<td>6.10</td>
</tr>
<tr>
<td>Constipation</td>
<td>7.60*</td>
<td>1.80***</td>
<td>3.30**</td>
<td>5.30 [37]</td>
<td>38.20*</td>
<td>38.20***</td>
<td>17.00 [38]</td>
<td>15.10</td>
</tr>
<tr>
<td>Insomnia</td>
<td>3.70*</td>
<td>2.30**</td>
<td>1.10**</td>
<td>1.10†</td>
<td>7.30*</td>
<td>7.30**</td>
<td>4.00 [38]</td>
<td>7.30</td>
</tr>
<tr>
<td>Pruritus</td>
<td>0.60*</td>
<td>1.80**</td>
<td>2.10**</td>
<td>0.60†</td>
<td>13.70*</td>
<td>13.70**</td>
<td>5.00 [38]</td>
<td>8.60</td>
</tr>
<tr>
<td>Vomiting</td>
<td>0.30*</td>
<td>1.30**</td>
<td>0.70**</td>
<td>3.90 [37]</td>
<td>17.10*</td>
<td>17.10**</td>
<td>8.00 [38]</td>
<td>6.90</td>
</tr>
<tr>
<td>Dizziness</td>
<td>5.40*</td>
<td>1.70**</td>
<td>1.30**</td>
<td>35.50 [37]</td>
<td>20.7*</td>
<td>20.7**</td>
<td>17.00 [38]</td>
<td>15.20</td>
</tr>
<tr>
<td>Somnolence</td>
<td>4.00*</td>
<td>0.30**</td>
<td>0.30**</td>
<td>19.70 [37]</td>
<td>21.30*</td>
<td>21.30**</td>
<td>12.00 [38]</td>
<td>9.40</td>
</tr>
<tr>
<td>Opioid abuse</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>3.34†</td>
<td>3.34 [39]</td>
<td>3.34 [38]</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Relative risk with PPI usage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptomatic ulcer</td>
<td>0.49 † [40]</td>
<td>0.25 [33]</td>
<td>0.37 [33]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
</tr>
<tr>
<td>Complicated GI bleed</td>
<td>0.49 † [40]</td>
<td>0.25 [33]</td>
<td>0.46 [33]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
</tr>
<tr>
<td>Dyspepsia</td>
<td>0.49 † [40]</td>
<td>0.25 [33]</td>
<td>0.43 [33]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
<td>0.49 † [40]</td>
</tr>
</tbody>
</table>

AE, adverse event; APAP, acetaminophen; BID, twice a day; CLBP, chronic low back pain; discon, discontinuation; ER, extended release; GI, gastrointestinal; IR, immediate release; OA, osteoarthritis; PDI, postdiscontinuation therapy; PPI, proton pump inhibitor; QD, once a day; Qsh, every 6 hours; RCT, randomized controlled trial.

* Meta-analysis of CLBP RCTs.
† Assumed the same as duloxetine.
‡ Assumed the same as oxycodeone.
** Meta-analysis of OA RCTs.
†† Expert opinion.
§ Assumed the same as no treatment.
¶ Assumed the same as tramadol.
** Meta-analysis of OA RCTs.
*† Assumed the same as the lowest comparator.
and from CLBP RCTs for duloxetine. Discontinuation in subsequent cycles was based on the initial rate of discontinuation and expert opinion.

Utilities

Treatment utilities were based on a meta-analysis and indirect comparison of pain scores in CLBP RCTs using the DerSimonian-Laird and Bucher methods [70–72]. Similar to the AE meta-analysis, searches were performed for English-language RCTs of 21 possible CLBP oral comparators. A 12-week minimum duration of treatment was among the inclusion criteria. RCTs were required to report baseline pain and change from baseline pain scores on a 1 to 10 or 1 to 100 scale. The outcome of interest was the difference in change from baseline scores between treatment and control arms.

Pain scores were converted to utilities by using a transfer-to-utility (TTU) [73–77] regression equation developed from patient-level pain and US preference-weighted EuroQol five-dimensional (EQ-5D) questionnaire scores of three duloxetine RCTs in CLBP [15–17]. Patient-level data from three Eli Lilly and Company duloxetine CLBP placebo-controlled trials were used in the analysis (ClinicalTrials.gov IDs NCT00767806, NCT00408876, and NCT00408876). Data from NCT00767806 and NCT00408876 were used to build the regression model, while data from NCT00424593 were used for validation.

The dependent variable was the last observed US EQ-5D questionnaire health index score or its transformation. The independent variables were the Brief Pain Inventory (BPI) 24-hour average pain score at the same observation as well as other potentially influencing variables such as sex and age. Histograms were plotted to assess normality of the variables. Transformations were attempted to improve normality and were assessed by skewness, kurtosis, and visual inspections of histograms and residuals plots. Variables and their transformations were considered for inclusion by using stepwise selection. Those that were statistically significant at $P < 0.10$ were considered as possible elements in the model. Goodness of fit was judged by the adjusted coefficient of determination ($R^2$), root mean square error (RMSE), and Akaike information criterion values.

Population utility weights by age and sex for the United States were taken from the National Health Measurement Study [78], while utility weights for AEs were derived from the literature [33,52,59,61,79].

Costs

Drug costs were calculated using dosing consistent with clinical trials and labeling. Average 2011 wholesale prices [29] were discounted 16%, consistent with current actual wholesale acquisition costs [80]. Opioid costs were adjusted in the first cycle of treatment to reflect titration and upon discontinuation to reflect tapering according to a Washington State Medicaid tapering calculator [30]. Titration schedules for duloxetine and pregabalin were found in the literature [15,16,23]. Titration and discontinuation-related medical service utilization was provided by expert opinion and priced by using Medicare reimbursement rates adjusted by using a Medicare/private payer cost ratio [31,32]. Costs for persistent during-AE health states were determined by using the Healthcare Cost and Utilization Project database of the Agency for Healthcare Research and Quality [43], International Classification of Diseases, Ninth Revision, codes from the Agency for Healthcare Research and Quality Inpatient Quality Indicators Program [44], and the literature [41,42,46–50]. Costs for persistent post-AE health states were taken from the literature [41,42,46–50]. When necessary, costs were inflated to 2011 US dollars by using the medical services component of the US Consumer Price Index [81].

<table>
<thead>
<tr>
<th>Table 2 – Persistent AE characteristics.</th>
<th>Adverse event</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symptomatic ulcer</strong></td>
<td><strong>Complicated GI bleed</strong></td>
</tr>
<tr>
<td><strong>Cost (3 mo) ($)</strong></td>
<td><strong>Cost (3 mo) ($)</strong></td>
</tr>
<tr>
<td>Utility weight</td>
<td>0.550 [51]</td>
</tr>
<tr>
<td>Excess mortality (3 mo) (%)</td>
<td>0.978 [33]</td>
</tr>
<tr>
<td>During</td>
<td>0.00 [33]</td>
</tr>
<tr>
<td>Post</td>
<td>0.00 [33]</td>
</tr>
<tr>
<td>Age-related relative risk</td>
<td>&lt;65 y</td>
</tr>
<tr>
<td></td>
<td>65 y +</td>
</tr>
</tbody>
</table>

AE, adverse event; GI, gastrointestinal.
predicting EQ-5D score produced an adjusted
A regression model with only the untransformed pain score
TTU Regression
pregabalin in CLBP are two phase 2 trials of unknown duration
seen in placebo arms of other RCTs. The only known RCTs of
lin was assumed to have the same efficacy as the placebo effect
dol/APAP as well as for oxycodone/APAP and oxycodone. Pregaba-
Equivalent efficacies were also assumed for tramadol and trama-
duloxetine. Celecoxib and naproxen were assumed to have the
pain scores for those taking NSAIDs, followed by opioids and
done/APAP . Our methods estimated the greatest improvement in
RCTs were found for celecoxib, pregabalin, tramadol, or oxyco-
tapentadol, one for naproxen, and one for hydromorphone. No
for tapentadol. In oxycodone-controlled trials, there were two for
etoricoxib, oxycodone, oxymorphone, and tramadol/APAP, and one
for tapentadol. In oxycodeone-controlled trials, there were two for
tation from Lilly September 8, 2011.

Sensitivity Analysis
The model enabled one-way and probabilistic sensitivity analyses on
22 and 17 variables of interest, respectively, between dulox-
etine and any other comparator. Outputs are a tornado diagram for one-way analysis and a scatter plot of incremental costs
versus quality-adjusted life-years (QALYs) as well as a cost-
effectiveness acceptability curve from probabilistic analysis. Variables chosen for sensitivity analysis were those for which there was substantial uncertainty, expectation of sensitivity, or recommendation by an advisory body.

Results

CLBP Efficacy Meta-Analysis
A systematic literature review of CLBP RCTs identified 16 clinical
trials for inclusion in a meta-analysis of efficacy of oral post-first-
line treatments in CLBP. The number of treatment arms found in
placebo-controlled trials was four for duloxetine, two each for etoricoxib, oxycodone, oxymorphine, and tramadol/APAP, and one
for tapentadol. In oxycodeone-controlled trials, there were two for
tapentadol, one for naproxen, and one for hydromorphone. No
RCTs were found for celecoxib, pregabalin, tramadol, or oxyco-
done/APAP. Our methods estimated the greatest improvement in
pain scores for those taking NSAIDs, followed by opioids and
duloxetine. Celecoxib and naproxen were assumed to have the same
efficacy as the pooled efficacy of etoricoxib and naproxen.
Equivalent efficacies were also assumed for tramadol and tram-
dol/APAP as well as for oxycodone/APAP and oxycodone. Pregaba-
lin was assumed to have the same efficacy as the placebo effect
seen in placebo arms of other RCTs. The only known RCTs of
pregabalin in CLBP are two phase 2 trials of unknown duration
described in an abstract that reports no statistically significant difference between treatment and placebo [21].

TTU Regression
A regression model with only the untransformed pain score
predicting EQ-5D score produced an adjusted $R^2$ of 0.3290.
Transformations and additional variables provided marginal improvements, and so the single independent variable model
was chosen with the equation:

\[ \text{EQ-5D} = 0.92859 - 0.04007 \text{BPI} \]

During validation, the Pearson correlation coefficient of the predicted versus actual EQ-5D scores was 0.5181.

Table 3 – Transient AE characteristics.

<table>
<thead>
<tr>
<th>Event</th>
<th>Cost ($)</th>
<th>3-mo costs ($)</th>
<th>Cost of physician visit ($)</th>
<th>Utility weight during event</th>
<th>Days of treatment*</th>
<th>Duration-adjusted cost ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyspepsia</td>
<td>28</td>
<td>49 [41,42]</td>
<td>76</td>
<td>0.730 [33]</td>
<td>27.8</td>
<td>119</td>
</tr>
<tr>
<td>Nausea</td>
<td>0</td>
<td>6’ [29]</td>
<td>76</td>
<td>0.887 [59]</td>
<td>12.5</td>
<td>77</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>0</td>
<td>6’ [29]</td>
<td>76</td>
<td>0.900 [59]</td>
<td>18.6</td>
<td>77</td>
</tr>
<tr>
<td>Constipation</td>
<td>539 [60]</td>
<td>66’ [29]</td>
<td>76</td>
<td>0.888 [59]</td>
<td>39.1</td>
<td>720</td>
</tr>
<tr>
<td>Insomnia</td>
<td>0</td>
<td>282’ [29]</td>
<td>76</td>
<td>0.887 [59]</td>
<td>34.5</td>
<td>259</td>
</tr>
<tr>
<td>Pruritus</td>
<td>0</td>
<td>47’ [29]</td>
<td>76</td>
<td>0.958 [59]</td>
<td>31.8</td>
<td>169</td>
</tr>
<tr>
<td>Vomiting</td>
<td>0</td>
<td>6’ [29]</td>
<td>76</td>
<td>0.887 [59]</td>
<td>3.8</td>
<td>76</td>
</tr>
<tr>
<td>Dizziness</td>
<td>0</td>
<td>0</td>
<td>76</td>
<td>0.887 [59]</td>
<td>15.5</td>
<td>76</td>
</tr>
<tr>
<td>Somnolence</td>
<td>0</td>
<td>0</td>
<td>76</td>
<td>0.887 [59]</td>
<td>28.8</td>
<td>76</td>
</tr>
<tr>
<td>Opioid abuse</td>
<td>5471 [60]</td>
<td>NA</td>
<td>NA</td>
<td>0.800 [61]</td>
<td>91.0</td>
<td>5471</td>
</tr>
</tbody>
</table>

AE, adverse event; NA, not applicable.
* Expert opinion.
1 Treatment and dosing from Lilly September 8, 2011.

Model Base-Case Results
The meta-analysis of CLBP RCTs and TTU conversion produced the highest utilities for NSAIDs, followed by opioids, duloxetine, and pregabalin. This was consistent with Kuijpers et al’s [10] 2011 systematic review, which found evidence of efficacy for NSAIDs and opioids, but not for antidepressants. (Duloxetine and pregaba-
lin were not included in the Kuijpers review.)

The cost-effectiveness model estimated that when AEs were considered, duloxetine became the most effective treatment, with the highest number of QALYs, as seen in Table 4. Naproxen, as the least expensive treatment, also appears on the cost-effectiveness
horizon in Figure 2. All other treatments were dominated. The incremental cost-effectiveness ratio (ICER) for duloxetine over
naproxen was estimated to be $59,473 per QALY. Subgroup analyses were performed (not shown) for a cohort aged 65 years and a cohort
at higher risk of CV and GI AEs. The model estimated ICERs for
duloxetine for these subgroups of $28,322 and $33,105, respectively.

Sensitivity Analyses
The output of a series of one-way sensitivity analyses is displayed as a tornado diagram. Figure 3 illustrates the results of
analyses with duloxetine and naproxen as comparators, with ICERs for duloxetine. The probabilities of CV AEs associated with
NSAIDs are the most sensitive inputs in the analysis. The values of
these inputs are from the NICE model and were derived by
NICE from the large landmark trials Celecoxib Long-term Arthri-
tis Safety Study, Multinational Etoricoxib and Diclofenac Arthritis
Long-term programme, and Therapeutic Arthritis Research and
Gastrointestinal Event Trial [82–85]. The second most sensitive
input is the discount rate, which in the base case (3%) is the
recommended rate for US CEAs [86]. The next most sensitive
input is the price of duloxetine. The fourth most sensitive set of
inputs, the probabilities of GI AEs associated with NSAIDs, was
also obtained directly from the NICE model.

A substantial amount of uncertainty surrounds the utilities of
duloxetine and naproxen due to the small number of RCTs in our
meta-analysis and the TTU method of calculating them. The
utility for naproxen is especially uncertain because it is based
chiefly on etoricoxib RCTs. However, our methods assigned
NSAIDs the highest estimated utilities. It is more likely that this
utility is too high rather than too low. Reducing the utility for
naproxen by 0.005 lowers the ICER to $52,080. Assuming that
naproxen and duloxetine have equal efficacy reduces it further to
$42,283. Exchanging the pain score–based utilities to an alternate
set based on Roland Morris Disability Questionnaire scores, however, has nearly no effect.
An incremental cost versus QALY scatter plot and a cost-effectiveness acceptability curve are common outputs of probabilistic sensitivity analysis. Figure 4 shows these outputs for a comparison of duloxetine versus naproxen. The scatter plot depicts the incremental costs versus QALYs from 1000 iterations of randomly selected inputs. A diagonal line across the plot represents a $50,000 threshold. The large central point indicates the mean of the incremental costs and QALYs. Placement slightly above the line is consistent with base-case results. The acceptability curve shows a nearly linear increase with the graph reaching 50% at approximately $60,000 and 82% at $70,000, consistent with base-case results. Figure 5 shows the outputs from an analysis comparing duloxetine to tramadol. In this case, almost all points in the scatter plot appear below the $50,000 threshold. The acceptability curve shows a nearly linear increase with the graph reaching 50% at approximately $60,000 and 82% at $70,000, consistent with base-case results. Figure 5 shows the outputs from an analysis comparing duloxetine to tramadol. In this case, almost all points in the scatter plot appear below the $50,000 threshold. The acceptability curve shows a nearly linear increase with the graph reaching 50% at approximately $60,000 and 82% at $70,000, consistent with base-case results.

Discussion

No previous CEA of oral medications in CLBP were identified. Therefore, a NICE model in OA was extended for cost-effectiveness evaluation in CLBP. Characteristics of the model that made it adaptable to this setting include a utility structure that takes into account treatment efficacy, general population HRQoL, and the effects of AEs. The accommodation of AEs includes those that have long-term effects on health and cost, as well as those that resolve upon the discontinuation of treatment.

Our model compares duloxetine to several commonly used oral treatments. Treatment-specific utilities were developed by using a meta-analysis of the RCTs available and the TTU technique. Where possible, data from the NICE economic model, such as the probabilities of persistent CV and GI AEs, were used. A novel maximum likelihood simulation method was used to estimate transient AE probabilities. This was done to overcome the common practice in the literature of reporting only those AEs that occur more frequently than an arbitrary reporting threshold. The model predicts duloxetine to be the most effective treatment for CLBP (most QALYs) and to be cost-effective when compared with all treatments other than generic NSAIDs. This information may be valuable to US payers and health care policymakers as well as patients and physicians when considering post–first-line pharmacological treatment options. Although a generic NSAID was included in the model (naproxen), its availability as an over-the-counter treatment may place it outside the scope of the US private payer perspective. In a scenario in which NSAIDs are inappropriate or no longer effective, duloxetine dominates strong opioids and is more effective than all other treatments.

### Table 4 – Base-case incremental results.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Total cost ($)</th>
<th>QALYs</th>
<th>Life years</th>
<th>Incremental cost ($)</th>
<th>Incremental QALYs</th>
<th>ICER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tapentadol ER</td>
<td>54,559</td>
<td>12.2029</td>
<td>17.3682</td>
<td></td>
<td></td>
<td>Dominated</td>
</tr>
<tr>
<td>Oxycodone ER</td>
<td>52,820</td>
<td>12.1974</td>
<td>17.3644</td>
<td></td>
<td></td>
<td>Dominated</td>
</tr>
<tr>
<td>Oxycodone/APAP</td>
<td>51,834</td>
<td>12.1973</td>
<td>17.3654</td>
<td></td>
<td></td>
<td>Dominated</td>
</tr>
<tr>
<td>Duloxetine</td>
<td>51,450</td>
<td>12.2123</td>
<td>17.3682</td>
<td>1,333</td>
<td>0.0224</td>
<td>$59,473</td>
</tr>
<tr>
<td>Pregabalin</td>
<td>51,338</td>
<td>12.1884</td>
<td>17.3696</td>
<td></td>
<td></td>
<td>Dominated</td>
</tr>
<tr>
<td>Tramadol IR</td>
<td>51,218</td>
<td>12.2043</td>
<td>17.3675</td>
<td></td>
<td></td>
<td>Dominated (extended)</td>
</tr>
<tr>
<td>Celecoxib</td>
<td>50,438</td>
<td>12.1887</td>
<td>17.3166</td>
<td></td>
<td></td>
<td>Dominated</td>
</tr>
<tr>
<td>Naproxen</td>
<td>50,117</td>
<td>12.1899</td>
<td>17.3252</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

APAP, acetaminophen; ER, extended release; ICER, incremental cost-effectiveness ratio; IR, immediate release; QALYs, quality-adjusted life-years.

* Costs discounted at 3%.

1 Life years and QALYs discounted at 3%.

An incremental cost versus QALY scatter plot and a cost-effectiveness acceptability curve are common outputs of probabilistic sensitivity analysis. Figure 4 shows these outputs for a comparison of duloxetine versus naproxen. The scatter plot depicts the incremental costs versus QALYs from 1000 iterations of randomly selected inputs. A diagonal line across the plot represents a $50,000 threshold. The large central point indicates the mean of the incremental costs and QALYs. Placement slightly above the line is consistent with base-case results. The acceptability curve shows a nearly linear increase with the graph reaching 50% at approximately $60,000 and 82% at $70,000, consistent with base-case results. Figure 5 shows the outputs from an analysis comparing duloxetine to tramadol. In this case, almost all points in the scatter plot appear below the $50,000 threshold. The acceptability curve shows a nearly linear increase with the graph reaching 50% at approximately $60,000 and 82% at $70,000, consistent with base-case results. Figure 5 shows the outputs from an analysis comparing duloxetine to tramadol. In this case, almost all points in the scatter plot appear below the $50,000 threshold. The acceptability curve shows a nearly linear increase with the graph reaching 50% at approximately $60,000 and 82% at $70,000, consistent with base-case results.

![Cost Effectiveness Plane](image-url)
Our analysis has a number of limitations. First, CLBP has not been defined consistently. For the purposes of this model, we have adopted a definition based on the criteria used in the clinical trials for which we had data: duration of at least 3 months and the presence of pain on most days. Second, the evidence base for oral medications in CLBP is weak. Although our
Comparators are commonly used CLBP treatments, searches found no CLBP RCTs for four comparators and only four RCTs were head-to-head. We have assumed efficacies for naproxen and celecoxib to be equal to those for etoricoxib, for tramadol to be equal to those for tramadol/APAP, and for oxycodone/APAP to be equal to that for oxycodone. We have assumed that immediate and extended-release formulations have similar efficacy. We have also assumed that AE rates for oral treatments are similar across musculoskeletal pain conditions when given at similar doses. Only the duloxetine AE rates are based on RCTs in CLBP. Ideally, long-term studies would be available for such a chronic condition. None were found.

A relative wealth of information is available for duloxetine: three high-quality CLBP RCTs, as well as a 41-week open label extension trial. Our meta-analysis found as many RCTs only for oxycodone and tapentadol, and most of these RCTs were not placebo controlled.

There is inadequate comparative information from clinical practice settings to confirm the findings of this model. A retrospective study of patients with CLBP in a privately insured claims database found that patients initiating duloxetine experienced similar costs for medical services and prescription drugs as matched patients initiating other pharmacological agents or noninvasive medical treatments. Analysis of a subset of patients employed in companies providing disability claims found that duloxetine-treated patients were associated with significantly lower costs for medical services, while drug costs were similar between the two groups [87]. Future research on the comparative costs of medical services related to medication titration, discontinuation, and the management of AEs in the care of patients with CLBP would provide further insight on the findings of this model.

A major assumption of the model is that patients receive medication every day. This would not be the case in acute or recurring LBP, and one could argue that it is unlikely in CLBP as well. However, duloxetine is intended for the management of chronic musculoskeletal pain of at least moderate severity. Most RCTs of comparators required a baseline pain score of at least 4 (moderate pain) on a 10-point scale. Moreover, duloxetine is intended to be taken daily and the RCTs on which the model is based included daily treatment, regardless of comparator.

The use of the TTU technique instead of utility scores produced by standard gamble or time trade-off methods may be criticized. However, TTU regression is being used frequently in illnesses as diverse as sleep dysfunction, OA, stroke, and bariatric surgery [73, 74, 88–90]. The regression model achieved an adjusted $R^2$ of 0.329 with an RMSE of 0.134. Validation showed that it tended to overestimate low utilities and underestimate higher utilities with a Pearson correlation coefficient of 0.5181. In comparison, Barton et al. [73] reports a preferred model for OA of five variables achieving an $R^2$ of 0.313 and an RMSE of 0.180. Grootendorst et al. [74] likewise developed a regression model for transferring Western Ontario and McMaster Universities Arthritis Index scores to Health Utilities Index Mark 3 utilities using 15 variables; it explained approximately 40% of the variance with an RMSE of 0.2065. The Grootendorst model was subsequently

![Fig. 5 - Probabilistic sensitivity analysis of duloxetine versus tramadol. ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.](image-url)
validated by Marshall et al. [91] on two sets of trial data that yielded RMSEs of 0.2066 and 0.1684. Marshall et al. [91] confirmed that predictive error was smaller with groups than with individual patients, with the error decreasing as the group size increased. Our BFI to EQ-5D regression equation explained as much variance as the Barton et al. model though somewhat less than the Grootendorst et al. model, yet yielded a smaller RMSE than either of them while using a much simpler model.

The additional AEs in our model that did not appear in the earlier NICE model are those that occur at substantially different rates among the comparators. Among these, AEs such as fracture, opioid abuse, and constipation are significant factors in the model. Other AEs such as nausea or somnolence are more likely to lead to discontinuation if not tolerated and therefore have little economic impact.

A number of plausible AEs and costs that would have benefited the cost-effectiveness of duloxetine were omitted. For example, epidemiological evidence has associated renal failure with the use of nonselective NSAIDs [92–96]. Very recently, two studies linked proton pump inhibitors to increased fracture and CV risks [97,98]. A large recent study associated significantly higher mortality with opioid misuse [99]. Costs have not been included for the management of long-term issues associated with opioid treatment, such as dose management. The effects of including these AEs and costs would be to reduce the QALYs and increase the costs of NSAID and opioid treatments, improving the general CV risks [97,98]. A large recent study associated significantly lower back pain. Arch Intern Med 2009;169:251–8.

Conclusions

Duloxetine offers an additional option for a condition for which no existing treatments have high-quality evidence of substantial benefit. Duloxetine has a different safety profile from existing treatments, which may make it a viable alternative for populations at high risk for some of the serious AEs associated with current treatments. This model estimated that duloxetine is a cost-effective post-first-line treatment for CLBP compared with all but generic NSAIDs. Therefore, duloxetine may offer a cost-effective alternative to current options for the management of CLBP.

Source of financial support: This research was supported by Eli Lilly and Company, Indianapolis, IN, USA.

REFERENCES