Research paper

Trajectories of depression symptom improvement and associated predictor analysis: An analysis of duloxetine in double-blind placebo-controlled trials

Hirofumi Tokuoka a,⁎, Hitoshi Takahashi b, Akichika Ozeki c, Atsushi Kuga a, Aki Yoshikawa d, Toshinaga Tsuji e, Madelaine M. Wohlreich f

a Eli Lilly Japan K.K., Neuroscience and Pain Products, Bio-Medicines, Medicines Development Unit Japan, Kobe, Hyogo, Japan
b Department of Psychiatry, Tokyo Women’s Medical University, Tokyo, Japan
c Eli Lilly Japan K.K., Statistical Science, Medicines Development Unit Japan, Kobe, Hyogo, Japan
d Eli Lilly Japan K.K., Scientific Communications, Medicines Development Unit Japan, Kobe, Hyogo, Japan
e Medical Affairs Department, Shinogi & Co., Ltd., Osaka, Osaka, Japan
f Eli Lilly and Company, Neuroscience, Indianapolis, IN, USA

A R T I C L E   I N F O

Article history:
Received 19 September 2015
Received in revised form 14 January 2016
Accepted 13 February 2016
Available online 17 February 2016

Keywords:
Duloxetine
Trajectory
Depression
Remission
Predictor
Placebo

A B S T R A C T

Background: In the treatment of major depressive disorder (MDD), it is not fully understood how individual symptoms improve over time (trajectory) in remitters. This study compared symptom improvement trajectories, as measured with the 17-item Hamilton Depression Rating Scale (HAM-D17), in remitters and nonremitters.

Methods: This analysis is based on 10 placebo-controlled, randomized, double-blind trials of duloxetine (40–60 mg/day) for treatment of MDD from baseline up to week 8. Remission was defined as a HAM-D17 total score ≤7 at week 8 (last observation carried forward). Trajectories of HAM-D17 items were assessed by mixed model repeated measures analysis for treatment and remitter-nonremitter comparisons. Grouping of the trajectories was performed by factor analysis. Predictor analysis using HAM-D17 items was conducted by logistic regression.

Results: There were 1555 patients in the duloxetine group (489 [31.4%] remitters) and 1206 patients in the placebo group (290 [24.0%] remitters; P < .0001). For most items, the difference in trajectories between remitters and nonremitters appeared at early time points and increased over time. Treatment response trajectories were very similar for duloxetine and placebo remitters, while duloxetine non-remitters improved more than placebo nonremitters. For duloxetine remitters, we found 3 trajectory groups of HAM-D17 items. The predictor analysis showed that improvement in 6 individual items at week 1 or 2 was significantly associated with remission at week 8.

Limitations: Generalizability of these results may be limited by the relatively short observation period used to define remission.

Conclusions: Early monitoring of some symptoms of depression may prove useful in guiding treatment decisions.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

For recovery from major depressive disorder (MDD), remission is a critical treatment goal. Failure to reach remission leads to higher probability of relapse (Pintor et al., 2003). Reaching remission is, moreover, needed to recover psychosocial functioning comparable to nondepressed people (Miller et al., 1998).

However, remission is not readily achieved. Trivedi et al. (2006), reporting results from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study, found that only 27.5% of patients treated with a selective serotonin reuptake inhibitor (SSRI) reached remission in 14 weeks when assessed with
the HAM-D. Therefore, it has been investigated what baseline characteristics of patients with MDD are associated with remission. For example, in the STAR*D study, gender, employment status, history, baseline function, and quality of life were shown to affect remission rates (Trivedi et al., 2006).

Another question regarding remission is whether or not early symptom improvement is predictive of remission. Lack of early (at 2 weeks) response to fluoxetine has been shown to predict a poor outcome at 8 weeks (Nierenberg et al., 1995, 2000). Furthermore, early improvement in the first 2 weeks as measured by Hamilton Depression Rating Scale (HAM-D) total score or subscale scores may predict later remission (Henkel et al., 2009; Katz et al., 2009; Szegedi et al., 2009). However, it is not clear whether some HAM-D items are better predictors than others.

Depressive disorders show many different types of symptoms. Accordingly, it is interesting to compare the degree of symptom improvement over time (trajectory) in patients who eventually reach remission (remitters) with those who do not (nonremitters). For example, responses on some items could improve faster or better than others in remitters, but not in nonremitters. Trajectories of individual symptoms in remitters and nonremitters were examined by secondary analysis of the STAR*D trial data (Sakurai et al., 2013). In that study, patients were treated with citalopram, and symptoms were evaluated using the 16-item Quick Inventory of Depressive Symptomatology, Self-Report (QIDS-SR16) and the 16-item Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C16) versions. Analysis of trajectories of symptoms revealed substantial separation between remitters and non-remitters for almost all symptoms. However, the differences between trajectories of remitters and nonremitters when assessed with the HAM-D, a more broadly used depression scale for clinical trials, are not known. Moreover, trajectories may depend on the antidepressant used, as each of them has its own pharmacological profile.

Duloxetine (DLX) is a potent inhibitor of serotonin (5-hydroxytryptamine) and norepinephrine reuptake and is relatively balanced in its binding affinity for serotonin and norepinephrine transporter sites (Bymaster et al., 2003; Wong and Bymaster, 2002). Acute administration of DLX increases extracellular monoamine levels (Karpa et al., 2002), thereby enhancing monoaminergic tone. DLX has demonstrated efficacy in the acute treatment of MDD in multiple randomized, double-blind, placebo (PLA)-controlled trials (Detke et al., 2002; Goldstein et al., 2002; Nemeroff et al., 2002). Moreover, a recent comparison of DLX and SSRIs suggested DLX has good efficacy for the core symptoms of MDD (Harada et al., 2015). However, despite the large amount of data supporting the efficacy of DLX as an antidepressant, it is not yet known precisely how DLX-treated patients experience improvement of symptoms. In addition, given the fact that a substantial number of patients experience remission even with PLA, it is still unknown whether there are any differences in the trajectories of improvement of individual symptoms between remitters treated with PLA and those treated with antidepressants.

In this analysis of PLA-controlled trials of DLX, we examined the trajectories of depression symptoms (items on the 17-item Hamilton Scale of Depression [HAM-D17]) in remitters, compared with nonremitters, treated with DLX. Furthermore, we examined these trajectories in PLA-treated patients, remitters versus non-remitters. A factor analysis (principal component analysis) was performed to determine if there are patterns of symptom improvement in DLX and remitters. Finally, a predictor analysis was utilized to assess if early improvement on any HAM-D17 items predict remission at endpoint.

2. Methods

2.1. Data sources

The data used in the present post-hoc analysis (Supplementary Table 1) were extracted from the integrated database of DLX clinical trials, which includes all clinical trials for MDD with DLX conducted by Eli Lilly and Co. The number of clinical trials included in the database is 39. This database allowed us to conduct a patient-level analysis. From the database, we included all of the trials which met the following inclusion criteria: acute (at least 6 weeks in duration), PLA-controlled, randomized, double-blind trials of DLX for the treatment of MDD; at least 1 DLX arm of ≥ 40 mg/day and ≤ 60 mg/day; and use of the 17-, 21-, or 24-item version of the Hamilton Depression Rating Scale (HAM-D17, HAM-D21, and HAM-D24, respectively). We excluded relapse-prevention and nonresponder trials.

As a result of screening, 10 trials met all criteria; they were used in the present post-hoc analysis (Supplementary Table 1). For the DLX group, only patients treated with ≥ 40 and ≤ 60 mg/day DLX were included in this analysis since this dose range has been confirmed as effective in the treatment of MDD and is commonly used globally (Ball et al., 2013; Cowen et al., 2005). Only data from the acute treatment phases of the studies throughout week 8 (up to day 70) were included in the analyses. No maintenance treatment phases were included in the analyses.

For the individual trials included in this analysis, patients were required to meet several inclusion criteria, including meeting the criteria for MDD, as defined by the Diagnostic and Statistical Manual-IV-Text Revision (DSM-IV-TR) (American Psychiatric Association, 2000), and were required to sign an informed consent document. Exclusion criteria typically included having any current Axis I disorder other than MDD, having a current or previous diagnosis of bipolar disorder or any psychotic disorder, having any organic mental disorder, dementia, or mental retardation, being at serious suicidal risk (in the judgment of the investigator), and having a recent history of substance abuse or dependence. Study protocols permitted minimum anxiolytic use by the patients.

The protocols for the individual studies were reviewed and approved by the applicable organizational ethical review boards. The studies were conducted in accordance with the ethical principles of the Declaration of Helsinki, Good Clinical Practice, and applicable laws and regulations.

The clinical trials included in this analysis were registered at ClinicalTrials.gov. Study Identifiers are as follows: NCT00036335; NCT00073411; NCT00406848 and NCT00536471. Studies HMAQa, HMAQb, HMAQc, HMAQd, HMBHa and HMBHb predate the registration requirement.

2.2. Time frame of analysis

The time frame of data collection was restricted to the acute phase of the clinical trial, beginning with the baseline visit and ending with the endpoint visit of the acute phase or 70 days after baseline, whichever occurred first. The time frame start definition excluded any “lead-in phase,” and the definition of time frame end excluded the data in any open-label trial extension. Analysed data included baseline (week 0), week 1 (day 7, range 1–10 days postbaseline), week 2 (day 14, range 11–21 days postbaseline), week 4 (day 28, range 22–35 days postbaseline), and week 8 (day 56, range 43–70 days postbaseline). If there were more than 2 data points within the allowance range of a specific time point for a patient, the data nearest the date of the specific time point was used. If there were more than 2 data points at the nearest date of the specific time point for a patient, the average of these data points was used for the specific time point for that patient.
2.3. Statistical analysis

Based on intent-to-treat principles, all randomized patients assigned to DLX (40 to 60 mg/day) or placebo with baseline and at least 1 postbaseline HAM-D assessment were included in the analyses. A last observation carried forward (LOCF) approach was used to determine remitter/nonremitter status for patients. Remission was defined as a HAM-D-17 total score ≤7 at week 8 or endpoint.

Trajectories (assessment of efficacy over time) were calculated and graphed for individual items, the HAM-D-17 total score (Hamilton, 1960), and subscale scores. Subscales included in the analysis were as follows: Bech (HAM-D-17 items 1, 2, 7, 8, 10, and 13) (Bech et al., 1975), Maier (HAM-D-17 items 1, 2, 7, 8, 9, and 10) (Maier and Philipp, 1985), Anxiety/somatization (HAM-D-17 items 10, 11, 12, 13, 15, and 17), Retardation (HAM-D-17 items 1, 7, 8, and 14), and Sleep (HAM-D-17 items 4, 5, and 6) (Cleary and Guy, 1977).

A mixed model repeated measures (MWMR) analysis was used for treatment and remitter status comparisons. In particular, the model was response = group + treatment + week + group*treatment + week*treatment + group*week + treatment + patient: (treatment or PLA; group: remitter or nonremitter; week: 0, 1, 2, 4, or 8; with patient = random effect. For sensitivity analysis, the same MRM method on restricted population with the complete dataset (available case = AC) was applied. Here, AC means patients data with all baseline and all postbaseline (weeks 0, 1, 2, 4, and 8) for all HAM-D items. Hence, remitters and nonremitters were defined without LOCF.

To group individual HAM-D-17 items in terms of a similar improvement pattern over time, factor analysis (principal component analysis) with a Varimax rotation was applied to the individual HAM-D-17 items at each time point. Data from DLX and PLA remitters were used. First, for each time point, change from baseline divided by the item maximum value was calculated for each of 17 items for each subject, and factor analysis was applied. The extraction process was terminated when the last eigenvalue was less than 1.00. Second, given the factors at each time point, we identified groups of HAM-D-17 items which showed the same pattern of factor categorization for all time points (weeks 1, 2, 4, and 8). See more details in Supplementary Methods.

A predictor analysis was also performed to investigate to what extent early improvement of each HAM-D-17 item predicts remission. For that analysis, a logistic regression, in which response equaled remitter versus nonremitter, was used to compare change on the HAM-D-17 total score, subscale scores, and individual item scores from week 0 to week 1 and week 0 to week 2. For example, DLX remitters with HAM-D-17 item 1 week 0 to week 1 change on trajectory regression model was Logit (Response) = item 1 Change (week 1–week 0).

We further used stepwise and backward algorithms for model selection based on all HAM-D-17 items. See more details in Supplementary Methods. Based on selected models, Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) were created using sensitivity and specificity to check predictability. AUC can be used as a summary measure of how well the selected models predict remitter/nonremitter status.

All statistical analyses were conducted using SAS v 9.2 (SAS Institute, Inc., Cary, NC, USA). All statistical tests were conducted at the .05 significance level (2-sided), with no adjustments for multiple comparisons. ROC curves were created with SAS v 9.2; other figures were created with R 3.1.0 for Windows (64 bit version).

3. Results

3.1. Patient baseline characteristics and disposition

From 10 PLA-controlled studies of DLX, a total of 2761 patients were included in the present analysis. There were 1555 patients in the DLX group, of which 489 (31.4%) were remitters; there were 1206 patients in the PLA group, of which 290 (24.0%) were remitters (P < .0001). Groups (DLX versus PLA) were well balanced in terms of their baseline demographics and disease characteristics (Table 1).

Patient disposition is presented in Table 2. DLX remitters, compared to DLX nonremitters, had a lower early discontinuation rate, primarily due to a lower frequency of discontinuations due to adverse events (AEs) in this group. PLA remitters also had a lower early discontinuation rate than PLA nonremitters, mostly because fewer remitters discontinued due to lack of efficacy. Overall, the rate of early discontinuation was lower in DLX-treated patients, compared to PLA-treated patients. For DLX-treated patients, the most common reason for early discontinuation was AEs, both for remitters and nonremitters. For PLA-treated patients, the most common reason for early discontinuation for remitters was “patient decision,” whereas for nonremitters the most common reason for early discontinuation was “lack of efficacy”.

3.2. Trajectories of HAM-D-17 items and subscales in duloxetine-treated patients

DLX remitters and DLX nonremitters demonstrated different trajectories in most of the HAM-D-17 items (MWMR with actual visit scores; Fig. 1A). For some items, lower scores for remitters, compared to nonremitters, were observed even at baseline (e.g. item 1 [depressed mood], item 7 [work and activity], and item 10 [anxiety: psychic]). For most items, the difference between remitters and nonremitters increased over time. Note that the ranges of scores is 0–2 for items 4–6, 12–14, 16, and 17, 0–4 for all other items.

There was a tendency for the greatest improvement by DLX remitters to be in those items for which the mean baseline scores were the highest. This observation is best exemplified by items 1 (depressed mood) and 7 (work and activity). Two other items, both with a baseline score greater than 1.5 (2 [feelings of guilt] and 10 [anxiety: psychic]) showed at least an improvement of 1.0. Among items with a score range of 0–2, item 13 (somatic symptoms: general) had the highest mean baseline scores and the largest improvement over the 8-week period.

The treatment response trajectories for DLX remitters and nonremitters on the HAM-D-17 total and subscales are shown in Fig. 1B. All subscales as well as the total score showed substantial differences between remitters and nonremitters at baseline (remitters had less severe symptoms), and these differences developed over time.

Fig. 2A presents comparisons of trajectories of depression symptoms between DLX- and PLA-treated patients in remitters and nonremitters for each of the items of the HAM-D17. A similar trajectory of symptom changes was observed for the DLX remitters and PLA remitters on virtually all items. Small separations between these trajectories were observed at baseline or other assessment time points, but not at endpoint (e.g. items 1 [depressed mood], 5 [insomnia: middle], 10 [anxiety: psychic], and 12 [somatic: gastrointestinal]). For nonremitters (DLX versus PLA), trajectories were not as similar. On several items (notably items 1 [depressed mood], 2 [feelings of guilt], 3 [suicide], 7 [work and activities], 8 [retardation], 10 [anxiety: psychic], and 15 [hypochondriasis]), there was more improvement in the DLX nonremitters, compared to the PLA nonremitters, at week 8 (P < .05.
affected patients; PLA

Abbreviations:

DLX

shown in Fig. 2B. As with the individual items (Fig. 2A), DLX remitters on total score and subscale scores of the HAM-D17 are for all).

Patient disposition.

Baseline demographics and disease characteristics.

Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DLX &gt; PLA (N=2761)</th>
<th>PLA</th>
<th>DLX Nonremitters (n=1066)</th>
<th>PLA Nonremitters (n=916)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, n (%)</td>
<td>Male 985 (35.7)</td>
<td>426 (35.3)</td>
<td>780 (64.7)</td>
<td>.6663</td>
</tr>
<tr>
<td></td>
<td>Female 1776 (64.3)</td>
<td>387 (36.3)</td>
<td>679 (52.7)</td>
<td>.0094</td>
</tr>
<tr>
<td>Race/ethnic origin, n (%)</td>
<td>Caucasian 2107 (76.3)</td>
<td>932 (77.3)</td>
<td>214 (73.8)</td>
<td>.1893</td>
</tr>
<tr>
<td></td>
<td>African 287 (10.4)</td>
<td>124 (10.3)</td>
<td>27 (9.3)</td>
<td>.0971</td>
</tr>
<tr>
<td></td>
<td>American Hispanic OR</td>
<td>309 (11.2)</td>
<td>126 (10.4)</td>
<td>44 (15.2)</td>
</tr>
<tr>
<td></td>
<td>Latino Other 2015 (73.0)</td>
<td>49.3 (16.9)</td>
<td>5 (1.7)</td>
<td>.0001</td>
</tr>
<tr>
<td>HAM-D17 total score, mean (SD)</td>
<td>20.3 (5.2)</td>
<td>20.5 (5.3)</td>
<td>18.9 (5.6)</td>
<td>21.2 (4.9)</td>
</tr>
<tr>
<td>Current MDD episode, n (%)</td>
<td>First 511 (18.5)</td>
<td>227 (18.8)</td>
<td>49 (16.9)</td>
<td>178 (19.4)</td>
</tr>
<tr>
<td></td>
<td>Other 2015 (73.0)</td>
<td>809 (67.1)</td>
<td>192 (66.2)</td>
<td>617 (67.4)</td>
</tr>
<tr>
<td></td>
<td>Missing 235 (8.5)</td>
<td>170 (14.1)</td>
<td>49 (16.9)</td>
<td>121 (13.2)</td>
</tr>
<tr>
<td></td>
<td>Age at first episode (years), mean (SD)</td>
<td>46.2 (15.9)</td>
<td>45.3 (15.5)</td>
<td>45.7 (15.2)</td>
</tr>
<tr>
<td>Parameter</td>
<td>DLX Remitters (n=489)</td>
<td>PLA Remitters (n=290)</td>
<td>DLX Nonremitters (n=1066)</td>
<td>PLA Nonremitters (n=916)</td>
</tr>
<tr>
<td>Early discontinuation, n (%)</td>
<td>368 (23.7)</td>
<td>338 (28.0)</td>
<td>318 (29.8)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Reasons for early discontinuations, n (%)</td>
<td>Adverse events 126 (8.1)</td>
<td>54 (4.5)</td>
<td>5 (1.7)</td>
<td>40 (5.3)</td>
</tr>
<tr>
<td></td>
<td>Patient decision 82 (5.3)</td>
<td>71 (5.9)</td>
<td>11 (3.8)</td>
<td>60 (6.6)</td>
</tr>
<tr>
<td></td>
<td>Lost to follow-up 62 (4.0)</td>
<td>48 (4.5)</td>
<td>9 (3.1)</td>
<td>49 (5.3)</td>
</tr>
<tr>
<td></td>
<td>Lack of efficacy 50 (3.2)</td>
<td>49 (4.6)</td>
<td>1 (0.3)</td>
<td>118 (12.9)</td>
</tr>
<tr>
<td></td>
<td>Physician decision 35 (2.3)</td>
<td>21 (1.8)</td>
<td>8 (2.8)</td>
<td>14 (1.5)</td>
</tr>
<tr>
<td></td>
<td>Sponsor decision 9 (.6)</td>
<td>9 (.7)</td>
<td>1 (.3)</td>
<td>8 (.9)</td>
</tr>
<tr>
<td></td>
<td>Death 1 (.1)</td>
<td>1 (.1)</td>
<td>0</td>
<td>1 (.1)</td>
</tr>
<tr>
<td></td>
<td>Other 1 (.1)</td>
<td>1 (.1)</td>
<td>0</td>
<td>1 (.1)</td>
</tr>
</tbody>
</table>

Abbreviations: DLX=duloxetine; HAM-D17=17-item Hamilton Depression Rating Scale; MDD=major depressive disorder; N=total number of patients; n=number of affected patients; PLA=placebo; SD=standard deviation.

a Compares “total” columns.

Table 2

Patient disposition.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DLX</th>
<th>PLA</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early discontinuation, n (%)</td>
<td>368 (23.7)</td>
<td>338 (28.0)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Reasons for early discontinuations, n (%)</td>
<td>Adverse events 126 (8.1)</td>
<td>54 (4.5)</td>
<td>5 (1.7)</td>
</tr>
<tr>
<td></td>
<td>Patient decision 82 (5.3)</td>
<td>71 (5.9)</td>
<td>11 (3.8)</td>
</tr>
<tr>
<td></td>
<td>Lost to follow-up 62 (4.0)</td>
<td>48 (4.5)</td>
<td>9 (3.1)</td>
</tr>
<tr>
<td></td>
<td>Lack of efficacy 50 (3.2)</td>
<td>49 (4.6)</td>
<td>1 (0.3)</td>
</tr>
<tr>
<td></td>
<td>Physician decision 35 (2.3)</td>
<td>21 (1.8)</td>
<td>8 (2.8)</td>
</tr>
<tr>
<td></td>
<td>Sponsor decision 9 (.6)</td>
<td>9 (.7)</td>
<td>1 (.3)</td>
</tr>
<tr>
<td></td>
<td>Death 1 (.1)</td>
<td>1 (.1)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Other 1 (.1)</td>
<td>1 (.1)</td>
<td>0</td>
</tr>
</tbody>
</table>

Abbreviations: DLX=duloxetine; N=total number of patients; n=number of affected patients; PLA=placebo.

a A completer was defined as a patient who completed the acute phase of the original study or at least continued beyond week 8 of data collection. Otherwise, the threshold for completion was the last date of HAM-D17 measurement assigned to week 8 or, if this was not available, day 70 postbaseline.

for all).

Comparisons of trajectories of DLX remitters and DLX nonremitters on total score and subscale scores of the HAM-D17 are shown in Fig. 2B. As with the individual items (Fig. 2A), DLX remitters and PLA remitters showed similar trajectories for total HAM-D17 scores and all subscale scores. DLX nonremitters showed better improvement than nonremitters for the HAM-D17 total score and all subscale scores (week 8, $P < .05$), except the Sleep subscale.

As a sensitivity analysis, the same analysis of trajectories of HAM-D items was applied to a more restricted population with a complete dataset (no missing data), for which imputation by LOCF was not needed. The results were almost identical to those shown in Fig. 1A and Fig. 2A (data not shown). Hence, the trajectory data based on the LOCF method we originally used for the definition of remitters and nonremitters can be regarded fairly reasonable and robust.

3.3. Grouping of trajectories

To better characterize the improvement in responses on HAM-D17 items of DLX-treated remitters, we assessed if there were groups of HAM-D17 items based on their trajectories.
Conceptually, we applied similar statistical method used by Kasper and Dienel (2002). A factor analysis (Table 3) identified 5-6 groups at each assessment point (weeks 1, 2, 4, and 8). Some items were categorized in the same group with respect to the highest loading at all time points, and these items were designated as a trajectory group. We found 3 such trajectory groups (A, B, and C). Group A

Fig. 1. Trajectories of depression symptoms in remitters and nonremitters treated with DLX. (A) Individual HAM-D17 items. The ranges of scores is 0-2 for items 4-6, 12-14, 16, and 17; 0-4 for all other items. (B) HAM-D17 total and subscales. The maximum possible value for the HAM-D17 total score is 52; the maximum possible values for the subscale scores are as follows: Bech, 22; Maier, 24; Retardation, 14; Sleep, 6; and Anxiety/somatization, 18. ***p < .001. Abbreviations: DLX = duloxetine; gastro. = gastro-intestinal; HAM-D17 = 17-item Hamilton Scale of Depression.
Fig. 2. Comparison of trajectories of depression symptoms between DLX- and PLA-treated patients in remitters and nonremitters. (A) Individual HAM-D17 items. The ranges of scores is 0-2 for items 4–6, 12–14, 16, and 17; 0-4 for all other items. (B) HAM-D17 total and subscales. The maximum value for the HAM-D17 total score is 52; the maximum values for the subscale scores are as follows: Bech, 22; Maier, 24; Retardation, 14; Sleep, 6; and Anxiety/somatization, 18. DLX remitters versus PLA remitters; †DLX nonremitters versus PLA nonremitters (1, 2 and 3 symbols refers to \( P < .05 \), \( P < .01 \), and \( P < .001 \), respectively). Abbreviations: DLX = duloxetine; gastro. = gastro-intestinal; HAM-D17 = 17-item Hamilton Scale of Depression; PLA = placebo.
Table 3
Grouping of HAM-D17 items’ trajectories in remitters treated with DLX based on factor analysis.

<table>
<thead>
<tr>
<th>HAM-D17 items</th>
<th>Primary factor at each week</th>
<th>Trajectory group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Week 1</td>
<td>Week 2</td>
</tr>
<tr>
<td>1 Depressed mood</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2 Feelings of guilt</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3 Suicide</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4 Insomnia early</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5 Insomnia middle</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>6 Insomnia late</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>7 Work and activities</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>8 Retardation</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>9 Agitation</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>10 Anxiety (psychic)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11 Anxiety (somnolence)</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>12 Somatic (gastro-intestinal)</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>13 Somatic (general)</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>14 Genital symptoms</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>15 Hypochondriasis</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>16 Weight loss</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>17 Insight</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

Note: Factor analysis with Varimax rotation was conducted to characterize HAM-D17 individual items at each week. A trajectory group was defined as a group of HAM-D17 items which showed the same pattern of primary factor at all time points. A minus sign (-) indicates an opposite vector sign in that particular factor and implies a change in score in the direction opposite to the other items with the same primary factor. At week 4, items 11, 14 and 17 were characterized with factor 6. However, the change in score for item 17 was opposite to items 11 and 14. Abbreviations: DLX=duloxetine; HAM-D17=17-item Hamilton Scale of Depression.

We used both a stepwise algorithm and a backward algorithm to analyse the prediction of remission (at week 8) by early improvement of each HAM-D17 item, partly to compare with the previous study by Sakurai et al. (2013) (Supplementary Fig. 2A). For DLX, the majority of HAM-D17 items showed a statistically significant positive association between early improvement (week 1 or week 2) and remission at week 8. In contrast, fewer HAM-D17 items demonstrated a statistically significant relationship with remission for the PLA group than for the DLX group (Supplementary Fig. 2A). These results on individual items demonstrated consistent trend with the stepwise algorithm result.

In both DLX- and PLA-treated patients, early improvement of the HAM-D17 total score and all subscale scores was significantly associated with remission at week 8 (Supplementary Fig. 2B). There were few marked differences between the DLX and PLA groups at week 1 or week 2. However, at week 1, the Sleep subscale was the most predictive of remission for the PLA group (OR=1.28, 95% CI=[1.17–1.39]), but was the least predictive subscale for the DLX group (OR=1.15, 95% CI=1.08–1.23).

4. Discussion

4.1. Trajectory of symptom improvement in remitters and nonremitters

In this report, we have demonstrated that response trajectories of HAM-D17 items are different in remitters and nonremitters treated with DLX. To our knowledge, this is the first report on the response trajectories of remitters and nonremitters using the HAM-D as the assessment scale in patients with depression.

It is clinically important how remitters and nonremitters differ early on in their responses to treatment for their depression. Early recognition of clinical response (or lack of response) would allow for earlier adjustment of dosage or change of pharmacotherapy to optimize treatment for the patient. Our data suggest some items show more improvement than others and that this improvement is evident as early as 1 or 2 weeks after the start of treatment. The largest improvement was shown in items 1 (depressed mood) and 7 (work and activities). These results are consistent with findings.
as shown in this study as well as in previous reports (Detke et al., 2000). Symptom improvement. A larger percentage of DLX-treated patients followed similar trajectories of DLX- and PLA-treated patients, and showed that PLA. The present report addressed this question by comparing antidepressants follow the same trajectory as remitters treated with placebo (imipramine and moclobemide) and placebo, and resilience is proposed as an underlying mechanism (Stassen et al., 2007). DLX may facilitate initiation of the remission process, while not affecting the speed of the recovery process from depression. Thus, those core symptoms could be important indicators to evaluate the treatment outcomes.

Trajectory-based analysis has been used for research of depression treatment and outcomes. For example, depressed mood and positive well-being improve in parallel, while physical symptoms and work functioning follow a different pattern of improvement (Aikens et al., 2008). It has been demonstrated with trajectory-based statistical analysis and growth mixture modeling that DLX-treated patients with MDD follow 1 of 2 trajectories: responders and non-responders (Gueorgueva et al., 2011). However, these studies did not focus on the differences in trajectories among different depression symptoms. Recently, Sakurai et al. (2013) demonstrated that depression symptoms show improvement with citalopram treatment over time in both remitters and non-remitters, as measured using the QIDS-SR16 and QIDS-C16. Our data support that finding, though with a different treatment (DLX) and using a different scale (the HAM-D17) to measure symptom improvement.

It has not been well investigated if remitters treated with antidepressants follow the same trajectory as remitters treated with placebo. The present report addressed this question by comparing trajectories of DLX- and PLA-treated patients, and showed that remitters from both treatment groups follow similar trajectories of symptom improvement. A larger percentage of DLX-treated patients experience remission compared with PLA-treated patients as shown in this study as well as in previous reports (Detke et al., 2002; Goldstein et al., 2002; Nemeroff et al., 2002). Therefore, more patients reach remission with DLX treatment, though remitters’ trajectories are similar with DLX and PLA treatment. One possible explanation for this similarity in remitters is that the speed of the recovery process from depression may be limited biologically, and DLX may not accelerate the process any further. In relation to this possibility, it has been reported that the time to onset of improvement in responders is virtually the same for antidepressants (imipramine and moclobemide) and placebo, and resilience is proposed as an underlying mechanism (Stassen et al., 2007). DLX may facilitate initiation of the remission process, while not affecting the speed of the remission process.

When DLX nonremitters and PLA nonremitters are compared, DLX showed greater improvement in some items. Thus, even patients who do not appear to be following a trajectory leading to remission may benefit from DLX treatment. This effect was most strongly demonstrated in item 1 (depressed mood), a core symptom of depression, and item 10 (anxiety: psychic). Improvement in symptoms of anxiety, even in nonremitters, makes sense given DLX’s demonstrated efficacy in the treatment of anxiety (Norman and Olver, 2008).

4.2. Grouping of trajectories

The results of the grouping by factor analysis have face validity and make intuitive sense. Group A, for example, comprises the 3 sleep-related items, which would be expected to improve together. Group C is composed of item 12 (somatic [gastrointestinal] symptoms and item 16 (weight loss), which also would be expected to be related. One might also expect that item 7 (work and activities), item 8 (psychomotor retardation), and item 13 (somatic symptoms: general) would improve together (Group B). These 3 items are joined by item 1 (depressed mood), ignoring week 1. Group B is composed of items closely related to the core symptoms of depression. The data suggest that these HAM-D17 items change in parallel for patients who will eventually be classified as remitters.

Grouping of depression symptoms using the HAM-D has been attempted previously (Cleary and Guy, 1977; Kasper and Dienel, 2002). Those studies used HAM-D scores at baseline or endpoint for the analysis, in a static manner. In the present study, we used HAM-D17 score change from baseline to assess whether symptoms can be grouped in terms of improvement pattern (in remitters) over time. Despite this methodological difference, those previous reports and our results highlight a close relationship between gastrointestinal symptoms and loss of weight, as well as the relationship among the 3 insomnia items.

4.3. Predictor analysis of early improvement for remission

Prediction of depression treatment outcome based on early symptom change has been extensively investigated. Regarding DLX, early symptom change, evaluated by HAM-D17 subscale scores, has been shown to predict remission (Katz et al., 2009).
However, it has not been shown how early improvement in each HAM-D17 item may predict the outcome. Early prediction was attempted in a study using the QIDS-SR16 and QIDS-C16 to evaluate depression symptom improvement with citalopram (Sakurai et al., 2013). They found some core symptoms (sad mood, negative self-view, feeling slowed down, low energy, and restlessness) to be associated with remission. Our results are consistent with the Sakurai et al. findings, suggesting the importance of core symptoms in early the prediction of treatment outcome with DLX as well. Moreover, the clinical trials used in this analysis employed the HAM-D17, which is more commonly used in clinical trials of MDD than QIDS-SR16 and QIDS-C16.

Both the stepwise algorithm and the backward algorithm for model selection lead to the same models of logistic analysis. The selected models were further verified by ROC curves. Our finding highlighted a few key items in the prediction of remission when treated with DLX. Notably, items 1 (depressed mood), 2 (feelings of guilt), 5 (insomnia middle), and 7 (work and activities) are consistently significant factors at both week 1 and 2. There were other items identified as significant factors: items 8 (retardation) and 16 (weight loss) at week 1, and items 9 (agitation) and 13 (somatic symptoms: general) at week 2. Those items are worthy of attention as well in assessing recovery. Notably, 6 of those identified items (1, 2, 7, 8, 9, and 13) are included in the Bech subscale (Bech et al., 1975) or Maier subscale (Maier and Phillipp, 1985), representing core symptoms of MDD. With the exception of item 9 (anxiety), these items also showed similar patterns of trajectories, as seen in Table 3. Thus, these data suggest that prediction of remission could be improved by focusing on the core symptoms of MDD.

In contrast, we identified fewer items in our analysis of the PLA group. Most identified items were also significant in the DLX group: items 1 (depressed mood), 5 (insomnia middle), and 7 (work and activities), supporting the importance of those items in depression treatment. Notably, the comparison with PLA manifests the involvement of item 2 (feelings of guilt) in the DLX group. The result that item 2 (feelings of guilt) is a significant predictor in the DLX group but not in the PLA group suggests DLX has an effect on item 2 (feelings of guilt) relevant to remission.

The ROC curve analysis supported those results from a different perspective. There was a consistent tendency that AUCs at week 1 were less significantly informative than those at week 2. For example, the DLX week 1 AUC (0.6557 [95% CI = 0.6255–0.6859]) was smaller than the DLX week 2 AUC (0.6865 [95% CI = 0.6563–0.7167]). It is a reasonable result because week 2 was closer to the endpoint than week 1. Therefore, it should be more predictable. Nevertheless, it should be noted that the data of individual items’ change at week 1 contains statistically significant information for prediction.

We also conducted predictor analyses for individual items. The results were similar to those from the stepwise algorithm. For example, for DLX week 1, 4 of the 6 items selected by the stepwise algorithm (Fig. 3) had the largest ORs estimated by individual logistic regressions (item 8 = 1.58, item 7 = 1.53, item 1 = 1.46, item 2 = 1.43; Supplementary Fig. 2). Similar results were found for DLX week 2 (5 out of 6 items overlapped), for PLA week 1 (3 out of 3 items overlapped), and for PLA week 2 (3 out of 4 items overlapped).

The HAM-D17 total score and subscale scores were all significantly predictive of remission. Of these, the Retardation subscale showed the highest OR, suggesting that symptoms related to this subscale, such as low energy, are worth paying close attention to when treating depression with DLX. This result is also consistent with the result of analysis for each item at week 1 by stepwise algorithm. One may notice that the most of the ORs for the subscales are smaller than those of the individual items. This is largely because the OR is shown per 1-point change in each scale.

Together with trajectory data, those prediction analyses suggest that most of the symptoms show better improvement over time in remitters compared with nonremitters, and particularly some core symptoms are more worthy of attention in assessing possibility of remission.

4.4. Limitations

There are several limitations which should be considered when evaluating the results of the present analyses. First, remission was defined by HAM-D17 total score at week 8 only. Trajectories and early prediction for remission after week 8 are possibly different. Note that remission rate increases over time, even after week 8 (Detke et al., 2004). Therefore, the nonremitter group in our analyses may contain patients who would have remitted if treatment had been continued. Second, differences in the assessment time points and frequency of HAM-D17 assessments in the included studies made it difficult to evaluate all patients in the same manner at each of the included time points (e.g. weeks 1, 2, 4, and 8). Third, the generalizability of our results to a clinical practice setting may be limited by the inclusion and exclusion criteria that are inherent to clinical trials. Fourth, the present analyses are based on patients whose DLX dose was ≥ 40 and ≤ 60 mg/day. Therefore, the results shown in this study may not be applicable to DLX doses outside of that range. These data (using treatment with DLX) may not easily generalize to other antidepressants, though the similarity between the DLX and results suggests there may be commonalities with other antidepressants.

Fifth, we used MMRM as a valid method for our dataset even though normality assumption was not fully met. However, the MMRM can provide intuitive visual comparison by treating individual items as continuous variable, rather categorical variables. Moreover, other statistical methods such as nonparametric methods (smoothing) would not be applicable due to sparse timepoints (weeks 0, 1, 2, 4, and 8) in our case.

Finally, because this was a post hoc analysis and no adjustments for multiple comparisons were made, the powering of the study should be considered when evaluating the outcomes of the variables.

4.5. Conclusions

Our analyses of trajectories in remitters and nonremitters, and the predictions of remission by individual HAM-D items, highlight the relevance of core depressive symptoms for remission in MDD patients treated with DLX. The results presented here suggest that early, substantial improvements in some core symptoms of depression may prove useful in guiding clinical decisions during the first few weeks of treatment with DLX.

Conflicts of interest


Contributors

All of the authors (H Tokuoka, H Takahashi, AO, AK, AY, TT, and MMW) contributed to the conception of the study, interpretation
of the data, drafting and critical revision of the manuscript, and provided final approval of the manuscript. In addition, AO performed the data analysis.

Acknowledgements
This study was sponsored by Eli Lilly Japan K.K., manufacturer/licensee of Cymbalta™, who was involved in the preparation of the manuscript. All authors participated in the drafting, critical revision, and approval of the final version of the manuscript. Medical writing assistance was provided by Rodney Moore, PhD, in-Ventiv Health Clinical, LLC, funded by Eli Lilly Japan K.K. Statistical support was provided by Ralf Jaeger, PhD, a full-time employee of Accovion GmbH.

Appendix A. Supporting information
Supplementary data associated with this article can be found in the online version at http://dx.doi:10.1016/j.jad.2016.02.039.

References