



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Dodd, S;Berk, M;Kelin, K;Zhang, Q;Eriksson, E;Deberdt, W;Craig Nelson, J

Title:

Application of the Gradient Boosted method in randomised clinical trials: Participant variables that contribute to depression treatment efficacy of duloxetine, SSRIs or placebo

Date:

2014-10-15

Citation:

Dodd, S., Berk, M., Kelin, K., Zhang, Q., Eriksson, E., Deberdt, W. & Craig Nelson, J. (2014). Application of the Gradient Boosted method in randomised clinical trials: Participant variables that contribute to depression treatment efficacy of duloxetine, SSRIs or placebo. *Journal of Affective Disorders*, 168, pp.284-293. <https://doi.org/10.1016/j.jad.2014.05.014>.

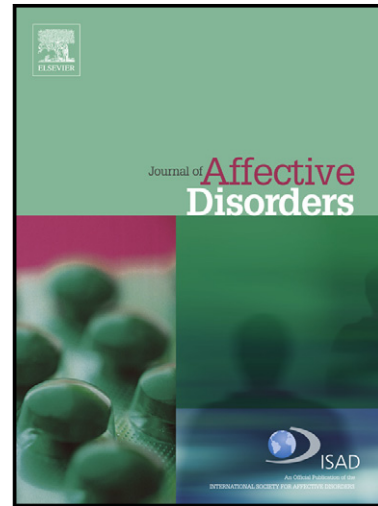
Persistent Link:

<https://hdl.handle.net/11343/43858>

Author's Accepted Manuscript

Application of the gradient Boosted method in randomised clinical trials: Participant variables that contribute to depression treatment efficacy of duloxetine, SSRIs or placebo

Seetal Dodd, Michael Berk, Katarina Kelin, Qianyi Zhang, Elias Eriksson, Walter Deberdt, J. Craig Nelson



www.elsevier.com/locate/jad

PII: S0165-0327(14)00302-4
DOI: <http://dx.doi.org/10.1016/j.jad.2014.05.014>
Reference: JAD6759

To appear in: *Journal of Affective Disorders*

Received date: 29 January 2014
Revised date: 13 May 2014
Accepted date: 14 May 2014

Cite this article as: Seetal Dodd, Michael Berk, Katarina Kelin, Qianyi Zhang, Elias Eriksson, Walter Deberdt, J. Craig Nelson, Application of the gradient Boosted method in randomised clinical trials: Participant variables that contribute to depression treatment efficacy of duloxetine, SSRIs or placebo, *Journal of Affective Disorders*, <http://dx.doi.org/10.1016/j.jad.2014.05.014>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Title page**Application of the Gradient Boosted Method in randomised clinical trials: Participant variables that contribute to depression treatment efficacy of duloxetine, SSRIs or placebo**

Authors:

Seetal Dodd^{1,2*}, Michael Berk^{1,2,3,4}, Katarina Kelin⁵, Qianyi Zhang⁶, Elias Eriksson⁷, Walter Deberdt⁸, J. Craig Nelson⁹

Affiliations:

- 1/. Department of Psychiatry, University of Melbourne, Parkville, Victoria, Australia
- 2/. IMPACT SRC (Innovation in Mental and Physical Health and Clinical Treatment), School of Medicine, Deakin University, Geelong, Victoria, Australia
- 3/. Florey Institute for Neuroscience and Mental Health, University of Melbourne, Parkville, Victoria, Australia
- 4/. ORYGEN Research Centre, Parkville, Victoria, Australia
- 5/. Eli Lilly Australia Pty Limited, West Ryde, , NSW, Australia
- 6/. Eli Lilly and Company, Indianapolis, USA
- 7/. Department of Pharmacology and the Institute of Physiology and Neuroscience, Sahlgrenska Academy, Goteborg University, Goteborg, Sweden
- 8/. Eli Lilly and Company, Brussels, Belgium
- 9/. University of California, San Francisco, USA

*Corresponding Author:

Associate Professor Seetal Dodd, Swanston Centre – Barwon Health, PO Box 281, Geelong, Victoria 3220, Australia. Email: seetald@barwonhealth.org.au; Phone: +61 3 4215 3299; Fax: + 61 3 4215 3491

Conflict of Interest

SD has received grant/research support from the Stanley Medical Research Institute, NHMRC, Beyond Blue, ARHRF, Simons Foundation, Geelong MRF, Eli Lilly, Glaxo SmithKline, Organon, Mayne Pharma and Servier, speaker's and advisory board fees from Eli Lilly and conference travel support from Servier. MB has received grant/research support from the Stanley Medical Research Institute, MBF, NHMRC, Beyond Blue, Geelong Medical Research Foundation, ARHRF, Bristol Myers Squibb, Eli Lilly, Glaxo SmithKline, Organon, Novartis, Mayne Pharma, Servier and Astra Zeneca, consultant fees

from Astra Zeneca, Bristol Myers Squibb, Eli Lilly, Glaxo SmithKline, Janssen Cilag, Lundbeck and Pfizer and speaker's fees from Astra Zeneca, Bristol Myers Squibb, Eli Lilly, Glaxo SmithKline, Janssen Cilag, Lundbeck, Organon, Pfizer, Sanofi Synthelabo, Solvay, Wyeth. JCN has received grant/research support from Health Resources and Services Administration and NIMH, and advisory board, consultant, or speaker fees from Bristol-Myers Squibb, Corcept, Eli Lilly, Eli Lilly Global, Forest, Lundbeck, Mylan, Otsuka USA, Otsuka Asia, Pfizer, Shire, and Sunovion. KK, QZ and WD are employees of Eli Lilly.

Contributions

SD, MB, KK, QZ, EE, WD and CN contributed to the conceptualisation of the study, preparation of the manuscript and gave final approval for the manuscript. QZ conducted the statistical analyses.

Accepted manuscript

Abstract

Background: Randomised, placebo-controlled trials of treatments for depression typically collect outcomes data but traditionally only analyse data to demonstrate efficacy and safety. Additional post-hoc statistical techniques may reveal important insights about treatment variables useful when considering inter-individual differences amongst depressed patients. This paper aims to examine the Gradient Boosted Model (GBM), a statistical technique that uses regression tree analyses and can be applied to clinical trial data to identify and measure variables that may influence treatment outcomes.

Methods: GBM was applied to pooled data from 12 randomised clinical trials of 4987 participants experiencing an acute depressive episode who were treated with duloxetine, an SSRI or placebo to predict treatment remission. Additional analyses were conducted on the same dataset using the logistic regression model for comparison between these two methods.

Results: With GBM, there were noticeable differences between treatments when identifying which and to what extent variables were associated with remission. A single logistic regression only revealed a decreasing or increasing relationship between predictors and remission while GBM was able to reveal a complex relationship between predictors and remission.

Limitations: These analyses were conducted post-hoc utilising clinical trials databases. The criteria for constructing the analyses data were based on the characteristics of the clinical trials.

Conclusions: GBM can be used to identify and quantify patient variables that predict remission with specific treatments and has greater flexibility than the logistic regression model. GBM may provide new insights into inter-individual differences in treatment response that may be useful for selecting individualised treatments.

Trial registration: IMPACT clinical trial number 3327; IMPACT clinical trial number 4091; IMPACT clinical trial number 4689; IMPACT clinical trial number 4298; NCT00071695; NCT00062673; NCT00036335; NCT00067912; NCT00073411; NCT00489775; NCT00536471; NCT00666757 (note that trials with IMPACT numbers predate mandatory clinical trial registration requirements)

Key words: randomised clinical trial; gradient boosted method; depression; duloxetine; SSRI

Introduction

Although randomised clinical trials of acute major depressive episodes suggest that all established antidepressant treatments are superior to placebo, these trials also demonstrate that only about one third of individuals achieve full remission, another third respond but retain residual symptoms, and the remaining third of individuals fail to respond to initial treatment (Tranter et al., 2002). Many of these individuals later achieve full remission with subsequent treatment regimens (Gaynes et al., 2012). Recently, there have been increased efforts to investigate inter-individual differences in response to pharmacological and non-pharmacological treatments of depression, as understanding these differences in treatment response is central to progressing personalised medicine goals for the treatment of depression.

Earlier work has shown illness history, especially a history of abuse, and/or comorbidity including alcohol and substance abuse and personality disorder, and/or a previous history of treatment non-response, to be the strongest predictors of non-response to treatment of a current episode of major depression (Dodd et al., 2004). Conversely, shorter periods of illness, no family history of illness and strong social support have been associated with a better prognosis (Trivedi et al., 2005). Riedel et al. (2011) found that lower symptom severity at baseline, episode length of less than 24 months and fewer previous hospitalisations were predictive of remission. Much of this earlier research has focused on identifying predictors of non-response to any treatment and has been of limited use for determining which specific treatment is the best option for each individual patient. However, other work has found significant differences between antidepressant agents. Kilts et al. (2009) found that for participants with greater severity of depression at baseline, change in HAMD24 score was greater for study participants treated with escitalopram compared to pooled antidepressants. Thase et al. (2007) found that duloxetine was superior to fluoxetine or paroxetine for the treatment of moderate to severe depression. In addition, the divergent adverse event profiles of various antidepressant agents have been used to determine best treatment options for individual patients, most notably where weight gain or sexual dysfunction is a concern (Gartlehner et al., 2011).

Personalised medicine goals require knowledge of which treatment is best for each individual, and promises increased efficiencies for health care systems. Much of the emphasis in the development of personalised therapies has been on genetic polymorphisms and blood biomarkers (Ozdemir et al., 2007), although a consistent pattern of data supporting such a predictive capacity is yet to emerge. Currently, determination of clinical variables by interview remains the most useful tool available to a treating physician for determining best treatment options on a case-by-case basis. Further insights into how specific treatments produce different outcomes for differences in clinical, demographic and other patient variables, may provide new information that may benefit personalised treatment goals.

In the past few years, a number of data mining methods have been introduced in clinical research. Decision tree methods (Breiman et al., 1984), which uses a tree-like structure to separate data into subsets with similar outcomes, are particularly attractive because of their ability to handle large number of predictors, their flexibility in model structure, their ability to ignore irrelevant predictors and automatically identify important interactions among predictors. Dietterich (2002) showed that an ensemble of trees gives improved performance over a single tree. One of the more recent and most prominent tree ensemble approaches is the Gradient Boosted Model (GBM) proposed by

Friedman (2001). GBM generally has more accurate prediction than a single decision tree (Hastie et al., 2001), and its ability to provide a measure of variable influence and partial dependence plots showing the relationship between selected predictors and the target variable is another advantage over a single tree. GBM can be applied to clinical datasets to uncover predictors of outcome, and has previously been used elsewhere to uncover predictors of probability of live births with in-vitro fertilisation (Banerjee et al., 2010).

In this study, GBM as implemented by Ridgeway (2007) in R, an open source language for statistical computing and graphics (available at <http://www.r-project.org/>), was applied to 12 randomised clinical trials of duloxetine compared to selective serotonin reuptake inhibitors (SSRIs) and/or placebo for the treatment of acute depression. The aim was to explore the utility of the Gradient Boosted Method in randomised clinical trials. Our goal was to use its variable influence and partial dependence plots to provide insight into differences in treatment response and remission between different antidepressants.

Method

Data was obtained from 12 industry-sponsored, randomised, clinical trials of antidepressant treatment of an acute episode of major depression (Table 1). The 17-item Hamilton Depression Rating Scale (HAMD17) (1967) was the primary outcome measure for each trial. Participants had a HAMD17 total score at baseline of 15 or greater. All trials included a treatment arm in which participants were randomised to receive treatment with duloxetine or another antidepressant and/or placebo. Data from comparator arms using placebo (Goldstein et al., 2002, 2004; Detke et al. 2002a, 2002b; Detke et al., 2004; Perahia et al., 2006; Raskin et al., 2008; Wise et al., 2007; Brannan et al., 2005; Oakes et al., 2012) or any of the selective serotonin reuptake inhibitors (SSRIs), citalopram, fluoxetine (Goldstein et al., 2002; Martinez et al., 2012), paroxetine (Goldstein et al., 2004; Detke et al., 2004; Perahia et al., 2006; Martinez et al., 2012; Lee et al., 2007), sertraline, and escitalopram (Martinez et al., 2012; Nierenberg et al., 2007) were included in the analyses, with subjects receiving any of the latter grouped together and defined as treated with an SSRI. Data for venlafaxine, a comparator arm in two of the trials (Perahia et al., 2008), was excluded as the sample size for venlafaxine treated subjects was much smaller than for duloxetine or SSRI treated subjects. There were data from a total of 4987 participants involved in the 12 trials. Study visits were completed by 81% of these participants, and the analysis included participants who had at least one post-baseline study visit. Duloxetine data from all 12 trials, SSRI data from 6 trials and placebo data from 8 trials were analysed.

The target outcome was remission (HAMD 17 total score ≤ 7) at study endpoint. The outcome value for participants who achieved remission by the study endpoint was coded as one; otherwise, the outcome value was zero. Data was pooled for study arms with the same treatment from the various trials. From each pooled study arm 80% of data was used to build the model and 20% hold out data was used to evaluate the model performance.

GBM was applied to predict the remission outcome based on 14 patient characteristics (age, gender, region, age at onset, BMI, ethnicity, duration of current MDD, number of previous episode and 6 HAMD scales/subscales). GBM is built by combining multiple small trees grown sequentially

(Ridgeway, 2007) to allow a flexible relationship between outcomes and predictors. When a tree is built for GBM, it automatically focuses on the data that previous trees don't predict well and improves the fit to those data. Therefore, GBM offers improved predictive performance.

According to Friedman (2001), tree-based GBM can provide a "relative" variable influence for each input predictor. The variable influence is estimated by averaging contributions of a predictor from different trees. The variable influences in this paper provide an insight to the importance of patient characteristics on treatment remission outcomes.

GBM is a non-parametric model that allows complex model structures; thus, it is unable to use simple numbers to indicate the changing direction of remission outcome by patient characteristics. However, GBM can generate a partial dependence plot to provide a general idea about the relationship between predictor and outcome. In this paper, partial dependence plots for predictors with high variable influence are provided to understand how the treatment remission rate changes by different values of the selected predictor. To account for the uncertainty of the estimation, a confidence interval was estimated and also displayed in the partial dependence plot. A non-parametric bootstrap method was applied to estimate the confidence interval by: (1) drawing a random data set with replacement from the available data; (2) computing the remission rates for specified values of the selected predictor using the given data set from step (1); (3) repeating (1) and (2) for a thousand times; (4) using the 2.5 and 97.5 percentile points for each specified value of the selected predictor obtained from (3).

Analysis using a traditional logistic regression (LR) model was also applied to the same dataset as a comparator. The LR model assumes a linear relationship between the predictors and the log odds of remission. The effect of the predictor is not allowed to change by the value of the predictor.

Results

GBM was applied to 80% of pooled data for each pooled arm (duloxetine, SSRI and placebo) and the holdout 20% data was used to evaluate the general performance of the model. The area under the Receiver Operating Characteristic (ROC) curve was used to measure the model performance based on the holdout data. The predictive power for placebo, duloxetine and SSRI remission was similar but modest. The estimated area under ROC curve was in the range of 0.6 – 0.67.

Variable influence on remission was calculated for placebo, duloxetine and SSRI treatment using GBM analysis for 14 separate variables and plotted as bar graphs (Figures 1a, 1b and 1c). The larger the value of influence, the more important the role the variable had in predicting remission. For each variable, its relative influence is estimated by aggregating all reduction in error due to using the variable as a splitting variable over all trees.

Variable influence differed notably between the three treatment types. Age was the strongest predictor of remission with placebo. It contributed to more than 25% model accuracy. The second influential predictor for placebo remission was the duration of current MDD, which also accounted for about 19% model accuracy. For duloxetine and SSRI treatment, there was no dominant remission predictor, as age was in predicting placebo remission. Their top influence values were less than 17. Moreover, the variable influence on remission had notably different patterns for duloxetine, placebo and SSRI treatment, which is most apparent when displayed graphically in Figure 1. HAMD17 core

symptoms and anxiety subscale scores had more impact on SSRI remission than for placebo. Age was a less important predictor in remission for SSRI treated participants than for duloxetine or placebo treated participants. The influences for BMI and duration of current episode were greater than 10 for all three treatments. Living in the USA was an influence for all treatment arms, especially for treatment with an SSRI.

GBM analyses also permitted the probability of remission with each treatment to be plotted for each variable. These plots are shown in Figure 2. An individual has a greater probability of achieving remission with the significantly superior treatment than with the comparator treatment if the 95% confidence intervals for the two treatments are non-overlapping. Duloxetine and SSRI confidence intervals were frequently non-overlapping when compared to placebo confidence intervals. Duloxetine confidence intervals were non-overlapping with SSRI confidence intervals for length of current episode of depression and for patients with no previous episode of depression. Duloxetine was superior to SSRI treatment for length of current episode of depression in a narrow window corresponding to the duration of current episode of approximately from 10 to 18 weeks; however, there was no significant difference between duloxetine and SSRIs for probability of remission for shorter or longer episodes of current depression outside of this narrow window.

LR model was also applied to the same analysis data to compare with the GBM. LR model assigns a positive or negative parameter to indicate that the remission rate will increase or decrease when the value of the corresponding variable is increasing. In comparison, the GBM is much more flexible as it is not limited to a single direction of increase or decrease. This is clearly demonstrated when both methods were used to investigate the relation between age and placebo remission, where the complexity of the relationship was captured by the GBM but the LR model could only provide the overall decreasing pattern while ignoring the increasing pattern among the young age group (Figure 3). Building the LR model on the subpopulations (age ≤ 27 and age > 27 , the cut-off point was determined by the GBM plot) separately and combining the two models could reflect the relationship similar to the GBM, but the parameters were no longer statistically different from zero.

To further compare LR and GBM in predicting the differences in treatment remission, three subpopulations were defined by the 25 and 75 percentiles of age in the entire data: age ≤ 33 , $33 < \text{age} < 53$ and age ≥ 53 . Similar to the above analyses, 80% of the data were used to build LR and GBM and the developed models were then used to predict treatment remissions using the remaining 20% hold-out data. The predicted differences in treatment remission were calculated by taking the difference in averaged remission rates between treatment groups. The difference between the predicted value and the observed value is the bias from the model. Negative bias indicates the model underestimating the treatment remission, and positive bias indicates the model overestimating the treatment remission. The higher absolute value in bias implies a worse model prediction. Using the bootstrap method similar to the calculation of the confidence interval, a thousand biases were calculated by randomly selecting 80% of data to build the LR and GBM and using the remaining 20% of hold-out data to predict the treatment remissions a thousand times. The averaged bias in predicting remission differences of duloxetine vs. placebo and SSRI vs. placebo for LR and GBM in different age subpopulations and the overall population are provided in Table 2. GBMs had more accurate prediction in treatment difference than LR models in subpopulations and the overall population except for the remission difference between SSRI and placebo in the age ≤ 33

subpopulation. The averaged bias for GBM ranged from -1.1% to 1.8% and the averaged bias for LR model ranged from -2.1% to 3.0%.

Discussion

GBM analysis is a powerful technique for investigating the characteristics of individuals who are most likely to achieve remission with a specific treatment. Our study found that there were differences between the profiles of individuals who are most likely to achieve remission with duloxetine, SSRIs or placebo, with these differences being most pronounced when placebo remitters were compared to those remitting with any of the two active treatments. A previous study of a subset of this dataset found that less severe depressive symptoms, younger age, less anxiety and a shorter current MDD episode were associated with remission with placebo treatment (Nelson et al., 2012). No major differences between active treatments for depression were apparent. This is consistent with research conducted elsewhere that suggests there is considerable overlap in the efficacy profiles of different antidepressants and only relatively minor differences (Gartlehner et al., 2011). GBM analysis appears to be a useful tool to probe clinical trial datasets and has greater flexibility than the traditional LR model. GBM was able to reveal complex relationship between predictors and the remission whereas the LR model only allows monotone decreasing and increasing relationships between predictors and remission. GBM allows not only for the detection of predictive factors but also to compare the impact of these variables for different treatments. This method can be used to give an indication of which values of predictors show differences or similarities between treatments.

GBM has other significant advantages over traditional statistical models. For example, the GBM gives advantages over the LR model when the number of covariates is large, as the tree-based model has no restriction on the size of the data and can detect interaction automatically without pre-specification. Missing data, another example, will be ignored by most traditional parametric models but not by the GBM where all data, including the observations with missing covariate, will be fully utilized.

A methodological concern arises when considering how to apply the GBM to a dataset with multiple treatment arms. In this study, GBM was applied to 14 patient characteristics to predict remission outcomes, which were inspected on the response for placebo, duloxetine, and SSRIs separately. More traditionally, GBM could have been applied to the all three treatment groups (duloxetine, SSRIs, and placebo) using treatment groups as an additional predictor to predict remission outcomes and to inspect for patient characteristics. Neither method is incorrect, but the approach applied here provides greater insight into understanding the impact of patient characteristics on different treatments. The results from these analyses can 1) evaluate the contribution of a specific characteristic to the response of a given treatment when compared with other characteristics, 2) distinguish predictive factors (identifying a population of patients who can get benefit from the treatment) from prognostic factors (providing information on the likely treatment outcome), 3) provide insights of the relation between treatment response and patient characteristics, and 4) be applied as the cornerstone for counterfactual (potential outcomes) approach.

Limitations

These data need to be seen in the context of the methodology used and the sample available and cannot be generalised to other antidepressant agents or psychosocial treatments. However, the results do suggest that GBM analyses have potential as a useful tool to explore these areas. A strength of the study is the large number of subjects in this pooled dataset and the large number of available variables. Limitations include that the GBM analyses were conducted post-hoc. Additional datasets with different clinical variables that may influence treatment response may be required, such as measures of personality traits or more detailed personal histories. These additional datasets may facilitate probing the influence of patient variables on treatment outcomes, and to confirm the utility of the GBM model. It is possible that probability of remission with antidepressant treatments may vary significantly for variables not included in the current study. These may include further demographic data as well as socioeconomic data, quality of life measures, measures of functioning, history of illness data, comorbidity information and personality traits.

Conclusion

In 12 randomised clinical trials of treatment for acute major depression some participants achieved remission in all treatment arms, including placebo. Conventional statistical analytical methods demonstrate superiority for active treatments compared to placebo and equivalent efficacy for the antidepressant treatments included in the study; however, the conventional methods do not deliver great insights into the characteristics of individuals that may increase their probability of remission if assigned to a specific treatment. Using GBM, patient variables can be analysed to determine the extent to which each variable influences the probability of achieving remission with a certain treatment and to distinguish the predictive factors from prognostic factors. The modest findings reported in this paper may reflect limitations in the clinical datasets rather than limitations in the GBM method. The GBM method may identify additional relationships if applied to clinical trial datasets where a greater range of clinical variables have been collected.

Applying the GBM method to datasets from 12 clinical trials of duloxetine, SSRIs and placebo revealed different profiles for patient variables that influence the probability of remission for the three different treatments. Significance testing found that participants with a current episode of depression with a duration of approximately 12 weeks had a greater probability of achieving remission if treated with duloxetine than if treated with an SSRI. As medical practice moves closer towards individualised treatments, identifying optimal treatments for each individual is becoming more important. This study demonstrates that the GBM method is a useful tool for understanding the relationship between patient characteristics and treatment response which may provide further insight in differentiating between different treatment options. This study confirmed the utility of GBM analysis in predicting patterns of treatment response, and suggests that it is a tool that is capable of generating hypotheses from diverse clinical trial datasets.

References

- Banerjee, P., Choi, B., Shahine, L.K., Jun, S.H., O'Leary, K., Lathi, R.B., Westphal, L.M., Wong, W.H., Yao, M.W., 2010. Deep phenotyping to predict live birth outcomes in in vitro fertilization. *Proceedings of the National Academy of Sciences of the United States of America* 107, 13570-13575.
- Brannan, S.K., Mallinckrodt, C.H., Brown, E.B., Wohlreich, M.M., Watkin, J.G., Schatzberg, A.F., 2005. Duloxetine 60 mg once-daily in the treatment of painful physical symptoms in patients with major depressive disorder. *Journal of Psychiatric Research* 39, 43-53.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. *Classification and Regression Trees*. Chapman & Hall/CRC, Boca Raton, FL.
- Detke, M.J., Lu, Y., Goldstein, D.J., Hayes, J.R., Demitrack, M.A., 2002a. Duloxetine, 60 mg once daily, for major depressive disorder: a randomized double-blind placebo-controlled trial. *Journal of Clinical Psychiatry* 63, 308-315.
- Detke, M.J., Lu, Y., Goldstein, D.J., McNamara, R.K., Demitrack, M.A., 2002b. Duloxetine 60 mg once daily dosing versus placebo in the acute treatment of major depression. *Journal of Psychiatric Research* 36, 383-390.
- Detke, M.J., Wiltse, C.G., Mallinckrodt, C.H., McNamara, R.K., Demitrack, M.A., Bitter, I., 2004. Duloxetine in the acute and long-term treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *European Neuropsychopharmacology* 14, 457-470.
- Dietterich, T.G., 2002. Ensemble learning. In: Arbib, M.A. (ed.) *The Handbook of Brain Theory and Neural Networks*, 2nd edition. The MIT Press, Cambridge, MA.

Dodd, S., Berk, M., 2004. Predictors of antidepressant response: A selective review. *International Journal of Psychiatry in Clinical Practice* 8, 91-100.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Annals of Statistics* 29, 1189-1202.

Gartlehner, G., Hansen, R.A., Morgan, L.C., Thaler, K., Lux, L.J., Van Noord, M., Mager, U., Gaynes, B.N., Thieda, P., Strobelberger, M., et al., Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression: An Update of the 2007 Comparative Effectiveness Review. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center, Contract No. 290-2007-10056-I.) AHRQ Publication No. 12-EHC012-EF. Rockville, MD: Agency for Healthcare Research and Quality www.effectivehealthcare.ahrq.gov/reports/final.cfm. 2011.

Gartlehner, G., Hansen, R.A., Morgan, L.C., Thaler, K., Lux, L.J., Van Noord, M., Mager, U., Gaynes, B.N., Thieda, P., Strobelberger, M., et al., Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression: An Update of the 2007 Comparative Effectiveness Review. www.effectivehealthcare.ahrq.gov/reports/final.cfm. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center, Contract No 290-2007-10056-I) AHRQ Publication No 12-EHC012-EF Rockville, MD: Agency for Healthcare Research and Quality 2011

Gaynes, B.N., Dusetzina, S.B., Ellis, A.R., Hansen, R.A., Farley, J.F., Miller, W.C., Sturmer, T., 2012. Treating depression after initial treatment failure: directly comparing switch and augmenting strategies in STAR*D. *Journal of Clinical Psychopharmacology* 32, 114-119.

Goldstein, D.J., Mallinckrodt, C., Lu, Y., Demitrack, M.A., 2002. Duloxetine in the treatment of major depressive disorder: a double-blind clinical trial. *Journal of Clinical Psychiatry* 63, 225-231.

Goldstein, D.J., Lu, Y., Detke, M.J., Wiltse, C., Mallinckrodt, C., Demitrack, M.A., 2004. Duloxetine in the treatment of depression: a double-blind placebo-controlled comparison with paroxetine. *Journal of Clinical Psychopharmacology* 24, 389-399.

Hamilton M., 1967. Development of a rating scale for primary depressive illness. *British Journal of Social Clinical Psychology* 6, 278-296.

Hastie, T., Tibshirani, R., Friedman, J.H., 2001. *The Elements of Statistical Learning*. Springer-Verlag, New York, NY.

Kilts, C.D., Wade, A.G., Andersen, H.F., Schlaepfer, T.E., 2009. Baseline severity of depression predicts antidepressant drug response relative to escitalopram. *Expert Opinions in Pharmacotherapy* 10, 927-936.

Lee, P., Shu, L., Xu, X., Wang, C.Y., Lee, M.S., Liu, C.Y., Hong, J.P., Ruschel, S., Raskin, J., Colman, S.A., Harrison, G.A., 2007. Once-daily duloxetine 60 mg in the treatment of major depressive disorder: multicenter, double-blind, randomized, paroxetine-controlled, non-inferiority trial in China, Korea, Taiwan and Brazil. *Psychiatry and Clinical Neuroscience* 61, 295-307.

Martinez, J.M., Katon, W., Greist, J.H., Kroenke, K., Thase, M.E., Meyers, A.L., Edwards, S.E., Marangell, L.B., Shoemaker, S., Swindle, R., 2012. A pragmatic 12-week, randomized trial of duloxetine versus generic selective serotonin-reuptake inhibitors in the treatment of adult

outpatients in a moderate-to-severe depressive episode. *International Clinical Psychopharmacology* 27, 17-26.

Nelson, J.C., Zhang, Q., Deberdt, W., Marangell, L.B., Karamustafalioglu, O., Lipkovich, I.A., 2012. Predictors of remission with placebo using an integrated study database from patients with major depressive disorder. *Current Medical Research and Opinion* 28, 325-334.

Nierenberg, A.A., Greist, J.H., Mallinckrodt, C.H., Prakash, A., Sambunaris, A., Tollefson, G.D., Wohlreich, M.M., 2007. Duloxetine versus escitalopram and placebo in the treatment of patients with major depressive disorder: onset of antidepressant action, a non-inferiority study. *Current Medical Research and Opinion* 23, 401-416.

Oakes, T.M., Myers, A.L., Marangell, L.B., Ahl, J., Prakash, A., Thase, M.E., Kornstein, S.G., 2012. Assessment of depressive symptoms and functional outcomes in patients with major depressive disorder treated with duloxetine versus placebo: primary outcomes from two trials conducted under the same protocol. *Human Psychopharmacology* 27, 47-56.

Ozdemir, V., Williams-Jones, B., Cooper, D.M., Someya, T., Godard, B., 2007. Mapping translational research in personalized therapeutics: from molecular markers to health policy. *Pharmacogenomics* 8, 177-185.

Perahia, D.G., Wang, F., Mallinckrodt, C.H., Walker, D.J., Detke, M.J., 2006. Duloxetine in the treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *European Psychiatry* 21, 367-378.

Perahia, D.G., Pritchett, Y.L., Kajdasz, D.K., Bauer, M., Jain, R., Russell, J.M., Walker, D.J., Spencer, K.A., Froud, D.M., Raskin, J., Thase, M.E., 2008. A randomized, double-blind comparison of duloxetine and venlafaxine in the treatment of patients with major depressive disorder. *Journal of Psychiatric Research* 42, 22-34.

Raskin, J., Wiltse, C.G., Dinkel, J.J., Walker, D.J., Desai, D., Katona, C., 2008. Safety and tolerability of duloxetine at 60 mg once daily in elderly patients with major depressive disorder. *Journal of Clinical Psychopharmacology* 28, 32-38.

Raskin, J., Xu, J.Y., Kajdasz, D.K., 2008. Time to response for duloxetine 60 mg once daily versus placebo in elderly patients with major depressive disorder. *International Psychogeriatrics* 20, 309-327.

Ridgeway, G., 2007. Generalized Boosted Models: A guide to the GBM package. <http://cran.r-project.org/web/packages/gbm/vignettes/gbm.pdf>.

Riedel, M., Moller, H.J., Obermeier, M., Adli, M., Bauer, M., Kronmuller, K., Brieger, P., Laux, G., Bender, W., Heuser, I., et al., 2011. Clinical predictors of response and remission in inpatients with depressive syndromes. *Journal of Affective Disorders* 133, 137-149.

Thase, M.E., Pritchett, Y.L., Ossanna, M.J., Swindle, R.W., Xu, J., Detke, M.J., 2007. Efficacy of duloxetine and selective serotonin reuptake inhibitors: comparisons as assessed by remission rates in patients with major depressive disorder. *Journal of Clinical Psychopharmacology* 27, 672-676.

Tranter, R., O'Donovan, C., Chandarana, P., Kennedy, S., 2002. Prevalence and outcome of partial remission in depression. *Journal of Psychiatry and Neurosciences*, 27, 241-247.

Trivedi, M.H., Morris, D.W., Pan, J.Y., Grannemann, B.D., John Rush, A., 2005. What moderator characteristics are associated with better prognosis for depression? *Neuropsychiatric Disease Treatment* 1, 51-57.

Wise, T.N., Wiltse, C.G., Iosifescu, D.V., Sheridan, M., Xu, J.Y., Raskin, J., 2007. The safety and tolerability of duloxetine in depressed elderly patients with and without medical comorbidity. *International Journal of Clinical Practice* 61, 1283-1293.

Accepted manuscript

Table 1. Descriptions of the 12 randomised clinical trials (identified by their 4-letter codes) from which data was used for GBM analyses to investigate participant variables that contribute to treatment efficacy.

Study	Study Title	Country	Dosage
HMAQ	Duloxetine versus Placebo in the Treatment of Major Depression	US	Duloxetine 20-60 mg PO BID Fluoxetine 20 mg PO QD Placebo
HMAT	Duloxetine versus Placebo and Paroxetine in the Acute Treatment of Major Depression	US	Duloxetine 20 mg or 40 mg PO BID Paroxetine 20 mg PO QD Placebo
HMAY	Duloxetine versus Placebo and Paroxetine in the Treatment of Major Depression	BG, HR, HU, PL, RO, RU, SK	Duloxetine 40 mg or 60 mg PO BID Paroxetine 20 mg PO QD Placebo
HMBH	Duloxetine Once-Daily Dosing versus Placebo in the Acute Treatment of Major Depression	US	Duloxetine 60 mg PO QD Placebo
HMBU	Duloxetine Versus Venlafaxine Extended Release in the Treatment of Major Depressive Disorder	AT, AU, DE, ES, FR, GB, IT	Duloxetine 60 to 120 mg PO QD Venlafaxine extended release 150 to 225 mg PO QD
HMBV	Duloxetine versus Placebo in the Treatment of Elderly Patients with Major Depressive Disorder	PR, US	Duloxetine 60 mg PO QD Placebo
HMCB	Duloxetine Once-Daily Dosing versus Placebo in Patients with Major Depression and Pain	US	Duloxetine 60 mg PO QD Placebo
HMCQ	Duloxetine Versus Venlafaxine Extended Release in the Treatment of Major Depressive Disorder	CA, US	Duloxetine 60 to 120 mg PO QD Venlafaxine extended release 75 to 225 mg PO QD
HMCR	Duloxetine versus Escitalopram and Placebo in the Treatment of Patients with Major Depression	US	Duloxetine 90 to 120 mg PO QD Escitalopram 20 PO QD
HMCV	Duloxetine versus Paroxetine in the Acute Treatment of Major Depression	BR, CN, KR, TW	Duloxetine 60 mg PO QD Paroxetine 20 mg PO QD
HMFS	Duloxetine Versus Placebo in Patients with Major Depressive Disorder (MDD): Assessment of Energy and Vitality in MDD	PR, US	Duloxetine 60 to 120 mg PO QD Placebo
HMFT	A 12-Week, Randomized, Open-Label Trial of Duloxetine Versus Generic SSRIs in the Treatment of a Severe Depressive Episode	US	Duloxetine 60 to 120 mg PO QD Citalopram 20-40 mg PO QD Fluoxetine 20-80 mg PO QD Paroxetine 20-50 mg PO QD Sertraline 50-200 mg PO QD

US, United States of America; BG, Bulgaria; HR, Croatia; HU, Hungary; RO, Romania; RU, Russia; SK, Slovakia; AT, Austria; AU, Australia; DE, Germany; ES, Spain; FR, France; GB, Great Britain (UK); IT, Italy; PR, Puerto Rico; CA, Canada; BR, Brazil; CN, China; KR, Korea; TW, Taiwan

Table 2. Averaged bias in predicting differences in treatment remission for age subgroups: Logistic Regression Model vs. Gradient Boosted Model.*

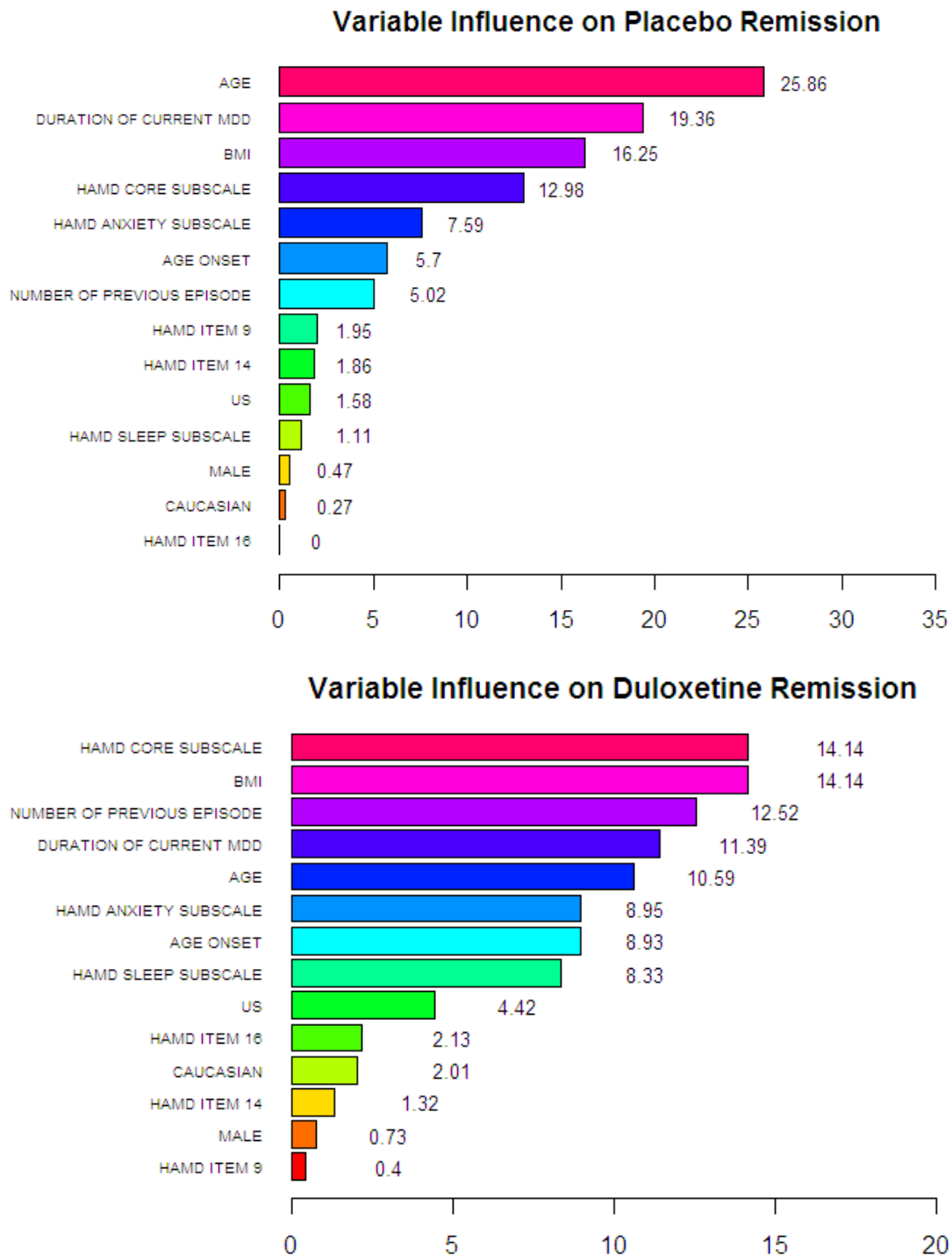
Treatment Difference	Age ≤ 33		33 < Age < 53		Age ≥ 53		ALL	
	DLX vs. PLA	SSRI vs. PLA	DLX vs. PLA	SSRI vs. PLA	DLX vs. PLA	SSRI vs. PLA	DLX vs. PLA	SSRI vs. PLA
LR	.0195	.0051	-.0212	-.0178	.0107	.0302	-.0025	.0004
GBM	.0183	.0080	-.0105	-.0114	-.0082	.0114	-.0018	<.0001

DLX, Duloxetine; GBM, Gradient Boosted Method; LR, Logistic Regression; PLA, placebo; SSRI, selective serotonin reuptake inhibitor

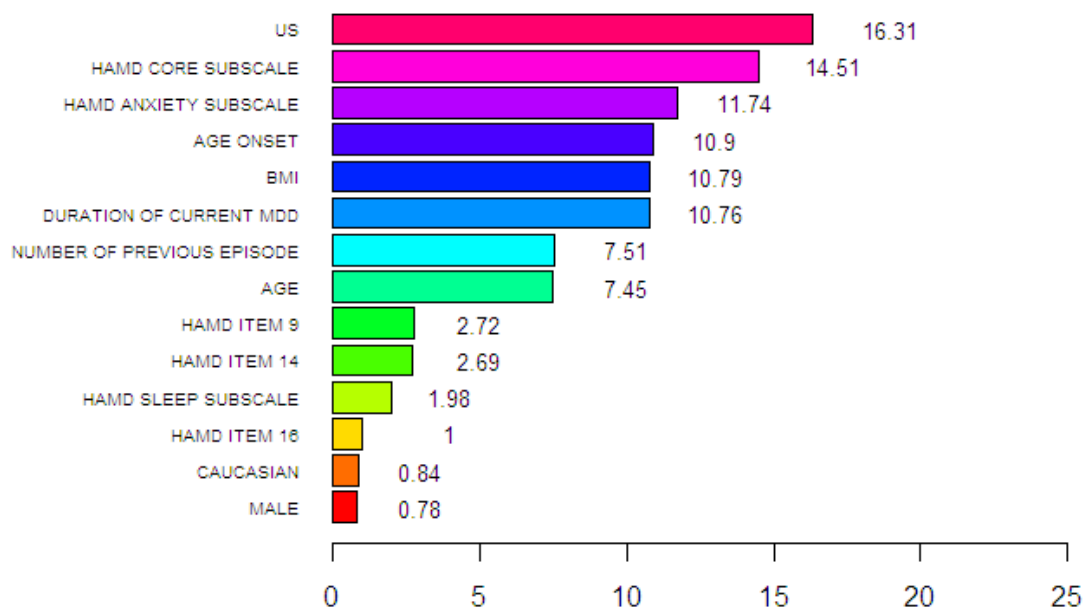
* The values shown in the table were averaged bias from 1000 bootstrap samples

Accepted manuscript

Fig. 1a, 1b and 1c: Variable influence on remission for treatment with placebo, duloxetine or SSRI treatment in major depressive disorder, demonstrating different variable influence profiles for each treatment type.



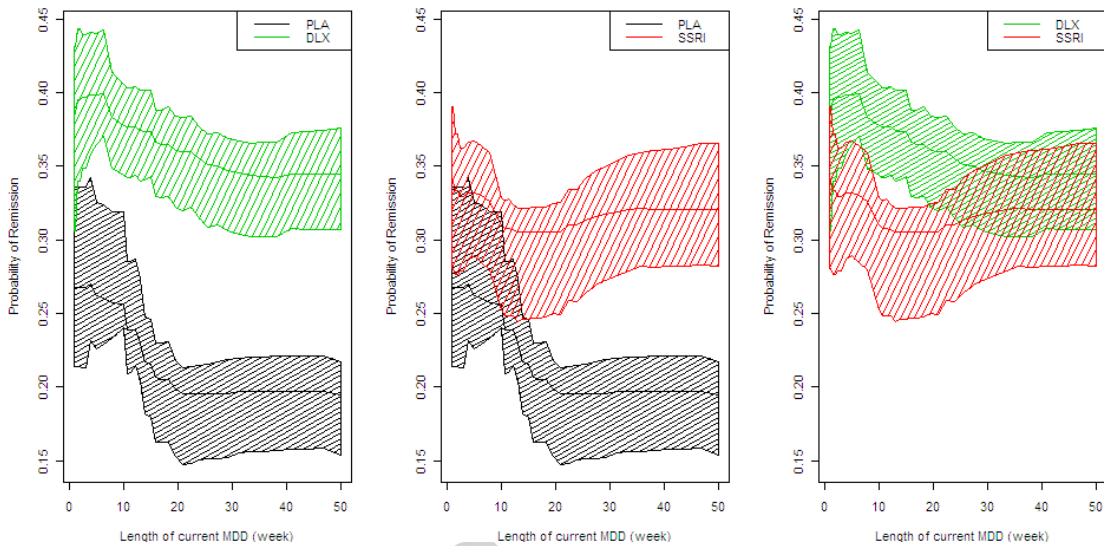
Variable Influence on SSRI Remission



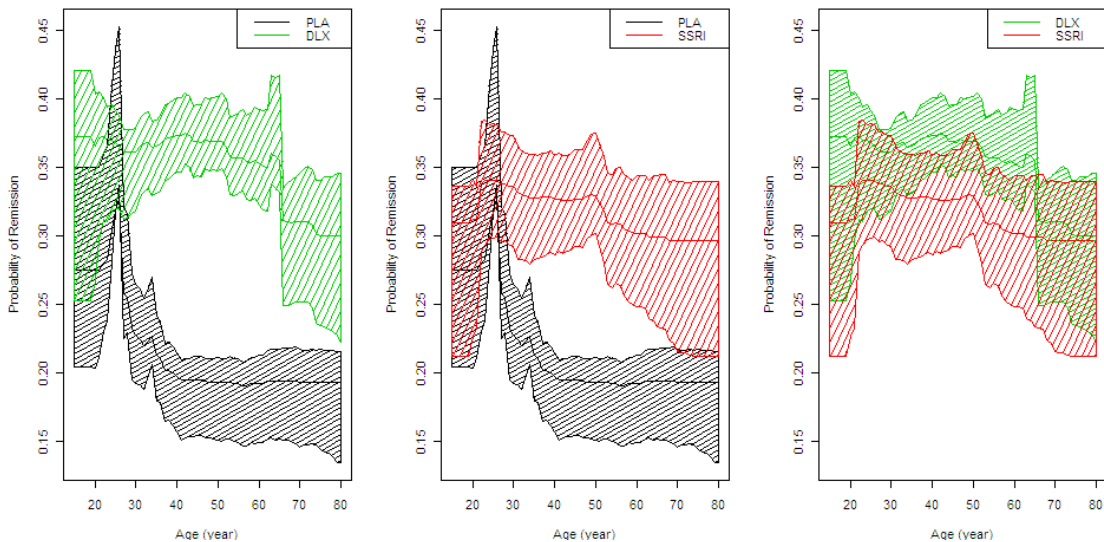
Accepted manu

Fig. 2: Plots comparing placebo, duloxetine and SSRI treatments for probability of remission for each variable of influence. Duloxetine is superior to SSRIs for medium duration length of current MDD (approx. 12 weeks). Duloxetine and SSRIs are superior to placebo over broader ranges of the length of current MDD. Significance occurs when the confidence intervals are non-overlapping.

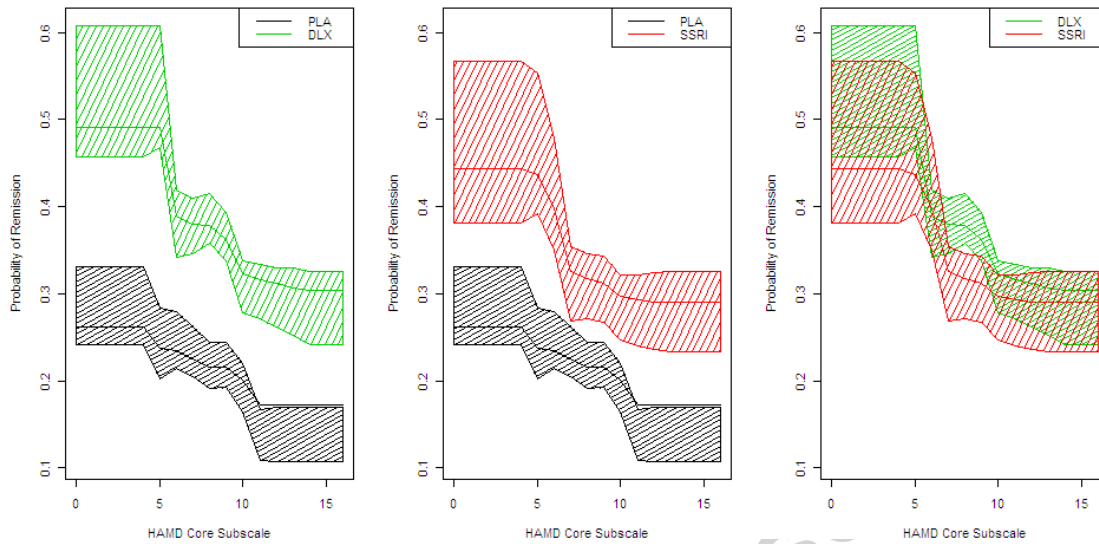
2A: Length of current episode of depression: Duloxetine was superior to SSRIs for medium duration length of current MDD (approximately 10-18 weeks). Duloxetine and SSRIs are superior to placebo over broader ranges of the length of current MDD.



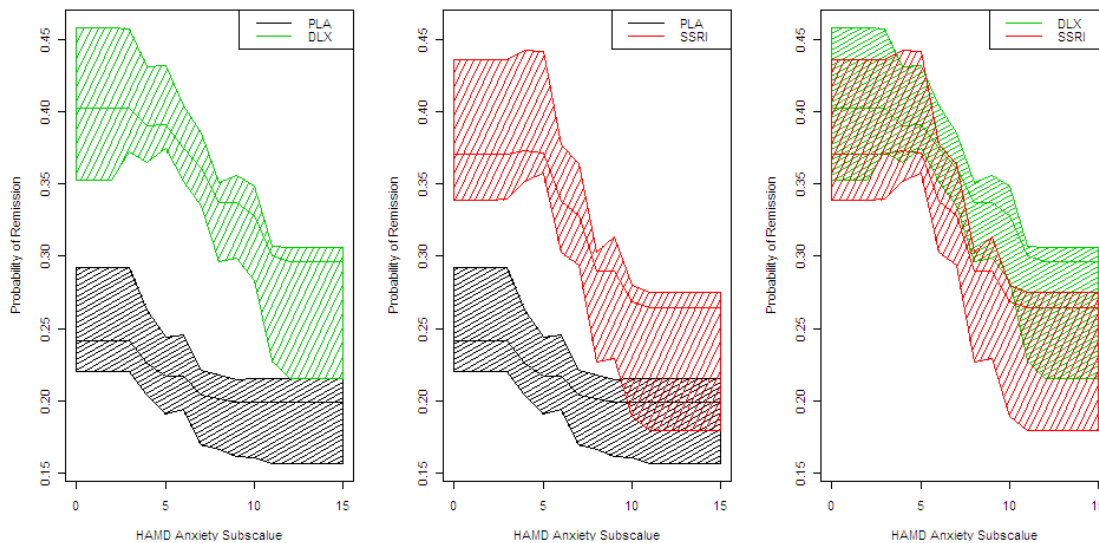
2B: The relationship between age and remission showing that young placebo treated patients are more likely to remit to placebo than older subjects.



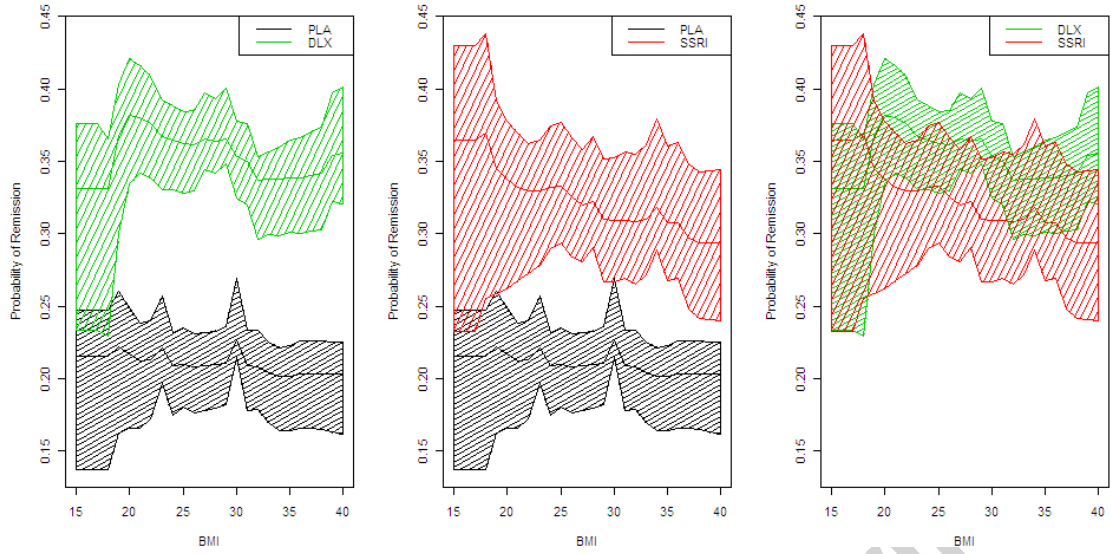
2C: The relationship between HAMD core subscale and remission showing that higher values are associated with a reduced probability of remission for both active treatments and placebo.



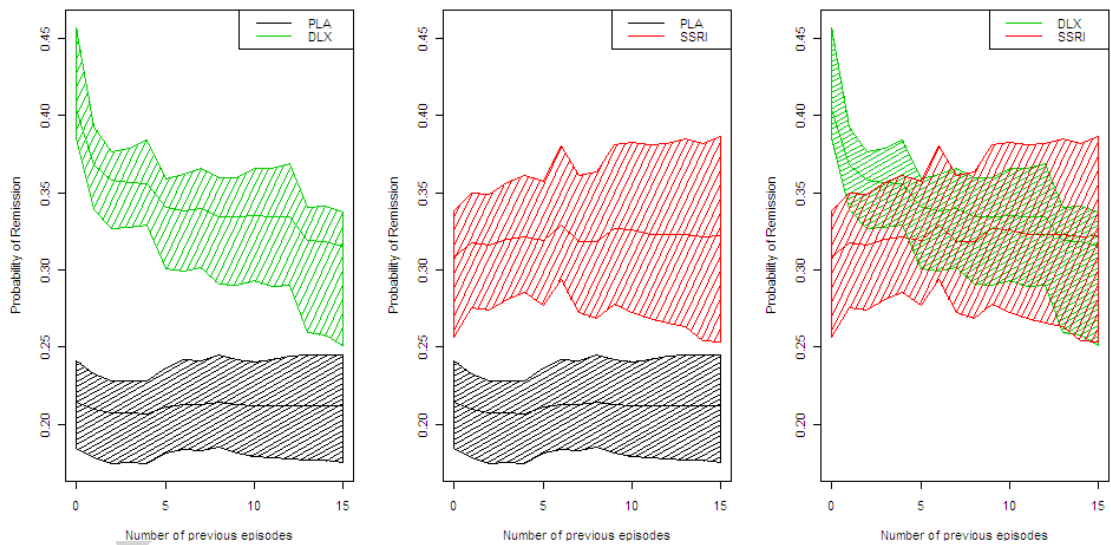
2D: The relationship between HAMD anxiety subscale and remission showing that all curves have a decreasing trend and that duloxetine and SSRI were superior to placebo for scores less than 11 and 9 respectively



2E: The relationship between BMI and remission showing that differences in BMI have little effect on the probability of remission



2F: The relationship between number of previous episode and remission showing duloxetine treated patients with no previous episodes to have a greater probability of remission than SSRI or placebo treated patients



2G: The relationship between remission and US and non-US based participants showing a higher probability for remission for non-US patients and an advantage for duloxetine compared to SSRIs for US based patients

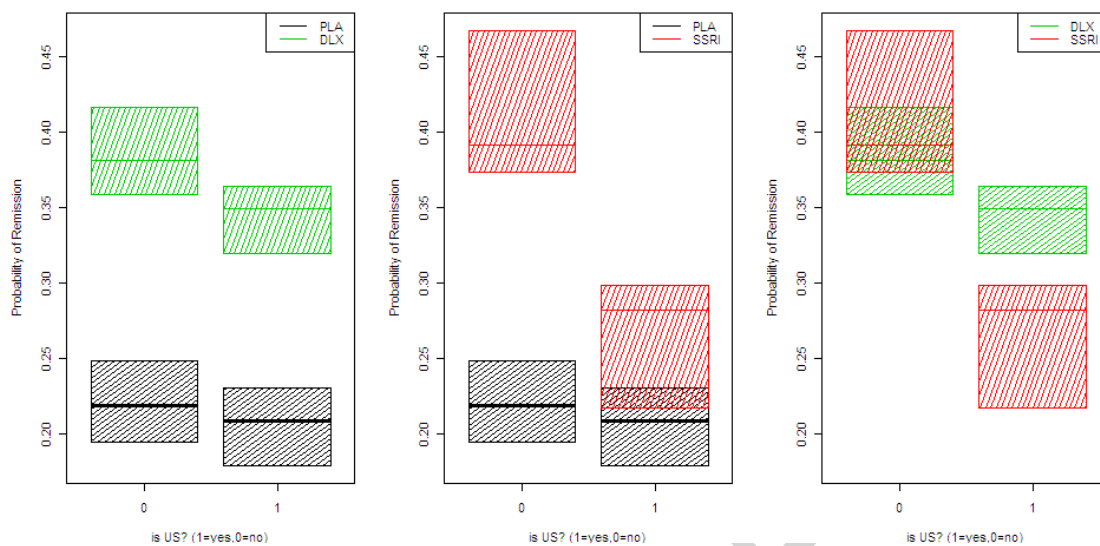
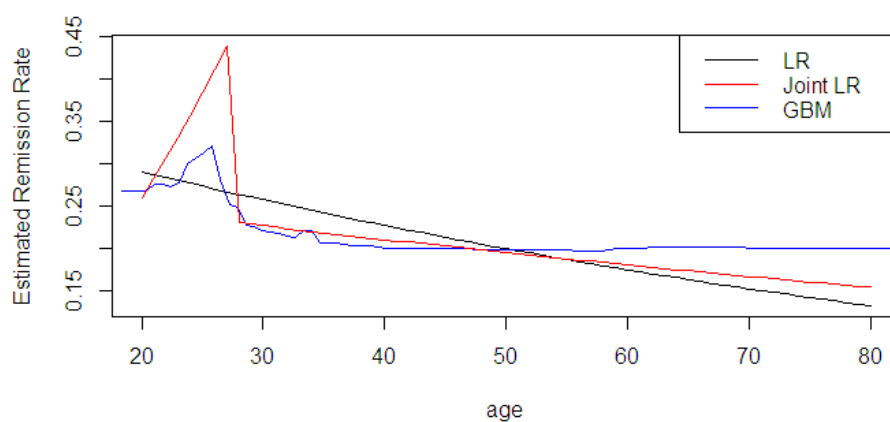


Fig. 3. Age and placebo remission demonstrated by the GBM and LR model. The GBM output for age implies that the relationship is not monotone decreasing.



LR, Logistic regression; Joint LR, Logistic regression by the subpopulation (age ≤ 27 and age >27);
GBM, Gradient Boosted Method

Table 1. Descriptions of the 12 randomised clinical trials (identified by their 4-letter codes) from which data was used for GBM analyses to investigate participant variables that contribute to treatment efficacy.

Study	Study Title	Country	Dosage
HMAQ	Duloxetine versus Placebo in the Treatment of Major Depression	US	Duloxetine 20-60 mg PO BID Fluoxetine 20 mg PO QD Placebo
HMAT	Duloxetine versus Placebo and Paroxetine in the Acute Treatment of Major Depression	US	Duloxetine 20 mg or 40 mg PO BID Paroxetine 20 mg PO QD Placebo
HMAY	Duloxetine versus Placebo and Paroxetine in the Treatment of Major Depression	BG, HR, HU, PL, RO, RU, SK	Duloxetine 40 mg or 60 mg PO BID Paroxetine 20 mg PO QD Placebo
HMBH	Duloxetine Once-Daily Dosing versus Placebo in the Acute Treatment of Major Depression	US	Duloxetine 60 mg PO QD Placebo
HMBU	Duloxetine Versus Venlafaxine Extended Release in the Treatment of Major Depressive Disorder	AT, AU, DE, ES, FR, GB, IT	Duloxetine 60 to 120 mg PO QD Venlafaxine extended release 150 to 225 mg PO QD
HMBV	Duloxetine versus Placebo in the Treatment of Elderly Patients with Major Depressive Disorder	PR, US	Duloxetine 60 mg PO QD Placebo
HMCB	Duloxetine Once-Daily Dosing versus Placebo in Patients with Major Depression and Pain	US	Duloxetine 60 mg PO QD Placebo
HMCQ	Duloxetine Versus Venlafaxine Extended Release in the Treatment of Major Depressive Disorder	CA, US	Duloxetine 60 to 120 mg PO QD Venlafaxine extended release 75 to 225 mg PO QD
HMCR	Duloxetine versus Escitalopram and Placebo in the Treatment of Patients with Major Depression	US	Duloxetine 90 to 120 mg PO QD Escitalopram 20 PO QD
HMCV	Duloxetine versus Paroxetine in the Acute Treatment of Major Depression	BR, CN, KR, TW	Duloxetine 60 mg PO QD Paroxetine 20 mg PO QD
HMFS	Duloxetine Versus Placebo in Patients with Major Depressive Disorder (MDD): Assessment of Energy and Vitality in MDD	PR, US	Duloxetine 60 to 120 mg PO QD Placebo
HMFT	A 12-Week, Randomized, Open-Label Trial of Duloxetine Versus Generic SSRIs in the Treatment of a Severe Depressive Episode	US	Duloxetine 60 to 120 mg PO QD Citalopram 20-40 mg PO QD Fluoxetine 20-80 mg PO QD Paroxetine 20-50 mg PO QD Sertraline 50-200 mg PO QD

US, United States of America; BG, Bulgaria; HR, Croatia; HU, Hungary; RO, Romania; RU, Russia; SK, Slovakia; AT, Austria; AU, Australia; DE, Germany; ES, Spain; FR, France; GB, Great Britain (UK); IT, Italy; PR, Puerto Rico; CA, Canada; BR, Brazil; CN, China; KR, Korea; TW, Taiwan

Table 2. Model accuracy comparisons for Logistic Regression Model and Gradient Boosted Method

	Prediction Error ^a Mean (Sd)*			AUC ^b Mean (Sd)*		
	PLA	DLX	SSRI	PLA	DLX	SSRI
LR	0.2225 (0.0298)	0.3635 (0.0251)	0.3364 (0.0347)	0.6409 (0.0484)	0.6096 (0.0288)	0.6387 (0.0429)
GBM	0.2199 (0.0301)	0.3671 (0.0251)	0.3339 (0.0356)	0.6206 (0.0456)	0.6024 (0.0296)	0.6317 (0.0429)

AUC, area under the curve; PLA, placebo; DLX, duloxetine; SSRI, selective serotonin reuptake inhibitor; LR, logistic regression; GBM, Gradient Boosted Method

^a prediction error: the lower the better

^b AUC: the higher the better

* Mean and standard deviation were calculated from 1000 resamples

Table 3. Coefficients for LR and joint LR model for the relationship between placebo remission and age.

	Coefficient for age
LR	-0.0165*
JLR	0.1148 (age \leq 27#) -0.0095 (age $>$ 27#)

LR, logistic regression; JLR, joint logistic regression

* significant at 0.05

the age cut-off point was determined by GBM plot

Role of the Funding Source

This manuscript was funded by Eli Lilly.

Contributions

SD, MB, KK, QZ, EE, WD and CN contributed to the conceptualisation of the study, preparation of the manuscript and gave final approval for the manuscript. QZ conducted the statistical analyses.

Conflict of Interest

SD has received grant/research support from the Stanley Medical Research Institute, NHMRC, Beyond Blue, ARHRF, Simons Foundation, Geelong MRF, Eli Lilly, Glaxo SmithKline, Organon, Mayne Pharma and Servier, speaker's and advisory board fees from Eli Lilly and conference travel support from Servier. MB has received grant/research support from the Stanley Medical Research Institute, MBF, NHMRC, Beyond Blue, Geelong Medical Research Foundation, ARHRF, Bristol Myers Squibb, Eli Lilly, Glaxo SmithKline, Organon, Novartis, Mayne Pharma, Servier and Astra Zeneca, consultant fees from Astra Zeneca, Bristol Myers Squibb, Eli Lilly, Glaxo SmithKline, Janssen Cilag, Lundbeck and Pfizer and speaker's fees from Astra Zeneca, Bristol Myers Squibb, Eli Lilly, Glaxo SmithKline, Janssen Cilag, Lundbeck, Organon, Pfizer, Sanofi Synthelabo, Solvay, Wyeth. JCN has received grant/research support from Health Resources and Services Administration and NIMH, and advisory board, consultant, or speaker fees from Bristol-Myers Squibb, Corcept, Eli Lilly, Eli Lilly Global, Forest, Lundbeck, Mylan, Otsuka USA, Otsuka Asia, Pfizer, Shire, and Sunovion. KK, QZ and WD are employees of Eli Lilly.

Acknowledgment

No acknowledgements.

Table 2. Averaged bias in predicting differences in treatment remission for age subgroups: Logistic Regression Model vs. Gradient Boosted Model.*

Treatment Difference	Age ≤ 33		33 < Age < 53		Age ≥ 53		ALL	
	DLX vs. PLA	SSRI vs. PLA	DLX vs. PLA	SSRI vs. PLA	DLX vs. PLA	SSRI vs. PLA	DLX vs. PLA	SSRI vs. PLA
LR	.0195	.0051	-.0212	-.0178	.0107	.0302	-.0025	.0004
GBM	.0183	.0080	-.0105	-.0114	-.0082	.0114	-.0018	<.0001

DLX, Duloxetine; GBM, Gradient Boosted Method; LR, Logistic Regression; PLA, placebo; SSRI, selective serotonin reuptake inhibitor

* The values shown in the table were averaged bias from 1000 bootstrap samples