OXFORD

REGULAR RESEARCH ARTICLE

Analysis of Time-Course, Dose-Effect, and Influencing Factors of Antidepressants in the Treatment of Acute Adult Patients With Major Depression

Qingqing Cheng, Jihan Huang, Ling Xu, Yunfei Li, Huafang Li, Yifeng Shen, Qingshan Zheng, Lujin Li

Center for Drug of Clinical Research, Shanghai University of Traditional Chinese Medicine, Shanghai, China (Ms Cheng, Dr Huang, Dr Xu, Dr Li, Dr Zheng, and Dr Li); Shanghai Mental Health Center, Shanghai, China (Drs Li and Shen).

Correspondence: Qingshan Zheng, PhD, No. 1200 Cailun Road, Shanghai, 201203, China (qingshan.zheng@drugchina.net) and Lujin Li, No. 1200 Cailun Road, Shanghai, 201203, China (lilujin666@163.com).
Q.C. and J.H. contributed equally to this work.

Abstract

Objective: Model-based meta-analysis was used to describe the time-course and dose-effect relationships of antidepressants and also simultaneously investigate the impact of various factors on drug efficacy.
Methods: This study is a reanalysis of a published network meta-analysis. Only placebo-controlled trials were included in this study. The change rate in depression rating scale scores from baseline was used as an efficacy indicator because a continuous variable is more likely to reflect subtle differences in efficacy between drugs.
Results: A total 230 studies containing 64346 patients were included in the analysis. The results showed that the number of study sites (single or multi-center) and the type of setting (inpatient or noninpatient) are important factors affecting the efficacy of antidepressants. After deducting the placebo effect, the maximum pure drug efficacy value of inpatients was 18.4% higher than that of noninpatients, and maximum pure drug efficacy value of single-center trials was 10.2% higher than that of multi-central trials. Amitriptyline showed the highest drug efficacy. The remaining 18 antidepressants were comparable or had little difference. Within the approved dose range, no significant dose-response relationship was observed. However, the time-course relationship is obvious for all antidepressants. In terms of safety, with the exception of amitriptyline, the dropout rate due to adverse events of other drugs was not more than 10% higher than that of the placebo group.
Conclusion: The number of study sites and the type of setting are significant impact factors for the efficacy of antidepressants. Except for amitriptyline, the other 18 antidepressants have little difference in efficacy and safety.

Keywords: antidepressant, efficacy, model-based meta-analysis

Significance Statement

Model-based meta-analysis (MBMA) is an important method for model informed drug discovery and development. This study not only involved a comprehensive quantitative evaluation of the efficacy of antidepressants but also described the time-course and dose-effect relationships of antidepressants and also simultaneously investigated the impact of various factors on drug efficacy using MBMA to provide necessary quantitative information for the current clinical practice guidelines of depression.

Introduction

The World Health Organization states that the rates of depression have risen by more than 18% during the past decade, and it is predicted to be the leading cause of disease burden by 2030 (Deardorff and Grossberg, 2014; Papadimitropoulou et al., 2017). Currently, commonly used antidepressants include selective serotonin reuptake inhibitors (SSRIs) (Ioannidis, 2008), serotonin-norepinephrine reuptake inhibitors (Amick et al., 2015), selective norepinephrine reuptake inhibitors (Clayton et al., 2003), noradrenergic antagonist-specific serotonin antagonists (Santarsieri and Schwartz, 2015), serotonin-modulating antidepressants, norepinephrine-dopamine reuptake inhibitors (Wang et al., 2016), etc. In the face of so many antidepressants, good evidence is needed to guide clinicians to make the best decisions in selecting which medication to prescribe (Amick et al., 2015).

A published network meta-analysis systematically compared the efficacy of 21 antidepressants (Cipriani et al., 2018). This network meta-analysis has the most abundant data in the field so far. However, this study has limitations created by the methodology of network meta-analysis. First, the efficacy data were obtained at different endpoints (ranging from 4 to 12 weeks) and were combined for analysis in this study, neglecting the effect of time on treatment efficacy. Second, the studies used response rates (defined as ≥50% reduction in initial depression rating-scale scores) as the primary outcome (Cleare et al., 2015), but this binary index will lose a lot of useful information compared with a continuous index (Khoo et al., 2015; Jakobsen et al., 2017). For example, a person who improves by 50% is called a responder, whereas one who improves by 49% is called a nonresponder, thus inflating the apparent difference between these patients. Third, this study did not distinguish between placebo-controlled trials and comparator-controlled trials. Many studies have shown that the efficacy of antidepressant drugs in a comparator-controlled trial is higher than that of a placebo-controlled trial (Rutherford et al., 2009); thus, the mixed analyses of these 2 types of trials may cause bias. In view of the above limitations, it is necessary to use a new method to reanalyze the data.

Model-based meta-analysis (MBMA) is an important method for model-informed drug discovery and development (Lalonde et al., 2007). MBMA can accurately describe the time-course and dose-effect relationships of drugs and can simultaneously investigate the impact of various factors on the efficacy parameters. Compared with a traditional meta-analysis, MBMA can make full use of the efficacy data at each time point (Boucher and Bennetts, 2016). Based on data shared by Dr Andrea Cipriani (Cipriani et al., 2018), this study involved a comprehensive quantitative evaluation of the efficacy of antidepressants using MBMA to provide necessary quantitative information for the current clinical practice guidelines of depression.

Methods

Data Source and Handling

The data for this study were derived from the database shared by Dr Andrea Cipriani (Cipriani et al., 2018). To correct for the heterogeneity of placebo response among the trials, only data from placebo randomized controlled trials were included for analysis. Because only the efficacy data at the end of the treatment were shown in the published database, the efficacy data at each time point during the trial period were added by a review of the original literature. If the efficacy data were presented as a graph, the digitizing software Engauge Digitizer (version 4.1) was used for data extraction. All the efficacy data were independently extracted by 2 researchers (Q.Q. Cheng and J.H. Huang), and the inconsistencies were determined by the third researcher (L. Xu). The data extraction error should not exceed 2% when reading the graph. If it exceeds 2%, the graph must be reread and the mean value of the 2 values is used as the final extracted value.

Various rating scales for depression were used, such as the Hamilton Depression Rating Scale (HAMD) (17 items, 21 items, 22 items, 24 items, 31 items which are five versions of HAMD), the Montgomery–Åsberg Depression Rating Scale, and the Inventory of Depressive Symptomatology-Self Report. To make baseline scores of different rating scales comparable, baseline scores were standardized as the percentage of maximum attainable scores on each rating scale. If the version of the HAMD in a study was not specified, the most commonly used 17-item version was used by default.

Risk of Bias Assessment

The risk of bias of the included studies has been assessed using the Cochrane Risk of Bias Tool in the database shared by Dr Cipriani. These studies were graded as high, moderate, or low quality based on the following criteria: (1) studies were considered to be of high quality when both randomization and allocation concealment were assessed as low risks of bias, and all other items were assessed as low or unclear risk of bias in a trial; (2) studies were considered to be of low quality if either randomization or allocation concealment was assessed as a high risk of bias, regardless of the risk of other items; and (3) studies were considered to be of moderate quality if they did not meet the criteria for high or low quality.

Model Building

The change rate in the rating scale score from the baseline was used as the evaluation index for modeling, thus eliminating the potential baseline impacts on the evaluation of drug efficacy. The change rate in rating scale score from baseline increased with time and finally reached the plateau, which was in line with the E_{max} model. The E_{max} model is widely used to describe the time course of placebo response (Formula 1) and drug efficacy (Formula 2). In this study, we assumed that the efficacy of the drug group consists of the placebo response and pure drug efficacy (Formula 3). Here, the pure drug efficacy is the relative efficacy of the drug, which is the efficacy after subtracting the corresponding placebo response.

$$E_{placebo,i,m} = -\frac{(E_{\max,placebo,i} + \eta_{1,i}{}^{study}) \times Time_m}{(ET_{50,placebo,i} + \eta_{2,i}{}^{study}) + Time_m} \quad (1)$$

$$E_{drug,i,j,m} = -\frac{(E_{max,drug,i,j} + \eta_{3,i}^{study} + \eta_{5,i,j}^{arm}) \times Time_m}{(ET_{50,drug,i,j} + \eta_{4,i}^{study} + \eta_{6,i,j}^{arm}) + Time_m} \quad (2)$$

$$E_{i,j,m} = E_{placebo,i,m} + E_{drug,i,j,m} + \frac{\varepsilon_{i,j,m}}{\sqrt{N_{i,j,m}/100}} \quad (3)$$

In Formula 1, $E_{max, placebo,i}$ is the maximum possible efficacy of placebo group in the ith study, $ET_{50, placebo,i}$ is the time to achieve 50% of $E_{max, placebo,i}$, which represents the speed of onset. Time ($Time_m$) represents the time of the mth observation time point (week). The inter-study variabilities ($\eta_{1,i}^{study}$ and $\eta_{2,i}^{study}$) of $E_{max, placebo,i}$ and $ET_{50,placebo,i}$ are assumed to be normally distributed with a mean of 0 and variances of ω_1^2 and ω_2^2, respectively.

In Formula 2, $E_{max, drug, i, j}$ is the maximum possible efficacy of the pure drug efficacy of the jth drug group in the ith study, and $ET_{50, drug,i,j}$ is the time to achieve 50% of $E_{max, drug,i,j}$. Time ($Time_m$) is the time of the mth observation time point (week). The inter-study variabilities ($\eta_{3,i}^{study}$ and $\eta_{4,i}^{study}$) of $E_{max, drug, i, j}$, and $ET_{50,drug,i,j}$ conform to the normal distribution with a mean of 0 and variances of ω_3^2 and ω_4^2, respectively. The inter-arm variabilities ($\eta_{5, i, j}^{arm}$ and $\eta_{6, i, j}^{arm}$) of $E_{max, drug, i, j}$ and $ET_{50,drug,i,j}$ conform to the normal distribution with a mean of 0 and variances of ω_5^2 and ω_6^2, respectively. It should be noted that for the stability of the model, inter-study variability and inter-arm variability of parameters of the final model may not be introduced simultaneously.

In Formula 3, $E_{i,j,m}$ is the observed efficacy at the mth time point of the jth drug group in the ith study, $E_{placebo,i,m}$ is the efficacy of the mth observation point of placebo group in the ith study, and $E_{drug,i,j,m}$ is the pure drug efficacy of the mth observation point of the jth drug group in the ith study. The residual error ($\varepsilon_{i,j,m}$) of the mth observation point of the jth drug group in the ith study will be weighted by the inverse of the square root of the sample size ($N_{i,j,m}$) normalized to 100 patients. $\varepsilon_{i,j,m}$ is assumed to be normally distributed with a mean of 0 and variance of σ^2.

The factors potentially affecting the pharmacodynamic parameters were evaluated, including age, female ratio, sample size per study, dosing schedule (fixed or flexible), type of scale ratings, type of setting (inpatient or noninpatient), standardized baseline score of scale ratings, sources of funding (industry or not), number of study sites, use of a placebo run-in period or not, and years of reporting. Categorical variables were tested using the proportional model (Formula 4), and continuous covariates were normalized to the median value and evaluated using both linear (Formula 5) and power (Formula 6) models.

$$P_{pop} = P_{Typical} + COV \times \theta_{cov} \quad (4)$$

$$P_{pop} = P_{Typical} + (COV - COV_{median}) \cdot \theta_{cov} \quad (5)$$

$$P_{pop} = P_{Typical} \times (COV/COV_{median})^{\theta\, cov} \quad (6)$$

In Formulas 4–6, P_{pop} is the model parameter under different covariate levels. The typical value of the model parameter ($P_{Typical}$) is the model parameter when the bivariate variable is defined as 0 or the continuous variable is defined as the median value. COV is the covariate value. COV_{median} is the median of

the covariate, and θ_{cov} is the covariate correction factor for the model parameters.

Nested models in covariate screening were compared statistically using a likelihood ratio test on the differences in the objective function value (OFV). The change was considered significant if the decrease in OFV was >3.84 (χ^2, df=1, P<.05) for the forward inclusion step, and the increase in OFV was >6.63 (χ^2, df=1, P<.01) for the backward elimination step (An et al., 2017). Missing covariate values were imputed by the median value.

Model Evaluation

Models were evaluated by scientific plausibility, the OFV parameter precision, and goodness-of-fit plots. A visual predictive check was performed to determine whether the fitted model provided an adequate description of the data (Nguyen et al., 2017). In this predictive check, 1000 datasets were simulated from the final model. The median, lower (2.5%), and upper (97.5%) quantiles of the simulated drug efficacy were then compared with the observed efficacy value. The robustness of the model was assessed using a nonparametric bootstrap with repetition of 1000 Nonlinear Mixed Effects Modeling (NONMEM) runs of the final model. The bootstrap median parameter values and the percentile bootstrap 95% confidence intervals (CI) were compared with the respective values estimated from the final model.

Typical Efficacy Analysis

After the final model is established, the Bayesian feedback algorithm will be used to obtain the individual parameter values of each study arm. When covariates that affect efficacy are found, the efficacy parameters need to be corrected to ensure that the drug efficacy are compared at the same covariant level. The different covariate model correction methods are as follows:

$$P_{i,corrected} = P_i - (COV - COV_{median}) \times \theta_{cov} \quad (7)$$

$$P_{i,corrected} = P_i \times (COV/COV_{median})^{-\theta_{cov}} \quad (8)$$

$$P_{i,corrected} = P_i - COV \times \theta_{cov} \quad (9)$$

In Formulas 7–9, P_i is the individual parameter value of the ith study arm obtained from the Bayesian feedback. $P_{i, corrected}$ is the corrected individual parameter of the ith study arm. COV is the covariate value of the ith study arm. COV_{median} is the median value of the covariate of all study arms. θ_{cov} is the correction coefficient of the covariate for the parameter.

The random effects model in the single-arm meta-analysis was then used to synthesize the corrected parameter values for each type of drug. Based on the distribution of parameters of each type of drug, 10000 Monte Carlo simulations were conducted to estimate the distribution of efficacy at each point in time for each type of drug.

Software

Model development and Bayesian feedback were performed by NONMEM (Version 7.4; Icon Inc, PA, USA). All model parameters were estimated using the first-order conditional estimation method with η-ε interaction in NONMEM. R (Version 3.4.4;

http://www.r-project.org) was used for visual diagnosis and model simulation. Meta-analysis was performed by Stata software (version 13.1, StataCorp LP, College Station, TX).

Results

Characteristics of Included Studies

A total of 230 studies containing 585 arms with 64 346 patients were included for analysis (supplementary Figure 1). A total 1682 data points were used for modeling, of which 1109 data points (65.9%) were additional data points outside the public database by Dr Andrea Cipriani.

Nineteen drugs were included for analysis: agomelatine (10 arms, n = 1481), amitriptyline (26 arms, n = 1582), bupropion (23 arms, n = 2467), citalopram (17 arms, n = 2061), desvenlafaxine (13 arms, n = 2645), duloxetine (22 arms, n = 2690), escitalopram (20 arms, n = 3032), fluoxetine (36 arms, n = 3307), fluvoxamine (7 arms, n = 518), levomilnacipran (8 arms, n = 1603), mirtazapine (15 arms, n = 825), nefazodone (13 arms, n = 809), paroxetine (44 arms, n = 4782), reboxetine (11 arms, n = 1332), sertraline (21 arms, n = 1948), trazodone (6 arms, n = 570), venlafaxine (27 arms, n = 2774), vilazodone (12 arms, n = 2264), and vortioxetine (24 arms, n = 3765). The information was shown in Figure 1. Because there was only 1 placebo-controlled trial of clomipramine and no placebo-controlled trial of milnacipran, these 2 drugs were not included for analysis.

In the included studies, the sample size of each arm ranged from 7 to 357 (median, 103). The treatment duration ranged from 4 to 12 weeks (median, 8 weeks). The mean age ranged from 31.9 to 56 years (median, 41 years). The percentages of female participants ranged from 10% to 100% (median, 63%). The standardized scores of rating scales in baseline ranged from 0.24 to 0.85 (median, 0.44). There were 289 (81.4%) drug arms from multi-center clinical trials, 257 (72.4%) drug arms sponsored by industry, 181 (51.0%) drug arms were given at a fixed-dose, and 208 (58.6%) drug arms used placebo run-in design. The detailed information of the included studies is shown in supplementary Table 1.

The bias risk assessments of the included studies are shown in Supplementary Table 2. According to the criteria of literature quality, 8.3% of the studies were rated as high quality, and the remaining 91.7% of studies were graded as moderate quality. Overall, the quality of the involved studies was relatively high.

Model Establishment and Assessment

The estimated model parameters are listed in Table 1. In the process of covariate investigation, we found that the sample size per study, the publication year (published before or after the year 2000), the number of study sites (single or multi-center), and type of setting (inpatient or noninpatient) had a significant effect on the parameter of $E_{max, drug}$. The results of subgroup analysis also showed that the above factors had a great influence on drug pure efficacy at week 8 (Table 2). It should be noted that the sample size per study was correlated with the publication year and the number of study sites such that the sample size was smaller in the trials before 2000 or in the single-center trials (supplementary Figure 2). After being screened by the forward inclusion and backward elimination steps, only the type of setting (inpatient or noninpatient) and the number of study sites (single or multi-center) were included in the covariate model. The final covariate model is expressed as follows:

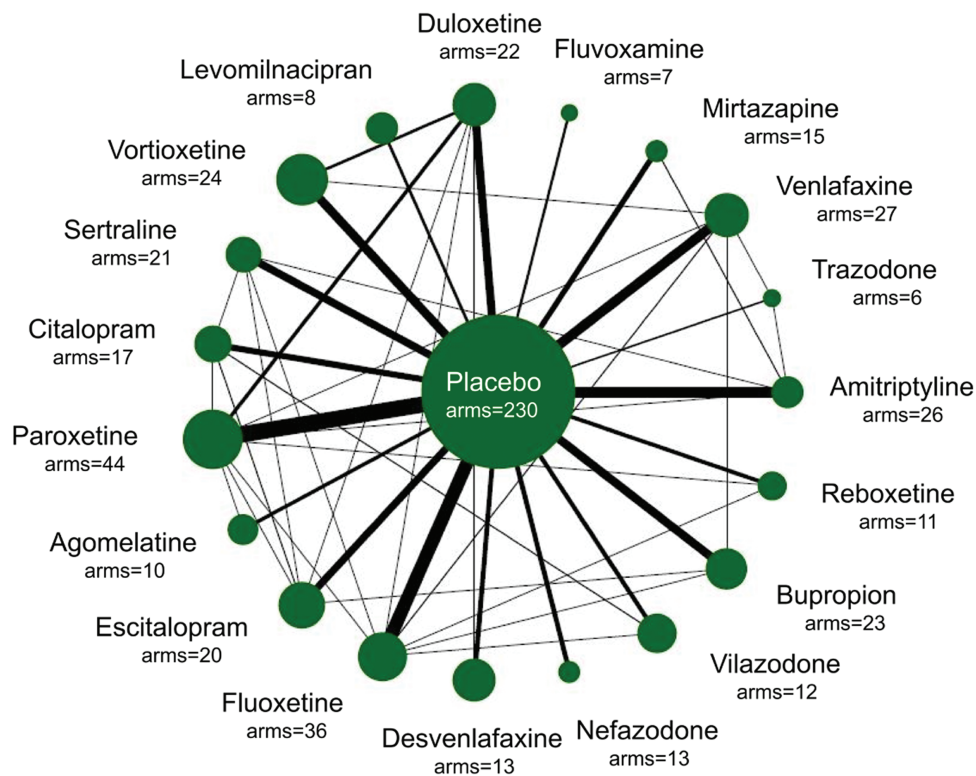$$E\max, drug = 16.4 + Type \times 18.4 + Center \times 10.2 \qquad (10)$$

Figure 1. Model-based meta-analysis of eligible comparisons for efficacy. Width of the lines is proportional to the number of arms comparing every pair of treatments. Size of every circle is proportional to the number of randomly assigned participants (i.e., sample size).

Table 1. Parameter estimation

Pharmacodynamic parameters	Estimate (95% CI)	991 Bootstrapped median (95% CI)
$E_{max-placebo}$,%	58.1 (54.2~62.0)	57.5 (50.5~62.0)
$ET_{50-placebo}$,wk	3.83 (3.24~4.42)	3.73 (1.14~4.46)
$E_{max-drug}$,%	16.4 (13.4~19.5)	15.8 (0.65~21.7)
$ET_{50-drug}$,week	5.36 (3.55~7.17)	5.45 (3.57~9.17)
θ_{Status} on E_{max},%	18.4 (7.94~28.9)	20.30(7.71~45.15)
θ_{Centre} on E_{max},%	10.2 (4.26~16.1)	11.40(3.84~36.93)
Variability parameters		
η ($E_{max-placebo}$),%	14.2 (11.4~17.0)	14.3 (11.4~17.3)
η ($ET_{50-placebo}$),%	1.68 (1.31~2.05)	1.73 (1.30~3.01)
η ($ET_{50-drug}$),%	2.56 (1.79~3.33)	2.45 (0.00~3.57)
ε ,%	2.54 (2.33~3.75)	2.53 (2.34~2.76)

Abbreviations: $E_{max-drug}$, maximal pure effect of drug; $E_{max-placebo}$, maximal effect of placebo; $ET_{50-drug}$, time to achieve 50% of $E_{max-drug}$; $ET_{50-placebo}$, time to achieve 50% of $E_{max-placebo}$; η, variability of pharmacodynamic parameter; ε, residual error.

In Formula 10, type is equal to 1 or 0 for inpatients and noninpatients, respectively. Center is equal to 1 or 0 for single-center or multi-center trials, respectively. The covariate model showed that the $E_{max, drug}$ value of inpatients was 18.4% higher than that of noninpatients, and the $E_{max, drug}$ value of single-center trials was 10.2% higher than that of multi-central trials. The $E_{max, drug}$ value of noninpatients in the multi-center trials was 16.4%.

The goodness-of-fit plots for the final model indicated that the population predictions and individual predictions fitted well along the identity line relative to the observed efficacy data, and the conditionally weighted residuals were well distributed along the zero line relative to population predictions and time point, which showed no obvious trends or model misspecification (supplementary Figure 3). The visual predictive check of the final model indicated the model describes the observed data well, and model predictions were generally within 95% prediction intervals (Figure 2). The median values from the bootstrap procedure

Table 2. Subgroup analysis of the pure efficacy of all antidepressants at week 8

All of drugs	Arms (sample size)	Efficacy at week 8, % (95% CI)	Corrected efficacy at week 8,% (95% CI)
Overall	355 (40455)	10.3 (1.10~28.1)	9.44 (0.67~22.2)
Sample size			
n<250	133 (7567)	14.4 (2.92~34.4)	10.5 (0.98~24.0)
n ≥ 250	222 (32888)	9.06 (0.92~18.8)	8.99 (0.70~18.8)
Trial design			
Placebo run-in	208 (20970)	10.4 (2.13~32.7)	9.75 (1.31~21.3)
Non placebo run-in	40 (5155)	10.4 (-0.01~29.8)	9.82 (-0.01~17.4)
Publication year			
Before 2000	112 (8729)	15.7 (4.42~34.1)	12.0 (2.23~23.3)
After 2000 (2000 included)	156 (22207)	9.52 (1.84~25.0)	9.43 (1.15~23.5)
Number of arms per trial			
≤2	55 (5975)	13.2 (2.45~38.5)	10.3 (2.45~21.8)
≥3	300 (34480)	9.68 (0.91~27.4)	9.33 (0.33~22.1)
Number of study sites			
1	43 (1883)	16.7 (6.13~34.6)	10.1 (0.74~24.2)
>1	289 (37242)	9.44 (0.86~22.1)	9.09 (0.63~20.4)
Funding source			
Industry-sponsored	257 (29986)	9.87 (1.06~25.4)	9.15 (0.90~20.0)
Non industry-sponsored	85 (9184)	11.8 (2.40~27.9)	10.3 (-0.29~25.0)
Type of subject			
Inpatients	21 (944)	21.7 (10.8~44.5)	8.89 (0.63~22.9)
Noninpatients	281 (31830)	10.3 (1.02~25.7)	9.73 (0.82~21.2)
Type of scale			
HAMD-17	166 (18908)	9.43 (1.19~27.5)	8.72 (0.85~21.0)
Non HAMD-17	189 (21547)	10.6 (1.17~28.1)	9.84 (0.51~23.7)
Dosing regimen			
Fixed dose	181 (22759)	9.23 (0.34~25.9)	8.89 (0.34~22.4)
Flexible dose	172 (17671)	11.5 (1.30~30.7)	10.2 (1.05~21.8)
Mean age, y			
<41	158 (17540)	9.99 (0.51~21.5)	9.41 (-0.48~19.9)
≥41	167 (20374)	10.63 (2.72~31.1)	9.61 (2.13~24.9)
Standardized baseline			
<0.44	162 (17018)	9.96 (0.59~25.6)	9.47 (0.59~21.1)
≥0.44	193 (23437)	10.4 (2.12~33.7)	9.44 (1.49~23.2)
Females, %			
<63%	94 (10983)	11.0 (0.11~27.2)	8.83 (-0.03~23.5)
≥63%	106 (13676)	9.38 (1.39~25.1)	8.92 (0.80~22.5)

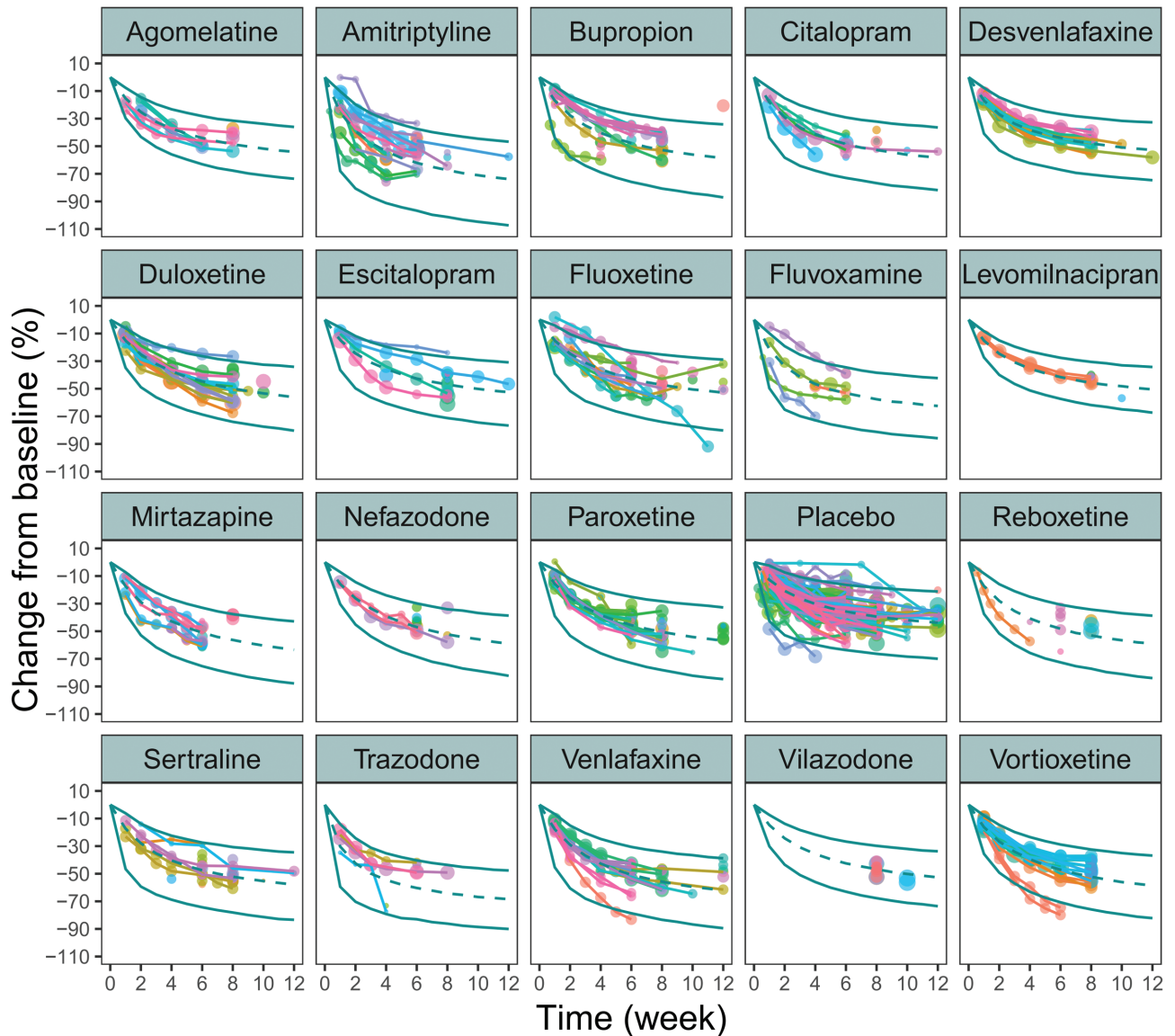Abbreviations: CI, confidence interval; HAMD, Hamilton Depression Rating Scale.

Figure 2. Visual predictive check of the final model for each treatment. The points represent the observed efficacy data and the symbol size is proportional to the sample size. The solid lines are the model predicted 95% confidence interval (CI) of each treatment. The dotted line is the median value of the model predicted efficacy.

were close to the parameter estimates from the NONMEM. In addition, more than 99% of the bootstrap runs were successful, indicating that the model was stable.

Typical Efficacies

Because the type of setting (inpatient or noninpatient) and the number of study sites (single or multi-center) had a significant impact on $E_{max,drug}$ value, it is necessary to correct E_{max} value for each drug arm to the same covariant level. The correction formula is as follows:

$$E\max, drug, i, corrected = E\max, drug, i - Type_i \times 18.4 - Center_i \times 10.2$$
(11)

In Formula 11, $E_{max,drug,i}$ is the $E_{max,drug}$ value of the ith drug group. $E_{max,drug,i,corrected}$ is the corrected value of $E_{max, drug, i.}$ $Type_i$ is the type of setting of the ith drug group (1 is for inpatients and 0 is for non-inpatient). $Center_i$ is the number of study sites of the ith

drug group (1 is for single center and 0 is for multi-center). This formula corrects the $E_{max,drug}$ values of different drug groups to the characteristics of multi-center clinical trials with noninpatients, thereby eliminating the influence of the above 2 covariates on the $E_{max,drug}$ values. After correction, the pure efficacy of drugs between different subgroups was basically comparable (Table 2).

For each kind of antidepressant (Table 3), the $E_{max,drug,corrected}$ of amitriptyline was 23.5% (95% CI = 19.1%–27.8%), which was the highest among all 19 antidepressants. The $E_{max,drug,corrected}$ of the remaining drugs were lower than 20%. The onset time of amitriptyline and trazodone is fast with $ET_{50,drug}$ values less than 3 weeks, while the $ET_{50,drug}$ values of the other drugs ranged from 3.9 weeks to 5.5 weeks. Based on the efficacy parameters described above, the distribution of drug efficacy at different time points can be simulated. Taking week 8 as an example, the effect of amitriptyline is 17.4% (95% CI = 14.1%–20.8%), which is obviously higher than other drugs. The other 18 antidepressants were distributed between 5.9% and 15.2% at week 8. Among them, we found the efficacy of venlafaxine was significantly

Table 3. Pharmacodynamic parameters and typical pure efficacy of each treatment

Group	Arms (sample size)	Emax (95% CI)	Emax, corrected (95% CI)	ET50 (95% CI)	Corrected efficacy at week 8 (95% CI)
Amitriptyline	26 (1582)	29.2 (24.9,33.5)	23.5 (19.1,27.8)	2.80 (2.15,3.45)	17.4 (14.1~20.8)
Trazodone	6 (570)	23.7 (14.8~32.6)	18.5 (9.9~27.2)	1.75 (0.51~2.98)	15.2 (8.0~22.8)
Venlafaxine	27 (2774)	18.8 (15.3~22.2)	17.7 (14.2~21.1)	4.04 (3.33~4.76)	11.7 (9.4~14.1)
Mirtazapine	15 (825)	21.1 (15.9~26.3)	17.6 (12.5~22.8)	3.93 (2.65~5.21)	11.8 (8.2~15.7)
Fluvoxamine	7 (518)	22.9 (14.3~31.4)	17.5 (9.0~26.1)	4.13 (2.6~5.67)	11.6 (5.8~17.6)
Duloxetine	22 (2690)	16.8 (13.2~20.4)	16.8 (13.2~20.4)	4.31 (3.49~5.13)	10.9 (8.5~13.4)
Levomilnacipran	8 (1603)	15.6 (9.6~21.6)	15.6 (9.6~21.6)	5.48 (3.85~7.11)	9.3 (5.6~13.2)
Vortioxetine	24 (3765)	15.3 (11.2~19.3)	15.3 (11.2~19.3)	5.41 (4.58~6.24)	9.1(6.6~11.6)
Sertraline	21 (1948)	15.9 (11.7~20.2)	15.2 (10.9~19.5)	4.90 (3.90~5.89)	9.4 (6.8~12.2)
Citalopram	17 (2061)	15.2 (10.9~19.6)	15.2 (10.9~19.6)	4.37 (3.34~5.40)	9.8 (6.9~12.8)
Paroxetine	44 (4782)	16.6 (13.7~19.5)	14.5 (11.6~17.3)	5.04 (4.38~5.71)	8.9 (7.1~10.7)
Agomelatine	10 (1481)	14.4 (8.9~19.9)	14.4 (8.9~19.9)	3.96 (2.21~5.70)	9.6 (5.8~13.8)
Escitalopram	20 (3032)	13.3 (9.3~17.4)	13.3 (9.3~17.4)	4.20 (3.27~5.13)	8.7 (6.0~11.4)
Fluoxetine	36 (3307)	13.9 (10.7~17.2)	12.1 (8.9~15.4)	4.78 (4.02~5.53)	7.6 (5.5~9.7)
Desvenlafaxine	13 (2645)	12.1 (7.6~16.7)	12.1 (7.6~16.7)	5.02 (3.81~6.23)	7.5 (4.6~10.5)
Nefazodone	13 (809)	15.3 (9.7~21.0)	11.6 (6.0~17.3)	5.17 (3.87~6.47)	7.1 (3.6~10.6)
Vilazodone	12 (2264)	11.1 (6.0~16.2)	11.1 (6.0~16.2)	5.27 (3.88~6.67)	6.7 (3.6~9.9)
Bupropion	23 (2467)	14.0 (9.4~18.7)	10.0 (6.1~13.8)	5.0 (3.73~6.27)	6.2 (3.8~8.7)
Reboxetine	11 (1332)	14.0 (4.5~23.4)	8.9 (2.6~15.3)	4.18 (2.93~5.44)	5.9 (1.7~10.1)
Placebo	230 (23 891)	56.6 (54.9~58.4)	—	3.53 (3.3~3.76)	39.3 (37.9~40.7)[a]

Abbreviations: CI, confidence interval; HAMD, Hamilton Depression Rating Scale.
[a]The efficacy of placebo group was not corrected; this value was the original value.

higher than bupropion, and their efficacy was at 11.7% (95% CI = 9.4%–14.1%) and 6.2% (95% CI = 3.8%–8.7%), respectively.

Time-Course Relationship

The time-course relationship was distinct in the 19 antidepressants (Figure 3). The efficacy of most drugs at week 4 was only about 50% of their $E_{max, drug}$ values, and this ratio increased to 70% at week 12. However, the onset time of trazodone and amitriptyline was faster, and their efficacy can exceed 80% of their $E_{max, drug}$ values by week 12.

Dose-Effect Relationship

In this study, the efficacy characteristics of each drug within the approved dose range were analyzed. The results showed that there was no significant dose-effect relationship among the 19 antidepressants within the approved dose range at week 8, and the efficacy of the high dose was similar to that of low dose (Figure 4).

Dropout Rate Due to Adverse Events

The dropout rate due to adverse events in the antidepressant groups was 4% higher than that of the placebo groups. Amitriptyline, trazodone, and fluvoxamine had higher dropout rates, which were 11% (95% CI = 7%–15%), 10% (95% CI = 2%–17%), and 9% (95% CI = 5%–13%) higher, respectively, than that of the placebo group. The dropout rate of agomelatine was low, which was comparable with the placebo group (Figure 5).

Discussion

There are many differences in the methodology of clinical trials of antidepressants (Wagner et al., 2017), such as having different control groups (positive drug or placebo control), different treatment durations, different number of study sites (single or multi-center), and so on. It has been reported that these heterogeneities will have significant impacts on drug efficacy. If the above heterogeneities were not corrected, it may affect the accurate judgment of drug efficacy. In this study, MBMA method was used to correct for the heterogeneities among studies. Our analysis showed that the publication year (before or after 2000), the number of study sites (single or multi-center), the sample size per study, and type of setting (inpatient or noninpatient) had a significant impact on $E_{max,drug}$ (Walsh et al., 2002). However, the publication year and the sample size per study were highly correlated with the number of study sites. That is, the earlier the publication year and smaller the sample size per study, the higher was the proportion of single-center trials (Undurraga and Baldessarini, 2012). Therefore, when the number of study sites (single or multi-center) was introduced into the covariate model, the significant influence of the publication year and the sample size per study disappeared, indicating that the publication year and the sample size per study were not essential impact factors for drug efficacy. The final covariate model only retained the number of study sites (single or multi-center) and the type of setting (inpatient or noninpatient). The model shows that the $E_{max,drug}$ value of single-center trials was 10.2% higher than that of multi-central trials, and the $E_{max,drug}$ value of inpatients was 18.4% higher than that of noninpatients. The above results are consistent with those reported in the literature (Undurraga and Baldessarini, 2012). Previous studies (Rutherford and Roose, 2013) have shown that with an increasing number of study sites, the screening for participants may be less stringent and the outcome assessment may be less consistent owing to the lack of experienced investigators, which results in a higher placebo response and much smaller differences in effectiveness between drug and placebo groups. The compliance of inpatients is generally better than that of outpatients; thus it is easier to observe a higher drug-placebo difference in the inpatients (Ansseaul et al., 1989).
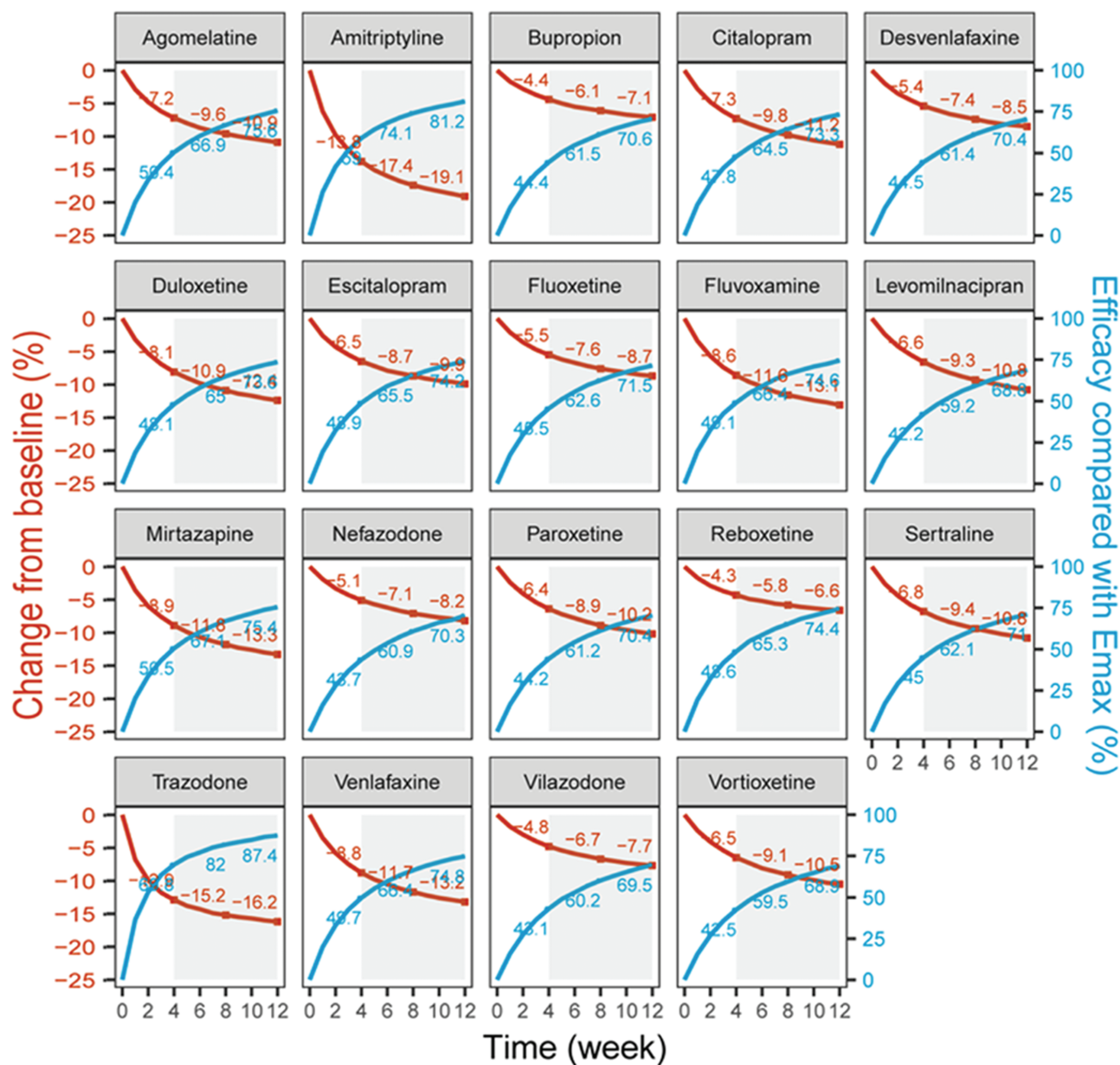
Figure 3. Typical predicted pure efficacy of each drug (red lines) and the corresponding efficacy ratio compared with the E_{max_drug} value (bule lines) at each time points.

A recently published meta-analysis (Munkholm et al., 2019) pointed out that the efficacy of antidepressants vs placebos was higher in trials with a placebo run-in study design compared with trials without such a design (P=.05); a lower efficacy was found in trials categorized as industry sponsored than in trials categorized as nonindustry sponsored (P=.005). However, this study found that whether there was a placebo run-in or not and the sources of funding had no significant effect on drug efficacy. The main reason is that the evaluation index in this study was defined as the change rate rather than change value in rating the scale score from baseline, which can further eliminate the effect of the baseline on drug efficacy. In addition, this study also deducted the impact of heterogeneities on the results, such as the type of setting and the number of study sites. Therefore, when there is a correlation between the impact factors, we should be careful to identify the impact factors to avoid the occurrence of false correlations, which may be caused by the imbalance of other factors.

Compared with the original analysis, this study not only minimized the heterogeneities among the different studies but also analyzed the time-course and dose-effect relationships of antidepressants, which can provide the necessary quantitative information for clinical guidelines. The efficacy of amitriptyline was ranked as the highest in the original study (Cipriani et al., 2018). However, in the clinical trials of amitriptyline, the proportion of single-center trials is high. When this factor was corrected for, amitriptyline remained the highest in pure efficacy among the 19 antidepressants, which is about 3 times as high as that of reboxetine at week 8. Amitriptyline is the representative drug for tricyclic antidepressants and is included in the World Health Organization List of Essential Medicines (Furukawa et al., 2016). Although the efficacy of amitriptyline is significantly higher than other antidepressants (Won et al., 2014), its dropout rate due to adverse reactions was 11% higher than that of the placebo group, which was also significantly
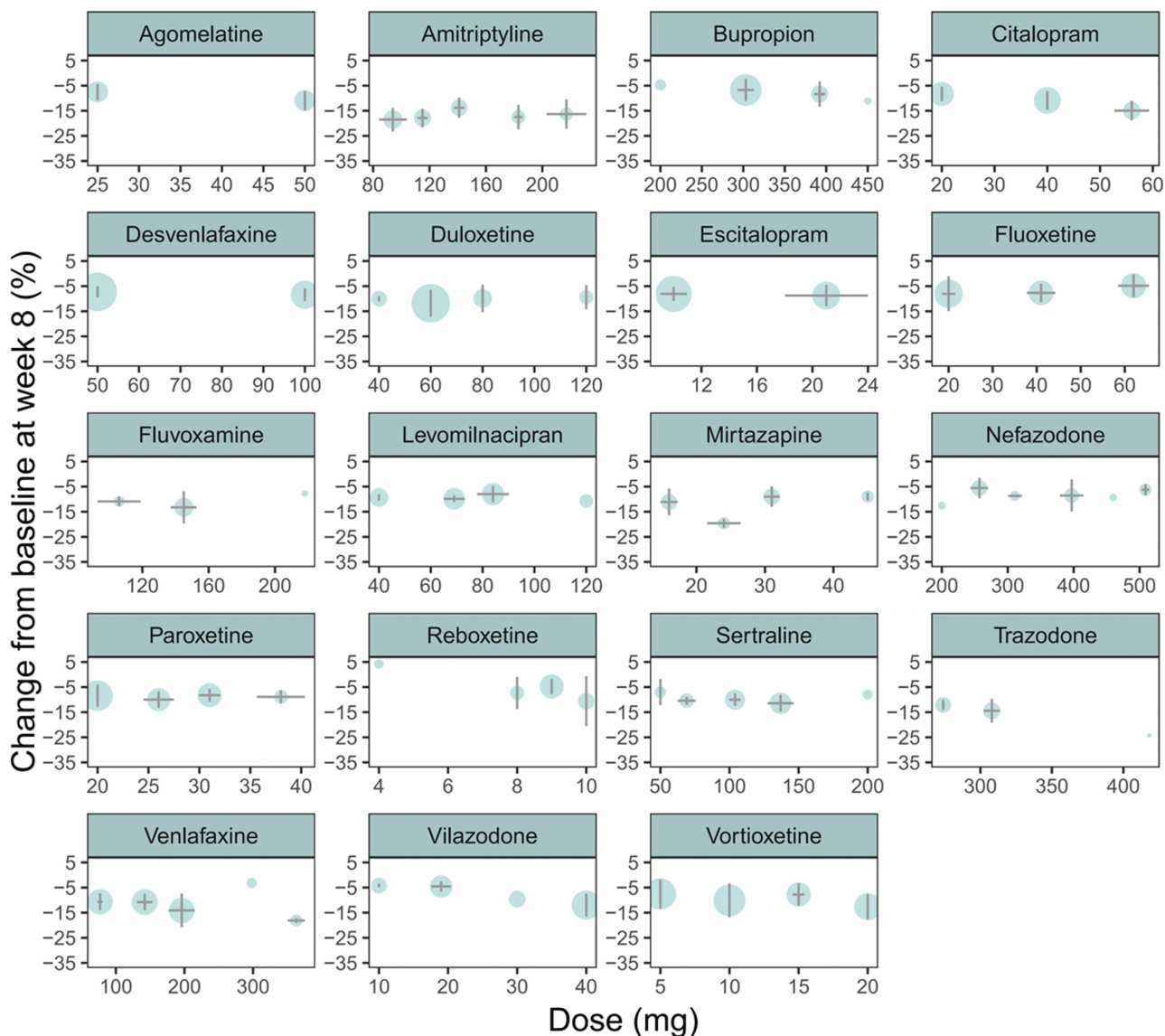
Figure 4. Dose-effect relationship of each drug. Horizontal error bars represent the SD of doses after the combination of similar doses. The vertical error bars represent the SD of the predicted pure efficacy of drugs at week 8 after the combination of same doses.

higher than other antidepressants. Among the remaining 18 antidepressants, excluding venlafaxine and bupropion, the 95% confidence intervals of pure efficacy of other 16 antidepressants at week 8 overlapped. However, the difference in pure efficacy between venlafaxine and bupropion was small, which was 5.5% at week 8. Therefore, we do not recommend ranking the efficacy of these 18 antidepressants owing to the small differences in efficacy among them.

Although most antidepressants had similar efficacy at week 8, there were some differences in the onset time. These results showed that the onset time of trazodone and amitriptyline were fast, and their ET_{50} values were less than 3 weeks. The onset time of levomilnacipran, vortioxetine, paroxetine, nefazodone, desvenlafaxine, vilazodone, and bupropion were relatively slow, and their ET_{50} values were longer than 5 weeks. It is worth mentioning that the rapid onset of trazodone and amitriptyline may be related to their sedative effect. Because amitriptyline and trazodone are H1 antagonists, they can quickly improve the sleep-related items of the HAMD (Becker, 2004). In addition,

treatment duration is an also important factor that affects the effectiveness of antidepressants (Bauer et al., 2013), and our model results showed that the efficacy of most antidepressants at week 12 is about 1.5 times that at week 4. However, the effect data at weeks 4–12 were combined for analysis in the previous meta-analysis, which ignored the impact of time on drug efficacy and thus may introduce deviation to the final conclusion (Cipriani et al., 2009).

The approved dose range for some antidepressants is large, such as the maximum approved dose for sertraline being 4 times stronger than that of its minimum approved dose (Cipriani et al., 2009). Recent meta-analyses (Hieronymus et al., 2016; Jakubovski et al., 2016; Furukawa et al., 2019a) have found a dose-response relationship for SSRIs; however, these meta-analyses were based on fixed-dose trials only, which do not reflect real-world conditions. In addition, these meta-analyses combined the analysis of different SSRIs and inevitably muddled the heterogeneity of different drugs. Therefore, the conclusions of the above studies are questionable (Jakubovski et al., 2016; Furukawa et al., 2019b). In
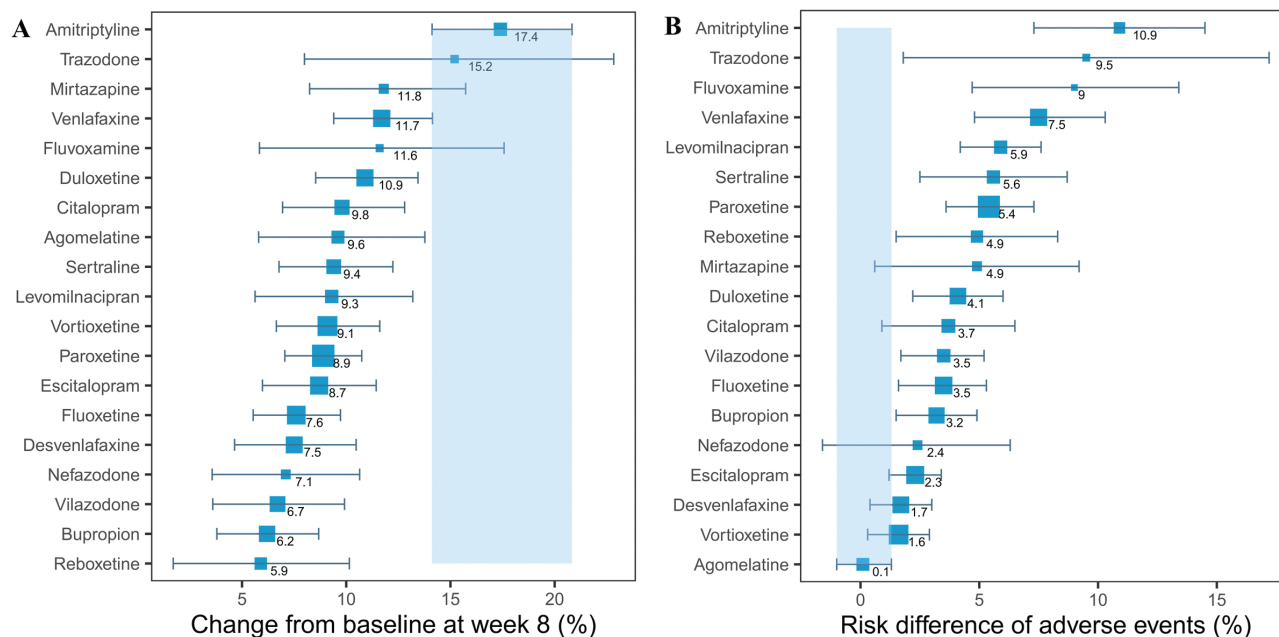
Figure 5. The predicted pure efficacy of each drug at week 8 (A) and the risk difference of adverse events between each drug and placebo (B). The error bar represents 95% confidence interval (CI) of the point estimate, the blue shadow area represents the distribution of efficacy (A) or adverse events (B) of the best drug, respectively.

this study, both fixed-dose and flexible-dose trials were included, and the dose-response relationship of each drug was analyzed. We found that 19 antidepressants had no significant dose-effect relationship within the approved dose range, and their pure efficacy at week 8 at high dose is comparable with that of low dose. This result is consistent with the conclusions of a previous study (Berney, 2005), which showed that the dose-response curve of SSRIs is flat. These results indicated that it is difficult to show a dose-effect relationship in clinical trials of antidepressants, and the reasons for that may be due to the low pure efficacy and large variation of antidepressants.

With the exception of amitriptyline, the dropout rate due to adverse events of the other drugs is no more than 10% higher than that of the placebo group. It should be noted that the dropout rate due to adverse events of agomelatine was almost the same as that of placebo. Agomelatine is a potent melatonergic MT1/MT2 receptor agonist with 5-HT2C receptor antagonist properties (Maniadakis et al., 2013). In addition to improving depression, agomelatine also has the advantages of being able to normalize sleep and circadian rhythm (Tuma et al., 2001), does not affect sexual function much, has good tolerance (Kennedy and Emsley, 2006), does not affect body weight, and does not have a strong withdrawal response. However, agomelatine has a risk of liver injury and requires regular monitoring of liver function (Norman and Olver, 2019).

Previous literature has suggested that there are differences between antidepressants in improving different depressive symptomologies (Baune et al., 2018). Therefore, it is necessary to carry out research to compare the efficacy of different antidepressants on different depressive symptoms to better guide patients to use drugs more effectively. However, this study was unable to assess the efficacy of different antidepressants on each depressive symptom due to the limited reports in the literature. In addition, this study also has the following limitations. This study only collected the study-level data for analysis, and some covariates were not tested as the range of summary level covariates was narrow. Despite the large amount of data

included in this study, many unpublished data were not included in the analysis and this may lead to some bias.

Conclusion

This study is a reanalysis of a previously published network meta-analysis that systematically compared the efficacy of 21 antidepressants. Compared with the original analysis method, our analysis adjusted the heterogeneities among the studies by modeling method and quantified the time-effect and dose-effect relationships and factors influencing the efficacy of 19 antidepressants. The results showed that the efficacy of amitriptyline was the highest and that of the other antidepressants were comparable or minimally different. Within the approved dose range, no significant dose-response relationship was observed. However, the time-course relationship is obvious for all antidepressants, with their pure efficacy at week 12 being about 1.5 times higher than that at week 4. In addition, we found that the efficacy of single-center clinical trials with inpatients was significantly better than that of multi-center clinical trials with noninpatients, which will need to be differentiated in future meta-analyses.

Supplementary Materials

Supplementary data are available at *International Journal of Neuropsychopharmacology (IJNPPY)* online.

Statement of Interest

The authors declare no conflict of interests.

References

Amick HR, Gartlehner G, Gaynes BN, Forneris C, Asher GN, Morgan LC, Coker-Schwimmer E, Boland E, Lux LJ, Gaylord S, Bann C, Pierl CB, Lohr KN (2015) Comparative benefits and harms of second generation antidepressants and cognitive behavioral therapies in initial treatment of major depressive disorder: systematic review and meta-analysis. Bmj 351:h6019.

An G, Ohls RK, Christensen RD, Widness JA, Mock DM, Veng-Pedersen P (2017) Population pharmacokinetics of darbepoetin in infants following single intravenous and subcutaneous dosing. J Pharm Sci 106:1644–1649.

Ansseaul M, Frenckelll RV, Mertens C, et al. (1989) Controlled comparison of two doses of milnacipran (F 2207) and amitriptyline in major depressive inpatients. Psychopharmacology 98:163–168.

Bauer M, Pfennig A, Severus E, Whybrow PC, Angst J, Möller HJ; World Federation of Societies of Biological Psychiatry. Task Force on Unipolar Depressive Disorders (2013) World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of unipolar depressive disorders, part 1: update 2013 on the acute and continuation treatment of unipolar depressive disorders. World J Biol Psychiatry 14:334–385.

Baune BT, Brignone M, Larsen KG (2018) A Network meta-analysis comparing effects of various antidepressant classes on the digit symbol substitution test (DSST) as a measure of cognitive dysfunction in patients with major depressive disorder. Int J Neuropsychopharmacol 21:97–107.

Becker PM (2004) Trazodone as a hypnotic in major depression. Sleep Med 5:7–8.

Berney P (2005) Dose-response relationship of recent antidepressants in the short-term treatment of depression. Dialogues Clin Neurosci 7:249–262.

Boucher M, Bennetts M (2016) The many flavors of model-based meta-analysis: part I-introduction and landmark data. CPT Pharmacometrics Syst Pharmacol 5:54–64.

Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, Watanabe N, Nakagawa A, Omori IM, McGuire H, Tansella M, Barbui C (2009) Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet 373:746–758.

Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, Leucht S, Ruhe HG, Turner EH, Higgins JPT, Egger M, Takeshima N, Hayasaka Y, Imai H, Shinohara K, Tajika A, Ioannidis JPA, Geddes JR (2018) Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet 391:1357–1366.

Clayton AH, Zajecka J, Ferguson JM, Filipiak-Reisner JK, Brown MT, Schwartz GE (2003) Lack of sexual dysfunction with the selective noradrenaline reuptake inhibitor reboxetine during treatment for major depressive disorder. Int Clin Psychopharmacol 18:151–156.

Cleare A, Pariante CM, Young AH, Anderson IM, Christmas D, Cowen PJ, Dickens C, Ferrier IN, Geddes J, Gilbody S, Haddad PM, Katona C, Lewis G, Malizia A, McAllister-Williams RH, Ramchandani P, Scott J, Taylor D, Uher R; Members of the Consensus Meeting (2015) Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2008 British Association for Psychopharmacology guidelines. J Psychopharmacol 29:459–525.

Deardorff WJ, Grossberg GT (2014) A review of the clinical efficacy, safety and tolerability of the antidepressants vilazodone, levomilnacipran and vortioxetine. Expert Opin Pharmacother 15:2525–2542.

Furukawa TA, Salanti G, Atkinson LZ, Leucht S, Ruhe HG, Turner EH, Chaimani A, Ogawa Y, Takeshima N, Hayasaka Y, Imai H, Shinohara K, Suganuma A, Watanabe N, Stockton S, Geddes JR, Cipriani A (2016) Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. BMJ Open 6:e010919.

Furukawa TA, Cipriani A, Cowen PJ, et al. (2019a) Optimal dose of selective serotonin reuptake inhibitors, venlafaxine, and mirtazapine in major depression: a systematic review and dose-response meta-analysis. Lancet Psychiatry 6:601–609.

Furukawa TA, Cowen PJ, Salanti G, Leucht S, Egger M, Cipriani A (2019b) Optimal dosing of antidepressant drugs - authors' reply. Lancet Psychiatry 6:806–807.

Hieronymus F, Nilsson S, Eriksson E (2016) A mega-analysis of fixed-dose trials reveals dose-dependency and a rapid onset of action for the antidepressant effect of three selective serotonin reuptake inhibitors. Transl Psychiatry 6:e834.

Ioannidis JP (2008) Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? Philos Ethics Humanit Med 3:14.

Jakobsen JC, Katakam KK, Schou A, et al. (2017) Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and trial sequential analysis. BMC Psychiatry 17:162.

Jakubovski E, Bloch MH (2016) Addressing difficulties in the study of dose-response relationships of SSRIs in depression: response to Hieronymus and Eriksson. Am J Psychiatry 173:836–838.

Jakubovski E, Varigonda AL, Freemantle N, et al. (2016) Systematic review and meta-analysis: dose-response relationship of selective serotonin reuptake inhibitors in major depressive disorder. Am J Psychiatry 173:174–183.

Kennedy SH, Emsley R (2006) Placebo-controlled trial of agomelatine in the treatment of major depressive disorder. Eur Neuropsychopharmacol 16:93–100.

Khoo AL, Zhou HJ, Teng M, Lin L, Zhao YJ, Soh LB, Mok YM, Lim BP, Gwee KP (2015) Network meta-analysis and cost-effectiveness analysis of new generation antidepressants. CNS Drugs 29:695–712.

Lalonde RL, Kowalski KG, Hutmacher MM, Ewy W, Nichols DJ, Milligan PA, Corrigan BW, Lockwood PA, Marshall SA, Benincosa LJ, Tensfeldt TG, Parivar K, Amantea M, Glue P, Koide H, Miller R (2007) Model-based drug development. Clin Pharmacol Ther 82:21–32.

Maniadakis N, Kourlaba G, Mougiakos T, Chatzimanolis I, Jonsson L (2013) Economic evaluation of agomelatine relative to other antidepressants for treatment of major depressive disorders in Greece. BMC Health Serv Res 13:173.

Munkholm K, Paludan-Müller AS, Boesen K (2019) Considering the methodological limitations in the evidence base of antidepressants for depression: a reanalysis of a network meta-analysis. BMJ Open 9:e024886.

Nguyen TH, Mouksassi MS, Holford N, et al. (2017) Model evaluation of continuous data pharmacometric models: metrics and graphics. CPT Pharmacometrics Syst Pharmacol 6:87–109.

Norman TR, Olver JS (2019) Agomelatine for depression: expanding the horizons? Expert Opin Pharmacother 20:647–656.

Papadimitropoulou K, Vossen C, Karabis A, Donatti C, Kubitz N (2017) Comparative efficacy and tolerability of pharmacological and somatic interventions in adult patients with treatment-resistant depression: a systematic review and network meta-analysis. Curr Med Res Opin 33:701–711.

Qaseem A, Barry MJ, Kansagara D; Clinical Guidelines Committee of the American College of Physicians (2016) Nonpharmacologic versus pharmacologic treatment of adult patients with major depressive disorder: a clinical practice guideline from the American College of Physicians. Ann Intern Med 164:350–359.

Rutherford BR, Roose SP (2013) A model of placebo response in antidepressant clinical trials. Am J Psychiatry 170:723–733.

Rutherford BR, Sneed JR, Roose SP (2009) Does study design influence outcome? The effects of placebo control and treatment duration in antidepressant trials. Psychother Psychosom 78:172–181.

Santarsieri D, Schwartz TL (2015) Antidepressant efficacy and side-effect burden: a quick guide for clinicians. Drugs Context 4:212290.

Tuma J, Strubbe JH, Mocaër E, Koolhaas JM (2001) S20098 affects the free-running rhythms of body temperature and activity and decreases light-induced phase delays of circadian rhythms of the rat. Chronobiol Int 18:781–799.

Undurraga J, Baldessarini RJ (2012) Randomized, placebo-controlled trials of antidepressants for acute major depression: thirty-year meta-analytic review. Neuropsychopharmacology 37:851–864.

Wagner S, Engel A, Engelmann J, Herzog D, Dreimüller N, Müller MB, Tadić A, Lieb K. (2017) Early improvement as a resilience signal predicting later remission to antidepressant treatment in patients with major depressive disorder: systematic review and meta-analysis. J Psychiatr Res 94:96–106.

Walsh BT, Seidman SN, Sysko R, Gould M (2002) Placebo response in studies of major depression: variable, substantial, and growing. Jama 287:1840–1847.

Wang SM, Han C, Lee SJ, Jun TY, Patkar AA, Masand PS, Pae CU (2016) Second generation antipsychotics in the treatment of major depressive disorder: an update. Chonnam Med J 52:159–172.

Won E, Park SC, Han KM, Sung SH, Lee HY, Paik JW, Jeon HJ, Lee MS, Shim SH, Ko YH, Lee KJ, Han C, Ham BJ, Choi J, Hwang TY, Oh KS, Hahn SW, Park YC, Lee MS; Clinical Research Center for Depression (2014) Evidence-based, pharmacological treatment guideline for depression in Korea, revised edition. J Korean Med Sci 29:468–484.