

Differential efficacy of escitalopram and nortriptyline on dimensional measures of depression

Rudolf Uher, Wolfgang Maier, Joanna Hauser, Andrej Marušič, Christine Schmael, Ole Mors, Neven Henigsberg, Daniel Souery, Anna Placentino, Marcella Rietschel, Astrid Zobel, Monika Dmitrzak-Weglarz, Ana Petrovic, Lisbeth Jorgensen, Petra Kalember, Caterina Giovannini, Mara Barreto, Amanda Elkin, Sabine Landau, Anne Farmer, Katherine J. Aitchison and Peter McGuffin

Background

Tricyclic antidepressants and serotonin reuptake inhibitors are considered to be equally effective, but differences may have been obscured by internally inconsistent measurement scales and inefficient statistical analyses.

Aims

To test the hypothesis that escitalopram and nortriptyline differ in their effects on observed mood, cognitive and neurovegetative symptoms of depression.

Method

In a multicentre part-randomised open-label design (the Genome Based Therapeutic Drugs for Depression (GENDEP) study) 811 adults with moderate to severe unipolar depression were allocated to flexible dosage escitalopram or nortriptyline for 12 weeks. The weekly Montgomery–Åsberg Depression Rating Scale, Hamilton Rating Scale for Depression, and Beck Depression Inventory were scored both conventionally and in a more novel way according to dimensions of observed mood, cognitive symptoms and neurovegetative symptoms.

Results

Mixed-effect linear regression showed no difference between escitalopram and nortriptyline on the three original scales, but symptom dimensions revealed drug-specific advantages. Observed mood and cognitive symptoms improved more with escitalopram than with nortriptyline. Neurovegetative symptoms improved more with nortriptyline than with escitalopram.

Conclusions

The three symptom dimensions provided sensitive descriptors of differential antidepressant response and enabled identification of drug-specific effects.

Declaration of interest

S.L. owns shares in GlaxoSmithKline. N.H. participated in clinical trials sponsored by pharmaceutical companies including GlaxoSmithKline and Lundbeck. A.F., P.M. and K.J.A. have received consultancy fees and honoraria for participating in expert panels from pharmaceutical companies including Lundbeck and GlaxoSmithKline. Funding detailed in Acknowledgements.

Less than 50% of people with depression respond to the first prescribed antidepressant, but the majority eventually respond to a different treatment.^{1,2} The rate and magnitude of response appear to be similar for tricyclic antidepressants and selective serotonin reuptake inhibitors (SSRIs).^{3–5} Psychiatrists are unable to predict which drug will work for whom and the choice of first and subsequent treatments has to progress by trial and error. The present study addresses two major methodological challenges that may have precluded identification of drug-specific effects in previous studies: symptomatic heterogeneity and statistical power.

Although depression is conceived as a single condition, its defining symptoms do not necessarily co-occur and individual symptoms may differ in their distribution across individuals and their response to treatments.⁶ This heterogeneity of depressive symptoms complicates exploration of drug effects. For example, the early improvement of sleep with tricyclic antidepressants may be unrelated to sustained response, but early improvement in anxiety precedes and predicts overall improvement.⁷ Such cross-sectional and longitudinal dissociations between symptom dimensions decrease the correlations between items of scales that combine mood, anxiety and sleep items in a single score, i.e. impair their internal consistency, to a degree where a summed test score is uninformative.^{8,9} We have sought to remediate this problem and, using categorical item factor analysis, we identified three dimensions of depressive symptoms with good psychometric properties: observed mood, cognitive and neurovegetative symptoms.¹⁰ The present study tests the hypothesis that escitalopram and nortriptyline differ in their effects on these dimensions.

A second challenge concerns the effectiveness of statistical analysis. Most previous trials were powered to compare active medication with placebo, but differences between active antidepressants are likely to be smaller.¹¹ To maximise the power for a specified sample size, it is essential that all information on outcome is used in the analysis. Many previous investigations used dichotomised outcomes (e.g. responder/non-responder). However, response to antidepressants is a matter of degree of change rather than a yes/no qualitative transformation, and dichotomising a continuous outcome is associated with a substantial loss of power.^{12,13} Furthermore, temporal characteristics of antidepressant response are lost in end-point analysis and the commonly used last observation carried forward procedure for missing data produces biased results.^{14–16} In the present report, we apply mixed-effect modelling that permits the use of data measured at multiple time points, and provides unbiased estimates in the presence of missing data.^{14,16,17} This approach also separates inter-individual variation in antidepressant response from measurement error and unmeasured centre differences. This partitioning allows estimation of the proportion of variance attributable to unmeasured individual-specific characteristics, including genes.

Method

Study design

Genome Based Therapeutic Drugs for Depression (GENDEP) is a partially randomised multicentre clinical and pharmacogenetic

study comparing two active antidepressants with contrasting modes of action. The study was undertaken in nine European clinical centres. GENDEP is registered at EudraCT2004-001723-38 (<http://eudract.emea.europa.eu>) and ISRCTN03693000 (www.controlled-trials.com).

Pragmatic design features were adopted to make GENDEP inclusive and acceptable to a large proportion of people with depression.¹⁸ These included non-random allocation of participants who would otherwise not be eligible, no use of placebo, flexible dosage, no post-allocation masking and open communication with general practitioners.

Interventions

Two antidepressants were selected that represent the two most common mechanisms of action among commonly used antidepressants and have a good efficacy record. Escitalopram is a highly selective inhibitor of the serotonin transporter with no effect on noradrenaline reuptake.¹⁹ Nortriptyline is a tricyclic antidepressant with a hundred times higher affinity for the noradrenaline transporter than for the serotonin transporter.²⁰ Nortriptyline was used in preference to the even more selective reboxetine as it has better established efficacy and was considered to be clinically at equipoise with escitalopram.

Study medication was started immediately after the first assessment in antidepressant-free participants or participants on low doses of other antidepressants. Two week wash-out was required for people on fluoxetine or monoamine oxidase inhibitors. Escitalopram was initiated at 10 mg daily and increased to a target dose of 15 mg daily within the first 2 weeks unless adverse effects limited dose increase, and could be further increased to 20 mg daily (and up to 30 mg if there was clinical agreement that a higher dose was needed). Nortriptyline was initiated at 50 mg daily and titrated to a target dose of 100 mg daily within the first 2 weeks unless adverse effects limited dose increase, and could be further increased to 150 mg daily (and up to 200 mg if there was clinical agreement that a higher dose was needed). Use of plasma levels to guide dose titration has been suggested for nortriptyline, but it is of uncertain benefit²¹ and could introduce a systematic difference between the two antidepressants. Therefore, dose titration of both antidepressants was informed by assessments of depressive symptoms and adverse effects rather than plasma levels. Adherence was recorded weekly as self-reported pill count and plasma levels of antidepressants were measured at week 8. Other psychotropic medication was prohibited with the exception of occasional use of hypnotics.

Allocation

Participants for whom the two antidepressants were clinically considered to be at equipoise were randomly allocated to receive escitalopram or nortriptyline using a random number generator, stratified by centre and performed independently of the assessing clinician. If there was a history of adverse effects, non-response or contraindications to one of the study medications, participants were allocated to the other drug non-randomly. Participants who could not tolerate the initially allocated medication or who did not experience sufficient improvement with adequate dosage within 8 weeks were offered the other antidepressant. Participants who swapped medication were then followed up for 12 weeks.

Outcome measures

The clinician-rated Montgomery-Åsberg Depression Rating Scale (MADRS),²² the 17-item Hamilton Rating Scale for Depression (HRSD-17)²³ and the self-report Beck Depression Inventory

(BDI)²⁴ were administered at baseline and then weekly for 12 weeks. The week 0, 8 and 12 assessments were face-to-face interviews with a psychiatrist and a research assistant, both trained in the administration of the instruments. The remaining assessments were conducted by telephone or face-to-face interviews with a trained psychologist or psychiatrist. Psychometric properties and interrater reliability have been reported.¹⁰ Using factor analysis of ordered categorical variables with robust weighted least squares estimator and item response modelling, the items of the three scales were integrated into three dimensional scores of observed mood, cognitive symptoms and neurovegetative symptoms.¹⁰ The dimensional scores for the present analyses were estimated based on a graded-response model using the previously reported item parameters¹⁰ applied in the MULTLOG 7 software for Windows.²⁵ The observed mood dimension comprised the symptoms of depressed mood, activity, anxiety and psychomotor disturbance rated by the clinician. The cognitive symptoms dimension consisted of guilt, pessimism, suicidal thoughts and most items of the self-report BDI. The neurovegetative factor included disturbed sleep, loss of appetite, weight loss and lack of libido. Full mapping of individual items to dimensions is available in a previous article.¹⁰ To facilitate interpretation, dimensional symptom scores have been converted to T-scores with a mean of 50 and standard deviation of 10, based on the baseline assessment. This makes a change of 10 on a dimensional score comparable with a change of 10 points on BDI, 7 points on MADRS or 5 points on HRSD-17.

Sample size and recruitment of participants

The sample size of over 800 gives GENDEP 90% power to detect drug differences corresponding to an effect size (Cohen's *d*) as small as 0.06 at $\alpha=0.05$.

Participants were recruited by generalist and specialist referrals and advertisement. Inclusion criteria were: White European ethnicity (to facilitate genetic association analyses), age 18 or older, onset of current depressive episode at age 65 or younger, and a diagnosis of major depressive episode of at least moderate severity defined by the ICD-10²⁶ or DSM-IV²⁷ and established using the Schedules for Clinical Assessment in Neuropsychiatry interview (SCAN version 2.1).²⁸ The exclusion criteria were: family history of bipolar affective disorder or schizophrenia in a first-degree relative, a personal history of hypomanic or manic episode, schizophrenia, mood incongruent psychotic symptoms, primary substance misuse, primary organic disease and pregnancy. Participants were also excluded if they had contraindications or a history of lack of efficacy or adverse reaction to both study medications. The study protocol was approved by the research ethics boards of all participating centres. After explanation of study procedures, all participants provided written consent.

Statistical analysis

Baseline characteristics were compared using chi-squared tests, Kruskal-Wallis tests or ANOVA for categorical, ordered and continuous variables respectively. Predictors of time to drop out or switch from initially allocated treatment were assessed by Cox proportional hazard regression with drug, allocation (random *v.* non-random), gender, age, baseline severity, taking antidepressants and benzodiazepines at baseline and number of previous episodes as explanatory variables.

To assess fair dosage of the two antidepressants, we followed the recommendation of a consensus group on antidepressant comparisons,¹¹ and used Cox proportional hazard regression to assess the impact of drug and allocation on time to reach a

mid-range dose, which is half-way between the lowest effective and highest recommended dose, i.e. 15 mg for escitalopram and 100 mg for nortriptyline.

Outcomes were analysed using mixed models with individual random intercepts and slopes, and fitted with full maximum likelihood.¹⁷ Participants who swapped medication were included under both medications, with the last measurement on the first antidepressant serving as a baseline for the effect of the second antidepressant, a fixed covariate capturing systematic differences between first and second run of medication, and individual-level clustering being controlled by the random effect of the individual. Centre was included as a higher-level random effect. Model selection was performed by means of likelihood ratio tests. The best fitting model included fixed linear and quadratic effects of time, and fixed linear effects of baseline severity, drug, allocation and age.

The mixed-effect models provide unbiased estimates, assuming the data is missing at random and the variables associated with missing values are included in the model.^{14,29} To assess the missing data mechanism, we explored the relationship between missingness and observed variables at baseline and at the last observed time point.

The combined analysis of randomised and non-randomised participants may be subject to confounding by baseline group differences on observed or unobserved variables. Therefore, to evaluate the sensitivity of our analysis to selection effects, the mixed-model analyses were repeated on the reduced sample of observations from randomised individuals while they were on their first course of medication.

All analyses were conducted in Stata 10 for Windows.³⁰

Results

Screening and reasons for non-inclusion

The flow of participants through the study is summarised in Figs 1 and 2. The reasons for exclusions at the screening stage were: not fulfilling diagnostic criteria for moderate or severe depressive episode (24%); bipolar disorder or psychotic symptoms (18%); unable to discontinue current psychotropic medication (16%); ethnicity (10%); primary alcohol or substance misuse (7%); family history of bipolar disorder or schizophrenia (7%); unable to attend the study centre (7%); contraindications (6%); age (3%); and pregnancy (2%).

Sample and baseline characteristics

From July 2004 to December 2007, 468 participants were randomised and 343 participants were allocated non-randomly (Fig. 1). More participants were non-randomly allocated to escitalopram than to nortriptyline. Sample characteristics at baseline are presented in Table 1 (full details are presented in online Table DS1). The non-randomly allocated participants differed from the randomised sample: fewer were married ($\chi^2(3)=11.72$, $P=0.008$) or employed ($\chi^2(5)=13.86$, $P=0.017$), they had later age at onset ($F(1, 809)=10.56$, $P=0.001$), fewer depressive episodes (Kruskal–Wallis $\chi^2(1)=45.70$, $P<0.001$) and less severe symptoms (MADRS $F(1, 809)=7.22$, $P=0.007$). Within the participants who could not be randomly allocated to treatment, those receiving nortriptyline had more previous episodes (Kruskal–Wallis $\chi^2(1)=5.04$, $P=0.025$) (Table 1) and were more likely to have a history of taking SSRI-type antidepressants ($\chi^2(1)=7.36$,

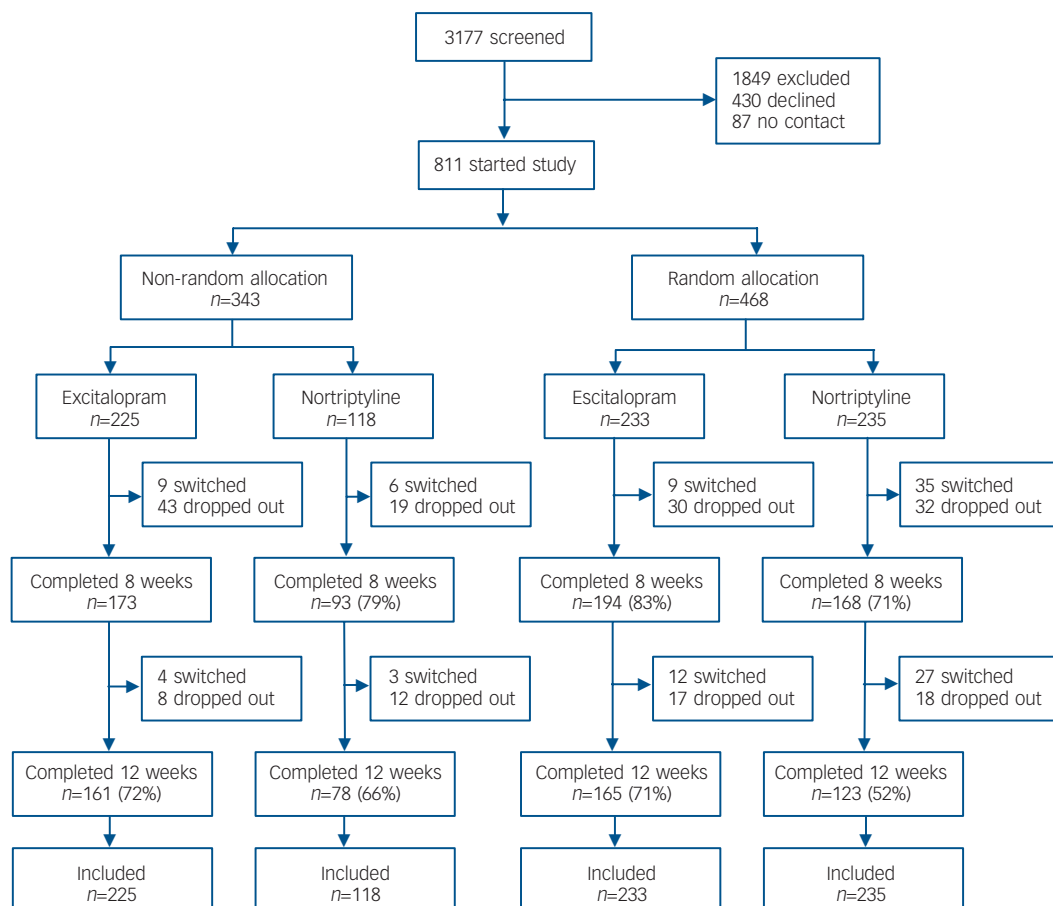


Fig. 1 Flow of participants through the study on initially allocated antidepressant.

Table 1 Baseline sample characteristics

	Randomised		Non-randomised	
	Escitalopram (n=233)	Nortriptyline (n=235)	Escitalopram (n=225)	Nortriptyline (n=118)
Female, n (%)	146 (63)	161 (69)	137 (61)	70 (59)
Age, years: mean (s.d.)	41.9 (11.6)	42.5 (11.6)	43.5 (11.7)	41.9 (12.6)
Education, years: mean (s.d.)	12.2 (3.2)	12.1 (3.1)	12 (3)	12.1 (3.1)
Occupation, n (%)				
Full-time work	88 (38)	84 (36)	85 (38)	32 (27)
Part-time work	28 (12)	29 (12)	23 (10)	15 (13)
Student	15 (6)	9 (4)	15 (7)	12 (10)
Home-maker	8 (3)	13 (6)	6 (3)	7 (6)
Retired	28 (12)	41 (17)	17 (8)	12 (10)
Unemployed	66 (28)	59 (25)	79 (35)	40 (34)
Age at onset, years: mean (s.d.)	32.2 (10.6)	31.4 (9.9)	34.4 (8.7)	33.3 (10.1)
Episodes, n (%)				
1	47 (20)	59 (25)	115 (51)	43 (36)
2	149 (64)	138 (59)	99 (44)	66 (56)
3+	37 (16)	38 (16)	11 (5)	9 (8)
Current episode duration, weeks: mean (s.d.)	19.5 (14.3)	17.8 (13.1)	18.9 (8.5)	20.7 (13)
Currently taking drug, n (%)				
Antidepressants	48 (21)	62 (26)	38 (17)	22 (19)
Benzodiazepines	90 (39)	80 (34)	84 (37)	36 (31)
Psychotropic medication	113 (49)	117 (50)	103 (46)	47 (40)
Baseline severity, mean (s.d.)				
Montgomery-Åsberg Depression Rating Scale	29.1 (6.6)	29.4 (6.7)	27.6 (6.7)	28.6 (6.9)
Hamilton Rating Scale for Depression-17	22.2 (5.1)	22.3 (5.4)	20.8 (5.4)	21.6 (5.2)
Beck Depression Inventory	28.6 (9.4)	28.6 (9.9)	26.9 (9.8)	28.0 (9.6)
Observed mood	49.8 (9.4)	50.2 (10.6)	48.3 (10.6)	48.6 (10.2)
Cognitive symptoms	49.9 (9.8)	50.1 (10.2)	47.9 (10.7)	49.2 (10.5)
Neurovegetative symptoms	50.8 (9.9)	50.1 (10.1)	48.6 (9.9)	49.8 (8.6)

SSRI, selective serotonin reuptake inhibitor.

$P=0.007$) than those non-randomly allocated to escitalopram (online Table DS1).

Retention of participants

Of the 811 participants, 628 (77%) completed 8 weeks and 527 (65%) completed 12 weeks on the originally allocated antidepressant (Fig. 1). Over the 12 weeks, 105 (13%) participants switched to the other antidepressant and an additional 4 switched after completing 12 weeks on the originally allocated drug. Reasons for switching were poor tolerance (39%), lack of effect (45%) or both (16%). Over the 12 weeks, 179 participants dropped out because of adverse reactions (31%), lack of effect (34%), improvement (8%), death (1%, see adverse events) and other reasons (25%). Of the 109 participants who switched antidepressant, 80 (73%) completed 8 weeks and 68 (62%) completed 12 weeks on the second antidepressant (Fig. 2).

The rate of drop out and switching was highest among participants randomly allocated to nortriptyline (hazard ratio (HR)=1.87, 95% CI 1.36–2.56, $P=0.001$ compared with random escitalopram; HR=1.47, 95% CI 1.02–2.13, $P=0.041$, compared with non-random nortriptyline; Fig. 3). There were no significant differences in drop-out and switching rate among the other three groups. Attrition was predicted by more severe baseline symptoms with a hazard ratio of 1.22 (95% CI 1.08–1.38, $P=0.002$) for one standard deviation increase in MADRS.

Missing data

The weekly data on depression severity were 92.9% complete and proportion of missing values did not differ between groups. Taking benzodiazepines at the time of recruitment was related

to the proportion of missing values; 4% data were missing in participants who were taking benzodiazepines at baseline compared with 9% in participants who were not taking benzodiazepines ($\beta=-0.045$, 95% CI -0.064 to -0.026 , $P<0.001$). Younger age was associated with more missing values ($\beta=-0.010$, 95%

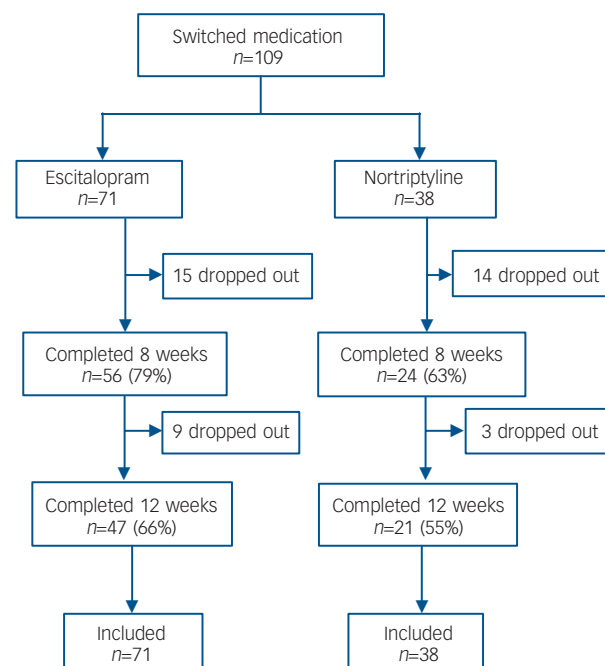


Fig. 2 Flow of participants through the study for those who swapped to the second antidepressant.

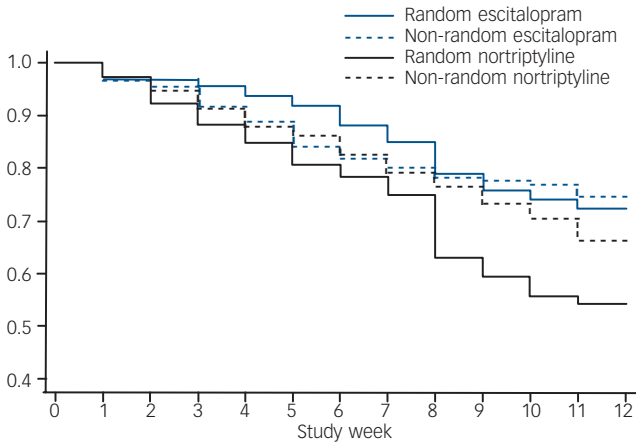


Fig. 3 Retention of participants on initially allocated medication (survival graph).

CI -0.018 to -0.001 , $P=0.030$). Other clinical and demographic variables were not related to missing data. Missing values at a specific time point (t) were not predicted by severity of depression on the preceding visit ($t=-1$), for example for MADRS ($\beta=-0.003$, 95% CI -0.012 to 0.005 , $P>0.1$).

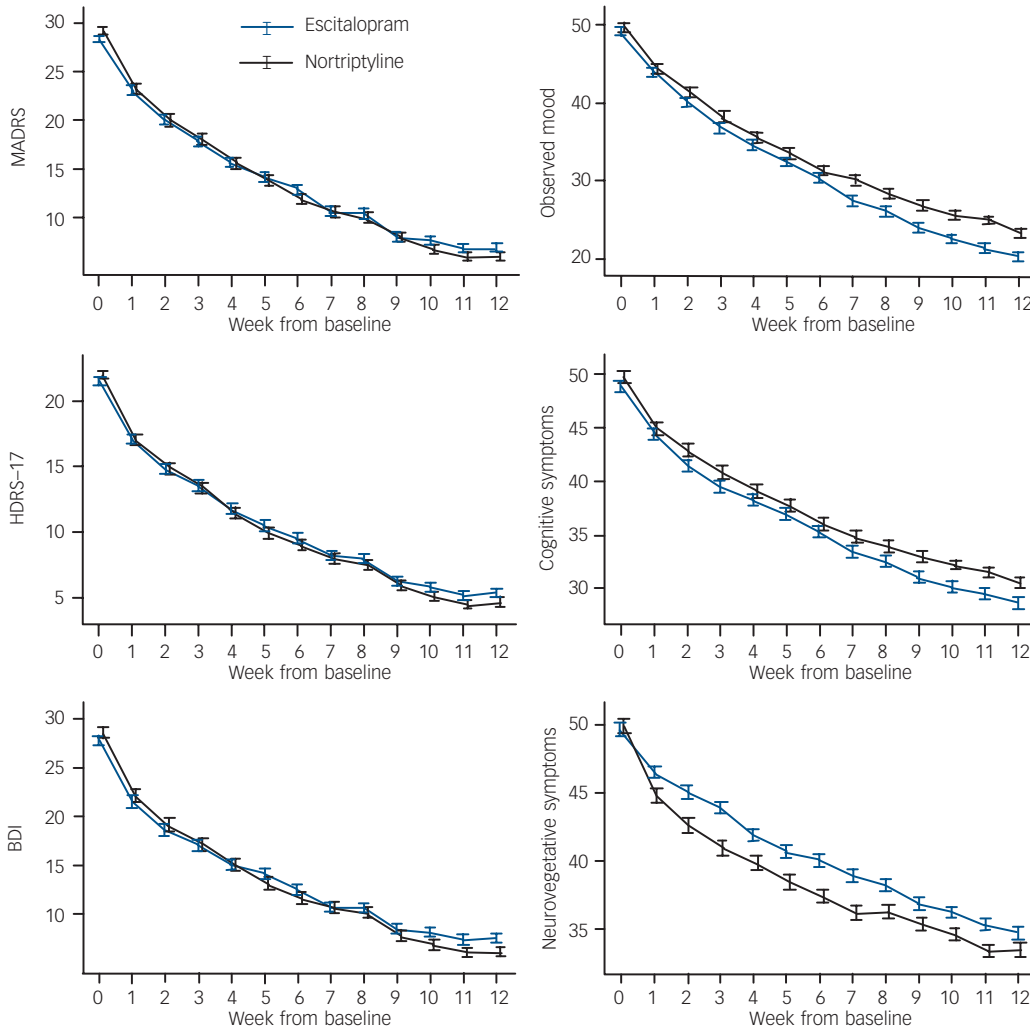


Fig. 4 Mean symptom scores by study week, drug and outcome measure. Symptom dimensions are represented as T-scores with a mean of 50 and standard deviation of 10 at baseline. Error bars represent 1 standard error of the mean. MADRS, Montgomery-Åsberg Depression Rating Scale; HDRS-17, Hamilton Rating Scale for Depression (17 items); BDI, Beck Depression Inventory.

Antidepressant dosage and adherence

For both antidepressants, the median time to reach mid-range dose was 3 weeks, and there was no significant effect of drug (HR=1.11, 95% CI 0.95–1.30, $P=0.198$), indicating similar rate of dose titration for both antidepressants. The mean dose by study group and week is presented in the online Table DS2. The self-reported adherence was high (98.4%) and did not differ between treatment groups ($P>0.1$). The average plasma levels at the eighth treatment week were nortriptyline 100.4 mcg/l (s.d.=57.9) and citalopram 30.7 mcg/l (s.d.=21.2), with no significant difference between randomly and non-randomly allocated participants ($P>0.1$).

Changes in depression symptoms

The weekly measurements of depressive symptoms on the three original scales and the three symptom dimensions are presented in Fig. 4. The mixed models included linear and quadratic functions of time, fixed effects of drug, randomisation status, baseline severity, age, gender, number of depressive episodes, history of taking antidepressants and benzodiazepines at baseline (the latter was included as it predicted missingness) and showed that drug did not affect the outcome measured by the HDRS-17, MADRS or BDI (all $P>0.1$, Table 2). However, there were

Table 2 Between-drug differences in the final mixed-effect models^a

	Whole sample analysis			Randomised sample analysis		
	β	95% CI	<i>P</i>	β	95% CI	<i>P</i>
Original scales						
Montgomery-Åsberg Depression Rating Scale	-0.006	-0.054 to 0.042	0.803	0.000	-0.104 to 0.104	0.998
Hamilton Rating Scale for Depression-17	0.008	-0.040 to 0.056	0.742	-0.004	-0.110 to 0.102	0.936
Beck Depression Inventory	-0.014	-0.061 to 0.033	0.551	0.016	-0.095 to 0.127	0.779
Symptom dimensions						
Observed mood	0.098	0.049 to 0.147	<0.001	0.075	-0.028 to 0.178	0.151
Cognitive	0.086	0.039 to 0.134	<0.001	0.079	-0.035 to 0.193	0.175
Neurovegetative	-0.156	-0.209 to -0.103	<0.001	-0.175	-0.290 to -0.061	0.003

a. β is the standardised regression coefficient of drug effect on symptom scores and can be interpreted as effect size in units of standard deviation. Positive values of β indicate more severe symptoms (less improvement) during treatment with nortriptyline compared with escitalopram. Negative values of β indicate greater improvement with nortriptyline compared with escitalopram. Whole sample analysis includes both medications in participants who switched. Randomised sample analysis only includes data from the first antidepressant course, when participants were treated by the randomly allocated medication.

significant effects of drug on outcome on each of the three symptom dimensions. The observed mood and cognitive symptoms improved more in escitalopram-treated participants. The neurovegetative symptoms improved more in those receiving nortriptyline (Table 2).

To control for selection bias, we performed a sensitivity analysis restricted to the first course of antidepressant treatment in the randomised participants. The results were very similar with all effect size estimates within one standard error of the whole sample estimates (Table 2). The degree of statistical certainty was reduced owing to the smaller sample size.

Younger age was associated with improvement on all measures (e.g. for MADRS: $\beta=0.08$, 95% CI 0.04–0.11 per 10 years of age, $P<0.001$). History of taking antidepressants predicted less improvement on all measures (e.g. for MADRS: $\beta=0.13$, 95% CI 0.04–0.23, $P=0.005$).

The fixed part of the models explained 35% of variability in antidepressant response on the observed mood dimension. Of the remaining variance, 8% was attributable to the unmeasured characteristics of centre, 69% was at the level of individual and 28% remained as level-three residuals, corresponding to measurement error and unmeasured time-varying factors.

Information on response and remission using last observation carried forward analysis is available in the online data supplement.

Adverse events and reactions

Two participants died during the study period. A woman randomised to nortriptyline died by suicide in the ninth week. A man randomly allocated to escitalopram died of a road traffic accident in the fifth week. Severe adverse events included two hospital admissions owing to suicide risk (ninth week on random escitalopram, third week on random nortriptyline), a manic episode in the third week of nortriptyline and an unintentional overdose of nortriptyline with full recovery. Commonly reported adverse reactions to escitalopram included nausea and vomiting (15%) and sexual dysfunction (30%). Common adverse effects of nortriptyline included dry mouth (80%), orthostatic dizziness (32%), drowsiness (27%) and constipation (24%).

Discussion

Differential effects of antidepressants

The present results demonstrate the utility of dimensional symptom measures derived by psychometric analysis to identify relative advantages of individual antidepressants. Escitalopram was more effective than nortriptyline in relieving mood and cognitive symptoms of depression. Nortriptyline was more

effective than escitalopram in improving neurovegetative symptoms such as disturbed sleep and poor appetite. None of these differences would have been revealed by summed scores on conventional depression rating scales that combine all three types of symptoms.

The observed mood dimension reflects the symptoms of depressed mood, anxiety, psychomotor retardation and activity. It has been noted that changes in core mood symptoms are more likely to reflect sustained antidepressant effect,⁷ differentiate active antidepressants from placebo,³¹ show dose-response relationship³² and moderation by polymorphism in the serotonin transporter gene.³³ The observed mood dimension contains information from most items that constitute the previously suggested core sub-scales of the HRSD,^{31,34} but has the advantages of using information from a larger number of items and not making indefensible assumptions about additivity and equal contribution of items.^{10,35} Therefore, the observed mood score is suitable for testing hypotheses related to pharmacological modulation of affect and biomarkers of the monoaminergic systems. The strong effect of escitalopram on observed mood indicates the utility of this antidepressant in people where core affective symptoms dominate the clinical picture.

The cognitive symptoms dimension comprises items reflecting dissatisfaction with oneself, pessimism, guilt and suicidal thoughts. It shows a modest advantage of escitalopram over nortriptyline. As suicidal ideation appears to lie on a continuum with cognitive symptoms,¹⁰ the cognitive dimension may be evaluated as a monitoring tool for treatment-emergent suicidality.³⁶

The most robust finding of the present study was that neurovegetative symptoms improved significantly more with nortriptyline than with escitalopram. The neurovegetative symptom dimension includes disturbed sleep, decreased appetite, weight loss and lack of sexual interest. These symptoms are characteristic of melancholic depression and may indicate the need for antidepressants with a broader spectrum of pharmacological effects.³⁷ It has been reported that the HRSD-17 with three sleep items may give an advantage to tricyclic antidepressants that improve sleep through their anticholinergic action over SSRIs that may disturb sleep, cause gastrointestinal discomfort and sexual dysfunction.³⁸ Sleep improvement may be independent of antidepressant action on mood⁷ and moderated by genes regulating the circadian rhythm.³⁹ The present findings add to the weight of evidence indicating that sleep and appetite should be measured separately from the core mood symptoms.

Our results suggest that failure to find differential efficacy of tricyclic antidepressants and SSRIs in previous studies³ may have been because such differences were obscured by the internal

inconsistency of scales such as the HRSD-17.⁸ As the item response theory scoring is independent of the number of administered items,³⁵ it could be used to derive equivalent scores for samples where either HRSD or MADRS is available.¹⁰ This raises the possibility of re-examining existing data-sets to attempt to replicate the present findings and extend them to placebo-controlled trials.

The size of the drug differences is comparatively small. However, it may be of clinical utility since it is approximately 25–50% of the size of the differences between antidepressants and placebo in contemporary trials.^{40,41} Increased efficacy of the item response theory-scored dimensions may also have substantial implications for the sample size and power of future comparisons between active drugs or between drugs and placebo.⁴² Moreover, small overall differences can point to large differences in subgroups of patients. A relatively small improvement in accuracy of symptom measurement can magnify the power to detect interactions between drug and individual characteristics, and facilitate identification of predictors of differential drug response.⁴³ Dimensional symptom scores will allow testing of specific pharmacogenetic hypotheses concerning mood,³³ neurovegetative³⁹ or cognitive symptoms.³⁶

The mixed-effect modelling estimated the sources of residual variability in symptom change over time. Although a number of predictors have been included in the models, these have jointly explained only 35% of the variance in the individual trajectories of depressive symptoms. Most of the residual variance is attributable to unmeasured individual characteristics that are stable over time. This large proportion of variance presents a challenge for future research, which should include exploration of genetic factors and early environmental influences.

Methodological considerations and limitations

Differential effects in clinical comparisons may be a result of genuine differences between treatments or may be false positives owing to chance, bias or confounding. Chance alone is unlikely to account for the present findings as the differential effects were identified with a high level of statistical certainty. Additional analyses excluded other potential sources of bias and confounding such as baseline differences between groups allocated to different drugs and inequality of dose titration.¹¹

The attrition rate was higher among participants randomly allocated to nortriptyline. This is consistent with previous reports.^{44,45} Interestingly, the differential attrition was a result of switching rather than drop out and did not generalise to participants who were non-randomly allocated to nortriptyline. This suggests that a high discontinuation rate on nortriptyline is not inevitable, and that clinical assessment based on medication history improves the fit between the individual and the antidepressant.

Differential drop out can lead to bias, especially with the last observation carried forward procedure.^{14,16,45} We applied maximum likelihood estimation with observed predictors of missingness included in the model. This method is robust to differential rates of missing data.^{14,15,17}

The GENDEP study aimed to include a sample representative of the treatment-seeking population of individuals with depression. Therefore, non-random allocation was allowed where the two antidepressants were not at equipoise and the participants and their general practitioners knew which medication they were receiving. These features increased the acceptability of the study to participants and to general practitioners and thus made the study more inclusive and externally valid. However, they have implications for the internal validity. The inclusion of non-randomly allocated

participants introduced systematic differences at baseline. However, the findings were qualified by a sensitivity analysis that demonstrated that observed differential effects of drugs on symptom dimensions were not a result of selection bias. The lack of masking introduces a potential for biased reporting of symptoms. It is, however, unlikely that a reporting bias would operate in opposite directions for different categories of symptoms.

In conclusion, dimensional measures distinguishing between observed mood, cognitive and neurovegetative symptoms of depression allowed the identification of relative advantages of escitalopram and nortriptyline. The differential drug effects were not a result of baseline sample characteristics, unfair dosage or differential attrition. These dimensional symptom measures provide a powerful tool to facilitate drug comparisons and find predictors of differential drug response.

Rudolf Uher, PhD, Institute of Psychiatry, King's College London, UK; **Wolfgang Maier**, MD, Department of Psychiatry, University of Bonn, Germany; **Joanna Hauser**, MD, Laboratory of Psychiatric Genetics, Poznan University of Medical Sciences, Poland; **Andrej Marušič**, PhD, Institute of Public Health, Ljubljana, Slovenia; **Christine Schmael**, MD, Central Institute of Mental Health, Division of Genetic Epidemiology in Psychiatry, Mannheim, Germany; **Ole Mors**, PhD, Centre for Psychiatric Research, Aarhus University Hospital, Riskov, Denmark; **Neven Henigsberg**, MD, Croatian Institute for Brain Research, Medical School, University of Zagreb, Croatia; **Daniel Souery**, PhD, Université Libre de Bruxelles, Erasme Academic Hospital, Department of Psychiatry, Brussels, Belgium; **Anna Placentino**, PsyD, Biological Psychiatry Unit and Dual Diagnosis ward IRCCS, Centro San Giovanni di Dio, FBF, Brescia, Italy; **Marcella Rietschel**, MD, Central Institute of Mental Health, Division of Genetic Epidemiology in Psychiatry, Mannheim, Germany; **Astrid Zobel**, MD, Department of Psychiatry, University of Bonn, Germany; **Monika Dmitrak-Weglarz**, PhD, Laboratory of Psychiatric Genetics, Poznan University of Medical Sciences, Poland; **Ana Petrovic**, MD, Institute of Public Health, Ljubljana, Slovenia; **Lisbeth Jorgensen**, MSc Psych, Centre for Psychiatric Research and Mood Disorders Research Unit, Aarhus University Hospital, Riskov, Denmark; **Petra Kalember**, MD, Croatian Institute for Brain Research, Medical School, University of Zagreb, Croatia; **Caterina Giovannini**, PsyD, Biological Psychiatry Unit and Dual Diagnosis ward IRCCS, Centro San Giovanni di Dio, FBF, Brescia, Italy; **Mara Barreto**, MD, Université Libre de Bruxelles, Erasme Academic Hospital, Department of Psychiatry, Brussels, Belgium; **Amanda Elkin**, MRCPsych, **Sabine Landau**, PhD, **Anne Farmer**, MD, FRCPsych, **Katherine J. Aitchison**, PhD, **Peter McGuffin**, PhD, FRCP, FRCPsych, Institute of Psychiatry, King's College London, UK.

Correspondence: Rudolf Uher, P080 SGDP, Institute of Psychiatry, 16 De Crespigny Park, SE5 8AF, London, UK. Email: r.uher@iop.kcl.ac.uk

First received 22 Jul 2008, final revision 7 Oct 2008, accepted 16 Oct 2008

Acknowledgements

The GENDEP study was funded by the European Commission Framework 6 grant, EC Contract Ref.: LSHB-CT-2003-503428. Lundbeck provided both nortriptyline and escitalopram free of charge for the GENDEP study. GlaxoSmithKline contributed by funding an add-on project in the London centre. The sponsors had no role in the design and conduct of the study, in data collection, analysis, interpretation or writing the report. We would like to thank the following collaborators for their contribution: Helen Dean, Bhanu Gupta, Joanna Gray, Cerisse Gunasinghe, Desmond Campbell, Richard J Williamson, Julien Mendlewicz, Thomas Schulze, Jana Strohmaier, Susanne Höfels, Anna Schuhmacher, Ute Pfeiffer, Sandra Weber, Erik Roj Larsen, Anne Schinkel Stamp, Dejan Kozel, Mojca Zvezdana Dernovek, Alenka Tancic, Jerneja Sveticic, Zrnka Kovacic, Pawe Kapelski, Maria Skibińska, Piotr M Czernski, Aleksandra Rajewska, Aleksandra Szczepankiewicz and Elzbieta Cegielska. We would like to specially acknowledge the contribution of Jorge Perez, who was the principal investigator at Brescia, Italy, and who passed away in October 2007. We also wish to acknowledge the important contribution made by Andrej Marušič, the principal investigator at Ljubljana, Slovenia, who passed away in June 2008.

References

- 1 Ruhe HG, Huyser J, Swinkels JA, Schene AH. Switching antidepressants after a first selective serotonin reuptake inhibitor in major depressive disorder: a systematic review. *J Clin Psychiatry* 2006; **67**: 1836–55.
- 2 Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry* 2006; **163**: 1905–17.
- 3 Anderson IM. Selective serotonin reuptake inhibitors versus tricyclic antidepressants: a meta-analysis of efficacy and tolerability. *J Affect Disord* 2000; **58**: 19–36.

- 4 Cipriani A, Geddes JR, Furukawa TA, Barbui C. Metareview on short-term effectiveness and safety of antidepressants for depression: an evidence-based approach to inform clinical practice. *Can J Psychiatry* 2007; **52**: 553–62.
- 5 Geddes JR, Freemantle N, Mason J, Eccles MP, Boynton J. SSRIs versus other antidepressants for depressive disorder. *Cochrane Database Syst Rev* 2000; **2**: CD001851.
- 6 Fava M, Uebelacker LA, Alpert JE, Nierenberg AA, Pava JA, Rosenbaum JF. Major depressive subtypes and treatment response. *Biol Psychiatry* 1997; **42**: 568–76.
- 7 Katz MM, Koslow SH, Frazer A. Onset of antidepressant activity: reexamining the structure of depression and multiple actions of drugs. *Depress Anxiety* 1996; **4**: 257–67.
- 8 Bagby RM, Ryder AG, Schuller DR, Marshall MB. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? *Am J Psychiatry* 2004; **161**: 2163–77.
- 9 Santor DA, Coyne JC. Examining symptom expression as a function of symptom severity: item performance on the Hamilton Rating Scale for Depression. *Psychol Assess* 2001; **13**: 127–39.
- 10 Uher R, Farmer A, Maier W, Rietschel M, Hauser J, Marusic A, et al. Measuring depression: comparison and integration of three scales in the GENDEP study. *Psychol Med* 2008; **38**: 289–300.
- 11 Lieberman JA, Greenhouse J, Hamer RM, Krishnan KR, Nemeroff CB, Sheehan DV, et al. Comparing the effects of antidepressants: consensus guidelines for evaluating quantitative reviews of antidepressant efficacy. *Neuropsychopharmacology* 2005; **30**: 445–60.
- 12 Ragland DR. Dichotomizing continuous outcome variables: dependence of the magnitude of association and statistical power on the cutpoint. *Epidemiology* 1992; **3**: 434–40.
- 13 Streiner DL. Breaking up is hard to do: the heartbreak of dichotomizing continuous data. *Can J Psychiatry* 2002; **47**: 262–66.
- 14 Mallinckrodt CH, Clark WS, David SR. Accounting for dropout bias using mixed-effects models. *J Biopharm Stat* 2001; **11**: 9–21.
- 15 Leon AC, Mallinckrodt CH, Chuang-Stein C, Archibald DG, Archer GE, Chartier K. Attrition in randomized controlled clinical trials: methodological issues in psychopharmacology. *Biol Psychiatry* 2006; **59**: 1001–5.
- 16 Lane P. Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. *Pharm Stat* 2008; **7**: 93–106.
- 17 Gueorguieva R, Krystal JH. Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Arch Gen Psychiatry* 2004; **61**: 310–7.
- 18 March JS, Silva SG, Compton S, Shapiro M, Califf R, Krishnan R. The case for practical clinical trials in psychiatry. *Am J Psychiatry* 2005; **162**: 836–46.
- 19 Sanchez C, Bergqvist PB, Brennum LT, Gupta S, Hogg S, Larsen A, et al. Escitalopram, the S-(+)-enantiomer of citalopram, is a selective serotonin reuptake inhibitor with potent effects in animal models predictive of antidepressant and anxiolytic activities. *Psychopharmacology (Berl)* 2003; **167**: 353–62.
- 20 Sanchez C, Hyttel J. Comparison of the effects of antidepressants and their metabolites on reuptake of biogenic amines and on receptor binding. *Cell Mol Neurobiol* 1999; **19**: 467–89.
- 21 Taylor D, Duncan D. Plasma levels of tricyclics and related antidepressants: are they necessary or useful? *Psychiat Bull* 1995; **19**: 548–50.
- 22 Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; **134**: 382–9.
- 23 Hamilton M. Development of a rating scale for primary depressive illness. *Br J Clin Psychol* 1967; **6**: 278–96.
- 24 Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatr* 1961; **4**: 561–71.
- 25 Thissen D, Chen WH, Bock D. *MULTLOG 7*. Scientific Software International, 2003.
- 26 World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. WHO, 1992.
- 27 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorder, (4th edn) (DSM-IV)*. APA, 1994.
- 28 Wing JK, Sartorius N, Ustin TB. *Diagnosis and Clinical Management in Psychiatry. A Reference Manual for SCAN*. World Health Organization, 1998.
- 29 Little T and Rubin D. *Analysis with Missing Data*. John Wiley & Sons, 1987.
- 30 StataCorp. *Stata statistical software: release 10*. Stata Corp LP, 2007.
- 31 Faries D, Herrera J, Rayamajhi J, DeBrotta D, Demitrac M, Potter WZ. The responsiveness of the Hamilton Depression Rating Scale. *J Psychiatric Res* 2000; **34**: 3–10.
- 32 Danish University Antidepressant Group (DUAG). Clomipramine dose-effect study in patients with depression: clinical end points and pharmacokinetics. *Clin Pharmacol Ther* 1999; **66**: 152–65.
- 33 Serretti A, Mandelli L, Lorenzi C, Pirovano A, Olgiati P, Colombo C, et al. Serotonin transporter gene influences the time course of improvement of 'core' depressive and somatic anxiety symptoms during treatment with SSRIs for recurrent mood disorders. *Psychiatry Res* 2007; **149**: 185–93.
- 34 Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bolwig TG. Quantitative rating of depressive states. *Acta Psychiatr Scand* 1975; **51**: 161–70.
- 35 Embretson SE, Reise SP. *Item Response Theory for Psychologists*. Lawrence Erlbaum Associates, 2000.
- 36 Perlis RH, Purcell S, Fava M, Fagerness J, Rush AJ, Trivedi MH, et al. Association between treatment-emergent suicidal ideation with citalopram and polymorphisms near cyclic adenosine monophosphate response element binding protein in the STAR*D study. *Arch Gen Psychiatry* 2007; **64**: 689–97.
- 37 Davidson J, Turnbull CD. Diagnostic significance of vegetative symptoms in depression. *Br J Psychiatry* 1986; **148**: 442–6.
- 38 Moller HJ. Methodological aspects in the assessment of severity of depression by the Hamilton Depression Scale. *Eur Arch Psychiatry Clin Neurosci* 2001; **251** (suppl 2): I113–20.
- 39 Serretti A, Cusin C, Benedetti F, Mandelli L, Pirovano A, Zanardi R, et al. Insomnia improvement during antidepressant treatment and CLOCK gene polymorphism. *Am J Med Genet B Neuropsychiatr Genet* 2005; **137**: 36–9.
- 40 Khan A, Kolts RL, Thase ME, Krishnan KR, Brown W. Research design features and patient characteristics associated with the outcome of antidepressant clinical trials. *Am J Psychiatry* 2004; **161**: 2045–9.
- 41 Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression: variable, substantial, and growing. *JAMA* 2002; **287**: 1840–7.
- 42 Leon AC, Marzuk PM, Portera L. More reliable outcome measures can reduce sample size requirements. *Arch Gen Psychiatry* 1995; **52**: 867–71.
- 43 Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol* 2004; **57**: 229–36.
- 44 Hotopf M, Hardy R, Lewis G. Discontinuation rates of SSRIs and tricyclic antidepressants: a meta-analysis and investigation of heterogeneity. *Br J Psychiatry* 1997; **170**: 120–7.
- 45 Joyce PR, Mulder RT, Luty SE, Sullivan PF, McKenzie JM, Abbott RM, et al. Patterns and predictors of remission, response and recovery in major depression treated with fluoxetine or nortriptyline. *Aust N Z J Psychiatry* 2002; **36**: 384–91.

