

Severity of depression and response to antidepressants: GENPOD randomised controlled trial

Nicola J. Wiles, Jean Mulligan, Tim J. Peters, Philip J. Cowen, Victoria Mason, David Nutt, Deborah Sharp, Debbie Tallon, Laura Thomas, Michael C. O'Donovan and Glyn Lewis

Background

Antidepressant prescribing is widespread. Nonetheless, response to antidepressants is variable. If it was possible to predict response to medication and thus tailor treatment accordingly, this would not only improve patient outcomes but may also have economic benefits.

Aims

To test the hypothesis that individuals with more severe depression would benefit more from noradrenaline reuptake inhibitors (NARIs) than selective serotonin reuptake inhibitors (SSRIs) compared with individuals with less severe depression.

Method

Individuals recruited from UK primary care who met ICD-10 criteria for a depressive episode and scored 15 or more on the Beck Depression Inventory (BDI) were randomised to either an SSRI (citalopram 20 mg daily) or a NARI (reboxetine 4 mg twice daily). Randomisation was by means of a remote automated telephone system. The main outcome was depressive symptoms measured by the BDI total score 6 weeks after randomisation. (Trial registration: ISRCTN31345163.)

Results

In total, 601 participants were randomised (citalopram: $n=298$, reboxetine: $n=303$). Ninety-one per cent were followed up at 6 weeks (citalopram: $n=274$, reboxetine: $n=272$). There was little evidence to support an interaction between treatment and severity of depression (interaction term: 0.02, 95% CI -0.59 to 0.62 , $P=0.96$). Adjustment for potential confounders (age, gender, employment status, history of depression, number of life events and social support) did not affect the findings (interaction term: 0.06, 95% CI -0.54 to 0.66 , $P=0.85$).

Conclusions

Treatment with NARIs does not confer any advantage over SSRI treatment for outcome in those with more severe depressive illness presenting in primary care.

Declaration of interest

P.J.C. has been a paid member of advisory boards of Eli Lilly, Servier, Wyeth and Xytis and has been a paid lecturer for Eli Lilly, Servier and GlaxoSmithKline. He has provided expert advice for solicitors representing GlaxoSmithKline. D.N. has acted as consultant and speaker for both Lundbeck and Pfizer. D.S. has acted as an advisor for Servier and has been a paid lecturer for Eli Lilly.

Antidepressants are often the first-line treatment for the management of depression in primary care. Antidepressant prescribing is widespread. In 2007–2008, 34 million prescriptions for antidepressants were written in England.¹ Similar prescribing of antidepressants is seen in other countries (for example Australia² and the USA³). Nonetheless, response to antidepressants is variable. The large US Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study found that only 30% of individuals will experience remission of their depressive symptoms after 14 weeks of treatment with citalopram.⁴ Subsequent response to treatment among non-responders is poor. Therefore, it would be extremely useful if it was possible to predict response and thus tailor treatment accordingly, the goal of 'stratified medicine'.⁵ Not only would this lead to improved patient outcomes but it would also enable clinicians to prescribe more effectively, which could have economic benefits.

Selective serotonin reuptake inhibitors (SSRIs) predominate as the treatment of choice. Most of the current evidence suggests that there is little difference in terms of efficacy between SSRIs and selective noradrenaline reuptake inhibitors (NARIs), although a multiple treatment meta-analysis has suggested that reboxetine could be less effective.^{6,7} However, the latter study by Cipriani *et al*⁷ assumed that those participants who were missing outcome data had not responded to treatment, but, as reboxetine was less well tolerated than SSRIs, this approach to handling missing data has the potential to introduce bias such that the outcome for those on SSRIs may appear more favourable.

The only NARI that has been licensed for use in the UK is reboxetine. A recent meta-analysis⁸ suggested that reboxetine was less effective than either placebo or SSRIs for the treatment of depression. Yet, an earlier publication by Papakostas *et al*⁹ that included data from four unpublished randomised controlled trials (RCTs) reached a different conclusion and reported no significant difference in response rates between reboxetine and SSRIs. This discrepancy has been acknowledged but not explained.⁸ Of note, Eyding *et al*⁸ found considerable heterogeneity in the meta-analysis of the eight RCTs comparing response rates for reboxetine *v.* placebo. This could not be fully explored as Eyding *et al*⁸ only had access to aggregate data but there was evidence that reboxetine was more effective than placebo among in-patients with current depression. Such individuals would be expected to have more severe illness than the community patients customarily recruited into modern regulatory trials of new antidepressants. Others have also reported a greater drug–placebo difference in individuals with more severe depression.¹⁰

Two other studies have suggested that NARIs are more effective in the more severe depressions.^{11,12} Historically, the presence of 'endogenous' or 'biological' symptoms were used as clinical indications for the prescription of antidepressants and such symptoms are viewed as markers of severity.¹³ The Genome Based Therapeutic Drugs for Depression (GENDEP) study has reported that even though overall outcome (as measured by scores on the Beck Depression Inventory (BDI)) was similar between

escitalopram and nortriptyline, the latter, a tricyclic antidepressant that predominantly blocks the reuptake of noradrenaline, led to more improvement in 'neurovegetative' symptoms of sleep, appetite and sexual interest and less improvement in mood and cognitive symptoms.¹⁴ However, GENDEP did not examine whether severity of depression predicted (primary) treatment response and there is little other literature in this area.¹⁵

The GENetic and clinical Predictors Of treatment response in Depression (GENPOD) study was designed to test two primary hypotheses: (a) those with the 'l/l' genotype of 5-HTTLPR would respond better to an SSRI than a NARI; and (b) those with more severe depression would show better response to a NARI than an SSRI. Importantly, given that the primary analyses were based on tests for interaction (that is, subgroup analyses), sample size was calculated accordingly. The primary outcome, principal comparisons and method of analysis were specified in advance as advocated in CONSORT guidelines.¹⁶ The results for one of the primary hypotheses, the genetic predictor of treatment response, have already been reported.¹⁷ The current paper reports the findings for the clinical predictor (the other primary hypothesis for the GENPOD study). We hypothesised that individuals with more severe depression would benefit more from NARIs than SSRIs compared with those with less severe depression.

Method

The GENPOD trial

The trial protocol has been published elsewhere.¹⁸ In brief, GENPOD was a multicentre RCT conducted in Bristol, Birmingham and Newcastle, UK. Individuals aged 18–74 years were referred to GENPOD by their general practitioner (GP) following agreement that an antidepressant should be prescribed. Participants were randomly allocated to receive either an SSRI (citalopram 20 mg daily) or a NARI (reboxetine 4 mg twice daily). Only those individuals who met ICD-10¹⁹ criteria for a depressive episode (F32) from the computerised Clinical Interview Schedule – Revised (CIS-R)^{20,21} and had a Beck Depression Inventory (BDI)^{22,23} score of ≥ 15 at the baseline assessment were eligible to participate.

Individuals who had taken antidepressant medication in the 2 weeks prior to the baseline assessment were excluded, as were those who could not complete self-administered questionnaires. General practitioners also excluded those with medical contraindications, psychosis, bipolar disorder, major substance or alcohol misuse and others whose participation was deemed inappropriate.

Ethical approval was obtained from the South West Research Ethics Committee (MREC 02/6/076) and research governance approval was granted by Bristol, Manchester and Newcastle Primary Care NHS Trusts. (Trial registration: ISRCTN31345163 and EudraCT number: 2004-001434-16.)

Randomisation procedure

Following the baseline assessment, eligible participants were asked to give written informed consent to randomisation. Randomisation was conducted by means of a computer-generated code, administered centrally and communicated by telephone, thereby concealed in advance from the recruiting researcher. Allocation was stratified by severity of overall symptoms (CIS-R total score < 28 or ≥ 28) and centre using variable block sizes to maximise concealment. The researcher gave the allocated medication to the participant. Neither researchers nor participants were masked to the allocated treatment.

Allocated treatments

Participants randomised to citalopram were prescribed 20 mg daily. Citalopram taken at this dose has been shown to occupy about 80% of serotonin transporter reuptake sites, which is the level of occupancy apparently required to produce reliable antidepressant effects.²⁴ Those randomly allocated to reboxetine were advised to start on 2 mg twice daily and increase this to 4 mg twice daily after about 4 days. Research has shown that acute doses of 4 mg of reboxetine increases cortisol levels indicative of increased noradrenergic function.²⁵ Further, reboxetine at the same dose produces peripheral autonomic effects consistent with noradrenaline reuptake blockade.²⁶

All participants were advised to contact their GP if they wished to increase the dose of their medication.

Outcome measures

Outcome data were recorded 6 and 12 weeks after randomisation. All outcomes were self-administered. The primary outcome was the total BDI score at 6 weeks. Secondary outcomes have been described elsewhere.¹⁸ In the present paper, we also report the proportion 'in remission' (defined as a total BDI score < 10) at 6 weeks, 12-item Short-Form Health Survey (SF-12)²⁷ mental and physical subscale scores and Hospital Anxiety and Depression Scale (HADS)²⁸ total scores. Adherence with medication was assessed at 6 weeks by both self-report and a pill count of returned medication.

Self-administered outcomes were used rather than clinician-administered measures such as the Hamilton Rating Scale for Depression (HRSD)²⁹ because of the potential for bias in a non-masked trial such as this.

Severity of depression

The CIS-R^{20,21} is a fully structured interview that measures 14 symptom groups and, as well as generating ICD-10 diagnoses, the CIS-R also generates a total symptom score (0–57). The assessment asks questions about the week prior to interview and the onset and duration of each episode. In order to address the specified hypothesis, the severity of depression (CIS-R_{dep}) was calculated as the sum of the scores for the following CIS-R items: depression, depressive ideas, fatigue, concentration and sleep (score range: 0–21).

Statistical analysis

The Trial Steering Committee agreed the analysis plan prior to the analyses being conducted. As described earlier, the GENPOD study was designed to test two primary hypotheses. The findings in relation to genetic predictors of treatment response have been reported elsewhere.¹⁷ The description of the analysis that follows is specific to the current hypothesis relating to severity and outcome. All analyses were conducted using Stata version 11.1 for Windows.

The primary analysis was a pre-specified subgroup analysis for the primary outcome of BDI score at 6 weeks, with analysis performed on an intention-to-treat (ITT) basis. A linear regression model was generated that included an interaction term between treatment (citalopram/reboxetine) and the predictor (severity of depression, CIS-R_{dep}) as a continuous variable. The model was adjusted for baseline BDI score, centre and the stratification variable of CIS-R total score (< 28 or ≥ 28). A further secondary analysis was conducted including only those participants who reported taking their allocated medication for at least 4 weeks. Analyses were repeated adjusting for the time

between randomisation and follow-up by incorporating it as a continuous variable in the models. Further adjustment was also made for potential confounders: sociodemographic factors (age, gender and employment status), history of depression, life events and social support. All differences were calculated by subtracting BDI scores for the citalopram group from those in the reboxetine group (so a positive value indicates a worse outcome for reboxetine).

Different instruments have been used to measure depressive symptoms in prior studies making direct comparison of the severity of participants' depression difficult. Therefore, an additional analysis that was not part of the original analysis plan was also conducted to test for differential effects in individuals with more severe depression. This was defined, *a priori*, as those in the top 10% of the CIS-R depression severity scale, CIS-R_{dep}.

Additional analyses (included in the original analysis plan) were undertaken for the secondary outcomes of BDI remission, HADS total score, SF-12 physical and mental subscale scores. Adjustment for baseline score, stratification variables and potential confounders were made as described earlier. For the binary outcome of 'remission', the reported odds ratios are the ratio of the odds of remission for those on reboxetine/citalopram so an odds ratio of more than one indicates a better outcome with reboxetine. Repeated measures linear/logistic regression was used to include data from both 6 and 12 weeks with a variable denoting time (6/12 weeks) in the regression model, together with the other adjustments mentioned in the primary analysis.

Finally, the baseline characteristics of those with or without outcome data at 6 weeks were compared. Missing data may not only result in a loss of precision but also introduces the potential for bias. The impact of missing data, primarily outcome data, on the findings was examined by adjusting for factors associated with 'missingness' in the various regression models. This method should address any bias under a missing at random assumption.³⁰

Justification of the sample size

Details of the sample size calculations for the trial and the impact of the final recruitment figures on the power of the study were given in the protocol paper.¹⁸ The sample size was primarily driven by the genetic hypothesis¹⁷ with a revised target of 570 participants set for the primary analysis. The original power calculations for the severity hypothesis assumed that 44% of participants would be classified as having 'severe' depression. In order to detect a differential effect of 75% remission in the low severity/SSRI and high severity/NARI groups compared with 55% remission in the other two groups, a total of 282 participants was required. Thus, the GENPOD trial was adequately powered to address the question of the differential response to treatment based on the severity of depression (that is, the interaction between severity and treatment allocation).

Results

Trial participation and follow-up

The CONSORT flow chart and baseline comparability of the randomised groups for the GENPOD trial has been published previously.¹⁷ In total, 601 participants were randomised to receive either citalopram ($n=298$) or reboxetine ($n=303$) between October 2005 and February 2008. The mean age of participants was 38.8 years (s.d.=12.4) and 68% ($n=408$) were female. Ninety-two per cent of participants ($n=550$) had moderate ($n=305$) or severe depression ($n=245$) according to ICD-10 criteria.

Ninety-one per cent ($n=546$) completed the 6-week follow-up (citalopram: $n=274$, reboxetine: $n=272$) and 81% ($n=486$)

completed the 12-week follow-up (citalopram: $n=253$, reboxetine: $n=233$).

Adherence to medication

Adherence to medication was higher among those randomised to citalopram at 6 weeks (on medication: citalopram $n=246$ (83%), reboxetine $n=193$ (64%)) and 12 weeks (on medication: citalopram $n=205$ (69%), reboxetine $n=141$ (47%)).

General practitioners increased the dose of citalopram for 55 participants (20%) (from 20 mg to: 30 mg ($n=11$), 40 mg ($n=33$) and to 60 mg ($n=11$)). A smaller number of those allocated to reboxetine ($n=13$ (5%)) also had their dose of medication increased (from 4 mg twice daily to: 10 mg ($n=3$), 12 mg ($n=9$) and to 16 mg ($n=1$)).

Baseline comparability of severity groups

Those with more severe depression were less likely to be employed (full or part time), have experienced more life events in the previous 6 months, and report lower levels of social support compared with those with less severe depression (Table 1). Those with more severe depression were slightly younger, more likely to be male and have a history of depression. However, there were no differences in terms of ethnicity, prior treatment with antidepressants or family history of depression according to severity of depression at baseline (Table 1).

Analysis for the severity hypothesis

The mean BDI scores at 6 weeks are given in Table 2 together with the differences between citalopram and reboxetine according to severity of depression. The primary analysis investigated the interaction between severity (as a continuous variable) and allocated treatment, with BDI score at 6 weeks as the outcome, adjusting for baseline BDI and stratification variables used in randomisation (CIS-R stratum and centre). This provided no evidence of a differential effect (interaction term: 0.02, 95% CI -0.59 to 0.62, $P=0.96$). Adjusting the primary analysis for the time interval between randomisation and the 6-week follow-up had no effect (interaction term: 0.02, 95% CI -0.59 to 0.62, $P=0.95$). Similarly, there was little evidence to support an interaction between treatment and severity of depression following adjustment for potential confounders (age, gender, employment status, history of depression, number of life events and social support) (interaction term: 0.06, 95% CI -0.54 to 0.66, $P=0.85$). Results from analyses that were restricted to the 474 participants who took their allocated medication for at least 4 weeks were consistent with the findings of the primary analysis (interaction term: 0.09, 95% CI -0.56 to 0.74, $P=0.79$ (adjusted for stratification variables, baseline BDI score and the potential confounders listed earlier)).

We also investigated a main effect of severity on outcome, irrespective of treatment, after adjustment for the random allocation, baseline BDI score and stratification variables. On average, a one unit increase in the CIS-R depression severity score was associated with a 0.45 unit increase in BDI score at 6 weeks adjusting for baseline (coefficient 0.45, 95% CI 0.05 to 0.85). However, following adjustment for potential confounders, the magnitude of this effect attenuated and the 95% confidence interval included the null (coefficient: 0.39, 95% CI -0.02 to 0.79).

An additional analysis that was not part of the original analysis plan was also conducted. The most severe 10% of respondents scored ≥ 19 on the CIS-R depression severity scale (CIS-R_{dep}) and this threshold was used to test for differential effects in participants with more severe depression. At baseline, mean BDI scores for the two groups were as follows: CIS-R_{dep} < 19:

Table 1 Characteristics of participants at baseline according to depression severity

	CIS-R depression score < 16 (<i>n</i> = 321)	CIS-R depression score ≥ 16 (<i>n</i> = 280)	<i>P</i> ^a
Age, years: mean (s.d.)	39.6 (12.7)	38.0 (11.9)	0.11
Female, <i>n</i> (%)	224 (69.8)	184 (65.7)	0.29
Ethnicity, <i>n</i> (%)			0.66
White	310 (96.6)	265 (94.6)	
Mixed	1 (0.3)	3 (1.1)	
Asian/British Asian	2 (0.6)	4 (1.4)	
Black/Black British	7 (2.2)	7 (2.5)	
Other	1 (0.3)	1 (0.4)	
History of depression, <i>n</i> (%)	223 (69.7)	212 (75.7)	0.10
Previous antidepressant treatment, <i>n</i> (%)	170 (53.5)	155 (55.4)	0.64
Family history of depression, <i>n</i> (%)	199 (62.2)	178 (63.6)	0.73
Employment status, <i>n</i> (%)			0.02
Working full time	140 (43.6)	103 (36.8)	
Working part time	69 (21.5)	45 (16.1)	
Student	9 (2.8)	14 (5.0)	
Retired	12 (3.7)	9 (3.2)	
Houseperson	42 (13.1)	35 (12.5)	
Unemployed job seeker	18 (5.6)	23 (8.2)	
Unemployed due to ill health	31 (9.7)	51 (18.2)	
Number of life events, mean (s.d.)	1.5 (1.2)	1.9 (1.5)	<0.001
Social support score, mean (s.d.)	12.6 (3.4)	11.3 (4.1)	<0.001
Beck Depression Inventory score, mean (s.d.)	29.5 (8.4)	38.4 (8.8)	<0.001
Suicidal thoughts, <i>n</i> (%)	21 (6.5)	60 (21.4)	<0.001

CIS-R, Clinical Interview Schedule – Revised.
a. *P*-values from chi-squared test for categorical variables or from *t*-test for continuous variables.

Table 2 Means and adjusted differences in Beck Depression Inventory scores at 6 weeks, in the citalopram and reboxetine groups, by depression severity (split into two groups at the median)

CIS-R depression score	Citalopram group		Reboxetine group		Adjusted differences ^a (95% CI)
	<i>n</i>	Mean (s.d.)	<i>n</i>	Mean (s.d.)	
< 16	141	17.0 (10.1)	151	17.3 (10.1)	1.11 (–1.12 to 3.34)
≥ 16	133	20.9 (11.3)	121	22.5 (11.7)	1.18 (–1.55 to 3.90)

CIS-R, Clinical Interview Schedule – Revised.
a. From linear regression models adjusted for centre, baseline severity strata, and baseline BDI score. Difference is reboxetine minus citalopram and higher scores denote a worse outcome.

n = 528, mean BDI score: 32.5 (s.d. = 9.1); CIS-R_{dep} ≥ 19: *n* = 73, mean BDI score: 42.0 (s.d. = 9.6). The mean BDI scores at 6 weeks are given in Table 3 together with the differences between citalopram and reboxetine according to severity of depression. Linear regression analysis was repeated with the BDI at 6 weeks as the outcome, adjusting for baseline BDI score, stratification variables and potential confounders to test for an interaction between allocated medication and severity (CIS-R_{dep} < 19 *v.* ≥ 19). In this analysis, the outcome in the lower severity group was worse for those randomised to reboxetine, whereas reboxetine was better for those with more severe depression (Table 3). There was weak evidence of an interaction between treatment and severity and BDI outcome at 6 weeks (interaction term: –5.07, 95% CI –10.4 to 0.27, *P* = 0.063) following adjustment for stratification variables and potential confounders. Among those with the most severe depression (CIS-R_{dep} ≥ 19), scores on the BDI for those who received reboxetine, on average, improved by 2.7 points more than those who received citalopram (Table 3). However, this improvement represented less than a third of a standard deviation change relative to the baseline BDI score for this group.

The influence of severity on outcome was also examined in repeated measures analyses that included data from both 6- and

12-week follow-ups. Again, there was no evidence to support an interaction between severity as a continuous variable and treatment allocation (interaction term: 0.09, 95% CI –0.47 to 0.65, *P* = 0.75) adjusted for stratification variables and potential confounders, as described earlier. However, in the repeated measures analysis examining the interaction between treatment and a binary indicator of the most severe depressions (CIS-R_{dep} < 19 *v.* = 19), there was some evidence that those with the most severe depression were more likely to benefit from reboxetine (interaction term: –5.28 (95% CI –10.2 to –0.36), *P* = 0.04).

Secondary outcomes

Analyses were repeated for the secondary outcomes of HADS, SF-12 physical and mental subscales scores (online Tables DS1–3), with adjustment for the stratification variables and potential confounders listed earlier. There was no evidence to support an interaction between severity and outcome for HADS (interaction term: 0.15, 95% CI –0.30 to 0.61, *P* = 0.51), SF-12 physical subscale scores (interaction term: 0.09, 95% CI –0.36 to 0.53, *P* = 0.69) or SF-12 mental subscale scores (interaction term: –0.24, 95% CI –0.95 to 0.46, *P* = 0.49). Repeated measures analyses were also conducted using data from both the 6- and 12-week follow-ups as described earlier. Again, there was no evidence to support an

Table 3 Additional analyses to test for differential effects on Beck Depression Inventory scores at 6 weeks in patients with most severe depression (CIS-R depression score ≥ 19)

CIS-R depression score	Citalopram group		Reboxetine group		Adjusted differences ^a (95% CI)
	<i>n</i>	Mean (s.d.)	<i>n</i>	Mean (s.d.)	
< 19	240	18.1 (10.6)	241	19.0 (10.7)	1.72 (−0.08 to 3.53)
≥ 19	34	24.6 (10.9)	31	24.2 (13.4)	−2.69 (−8.71 to 3.33)

CIS-R, Clinical Interview Schedule – Revised.
 a. From linear regression models adjusted for centre, baseline severity strata, and baseline Beck Depression Inventory score. Difference is reboxetine minus citalopram and higher scores denote a worse outcome.

interaction between severity and treatment allocation (interaction terms: HADS 0.10, 95% CI −0.32 to 0.52, $P=0.65$; SF-12 physical subscale 0.30, 95% CI −0.10 to 0.69, $P=0.14$; SF-12 mental subscale −0.24, 95% CI −0.86 to 0.37, $P=0.44$).

Finally, the proportion of patients in ‘remission’ (BDI score < 10) at 6 weeks is presented in Table 4 according to severity stratum. There was no evidence of a differential effect of severity on remission outcomes (interaction odds ratio (OR) = 0.99, 95% CI 0.85–1.15, $P=0.88$) adjusted for all stratification variables and potential confounders, as previously). Results of repeated measures analyses also showed no evidence to support an interaction (interaction OR = 0.92, 95% CI 0.74–1.14, $P=0.43$).

Missing data

Attrition was low, with less than 10% of participants lost to follow-up 6 weeks after randomisation. However, those lost to follow-up at 6 weeks were younger and reported more life events at baseline (online Table DS4). Both of these factors were included in the list of potential confounders that were adjusted for in the main analyses, and hence no further analyses were deemed necessary.

Discussion

Main findings

The study was designed to test the hypothesis that individuals with more severe depression would benefit more from NARIs than SSRIs. However, there was very little evidence to support this. Most of the participants in GENPOD had moderate or severe depression according to ICD-10 criteria. Only when we restricted the analysis to examine treatment response among the most severely affected (top 10% CIS-R_{dep}: mean BDI score 42) was there weak evidence to support the hypothesis. Even among this group, the mean difference in BDI scores after 6 weeks of treatment for those on reboxetine compared with citalopram represented a reduction of less than a third of a standard deviation, which is an effect size often cited as clinically relevant. It is possible that a more refined method of identifying a subgroup of individuals with depression – for example, by identifying those with endogenous or biological symptoms – could be of value. However, we can conclude that, for the vast majority of those attending UK primary care with depression, which is the setting in which most

depression is treated, treatment with NARIs does not confer any advantage for outcome in those with more severe illness.

Strengths and limitations

The main strength of the study was the measures taken to reduce the possibility of a chance finding. This included prior publication of the trial protocol and hypotheses, and agreement of a pre-specified analysis plan that restricted opportunities for multiple testing and selective reporting. Furthermore, the sample size calculation for the GENPOD study was primarily driven by the genetic hypothesis,¹⁷ and hence the number of participants recruited was more than adequate to ensure that the study was adequately powered to address the hypothesis relating to severity and treatment response. In fact, the study recruited more than twice the number of participants that were required for these analyses (due to the larger numbers required for the other main hypothesis).

Non-adherence to the allocated treatment may have resulted in an underestimate of the influence of severity on outcome. Nevertheless, when analyses were restricted to those who took at least 4 weeks of medication, the findings did not differ. Follow-up rates were high (91% at 6 weeks), but even a small amount of missing data may introduce bias. However, adjustment for factors associated with missing data did not materially affect the findings.

Allocated medications were prescribed at doses that are standard for UK primary care (<http://bnf.org/bnf/index.htm>). General practitioners retained clinical responsibility for patient care throughout the GENPOD study and were free to increase the dose of allocated medication where appropriate and, indeed, the dose of medication was increased for 20% of GENPOD participants. The majority of participants in GENPOD were taking a lower dose of citalopram than used in the large US STAR*D study (primary care settings: mean exit dose 40.6 mg/day).⁴ Generally, 20 mg/day of citalopram is regarded as the minimum effective dose³¹ and there is no evidence that 40 mg/day is more effective at reducing depressive symptoms than 20 mg/day.³² However, citalopram is a racemic mixture and there is evidence that its action at the serotonin transporter may be inhibited by the otherwise inactive R-enantiomer.³² This may have implications for the efficacy of citalopram relative to certain other antidepressants.⁷

Table 4 Proportion of participants with a Beck Depression Inventory (BDI) score of less than 10 (‘remission’) at 6 weeks

CIS-R depression score	Citalopram			Reboxetine		
	Total <i>n</i>	BDI < 10		Total <i>n</i>	BDI < 10	
		<i>n</i>	% (95% CI)		<i>n</i>	% (95% CI)
< 16	141	39	27.7 (20.5–35.8)	151	41	27.2 (20.2–35.0)
≥ 16	133	21	15.8 (10.0–23.1)	121	20	16.5 (10.4–24.4)

CIS-R, Clinical Interview Schedule – Revised.

Finally, it is important to consider the generalisability of these trial findings. It is difficult to recruit a truly representative sample of individuals with depression. As has been previously reported,¹⁷ GPs referred a total of 842 patients to the GENPOD study over a period of 28 months. The mean BDI scores in GENPOD are similar to other UK depression trials conducted in primary care^{33,34} and the large GENDEP study that ran in nine European centres.¹⁴ In addition, the response rate to antidepressants in GENPOD was similar to the large US STAR*D study,⁴ which aimed to recruit a representative sample of individuals with depression in the US healthcare system. Furthermore, published formula³⁵ enable the conversion of HRSD scores commonly used in US studies with the CIS-R scores available for GENPOD. The mean HRSD score of 22–23 reported in STAR*D⁴ and other major US trials of cognitive-behavioural therapy *v.* antidepressant medication³⁶ or combinations of medication for depression³⁷ is equivalent to a CIS-R score of 23–24, which is slightly lower than the mean of 31 in GENPOD.¹⁷ Thus, it would seem reasonable to generalise our findings to other European and US populations of people with depression seen in primary care and out-patient settings.

Comparisons with existing literature

Others have found that NARIs are more effective in the most severely affected individuals.^{11,12} Massana *et al*¹¹ found that, among participants who were rated the most severely ill at baseline, those randomised to reboxetine experienced a greater reduction in scores on the HRSD rating scale compared with those randomised to receive fluoxetine for 8 weeks. However, no formal test of an interaction of treatment by severity was undertaken. Venditti *et al*¹² combined data from two 8-week multinational trials with a specific focus on the outcome of self-rated social functioning and found that reboxetine was more effective in improving social functioning for the more severely ill participants after 4 weeks of treatment, although this difference diminished with time. However, Venditti *et al*¹² incorporated the severity of depression in the participants measured during the trial in the regression model rather than focusing on a baseline measure of severity as a predictor of outcome.

Different instruments were used to measure depressive symptoms in these studies, thus it is difficult to directly compare the severity of participants' depression. However, converting scores as above demonstrates that the samples in these trials were similar to GENPOD (mean HRSD: 27–29 equivalent to CIS-R scores 29–31).³⁵ Nonetheless, in the findings described above, severity of depression was measured using the Clinician Global Impression – Severity of Illness scale (CGI-SI), and it is difficult to directly compare this measure to the CIS-R depression scale used in the present analyses. Although the CGI-SI will reflect the severity of symptoms, such a rating may also capture the impact of the depressive symptoms on the individual. Thus, it may be that the anecdotal clinical evidence that those with more severe depression are more likely to respond to treatment with NARIs may reflect a more global view of severity, which would align with the CGI-SI measure and thus the evidence from prior studies^{11,12} rather than GENPOD.

Others have suggested that, rather than improve mood, different antidepressants may lead to improvement in different symptoms. For example, in the GENDEP study, the tricyclic antidepressant nortriptyline led to more improvement in sleep, appetite and sexual interest and less improvement in mood and cognitive symptoms.¹⁴ It is possible that such effects may have been overlooked in the present analyses as the measure of severity used combined scores on a range of symptoms (including depressive ideas, fatigue, concentration and sleep).

As discussed earlier, a recent meta-analysis⁸ suggested that reboxetine was less effective than either placebo or SSRIs for the treatment of depression but this was at odds with an earlier publication that reported no difference in response rates between reboxetine and SSRIs.⁹ Moreover, there was evidence that reboxetine was more effective than placebo among in-patients⁸ who would be likely to have more severe depression.⁸ Even if we assume that reboxetine is a less effective treatment, then this would have made it more likely that we would have found an interaction between severity of depression and treatment allocation, in contrast to our findings. Moreover, drop-out from treatment is higher with reboxetine compared with SSRIs.^{7,8} Such differential non-adherence needs to be accounted for in order to obtain valid estimates of the comparative efficacy of different treatments. Methods to allow for this have recently been proposed³⁸ but, to the best of our knowledge, no such allowances were made in the previous comparisons.^{7,8} Indeed, the approach by Cipriani *et al*⁷ of assuming non-response to treatment may, in the presence of differential adherence, introduce bias such that reboxetine appears less effective.

Implications and further research

The ability to tailor treatments is key in the drive towards stratified medicine.⁵ Not only would this speed recovery for individuals but it would have economic benefits. The GENPOD study was designed to test the hypothesis that individuals with more severe depression (measured at entry to the study) would benefit more from NARIs rather than SSRIs. However, there was very little evidence to support this in individuals with depression seen in primary care. Given the multifaceted nature of depressive illness, perhaps it is not surprising that a single factor in isolation is not predictive of treatment response. More consideration of the wider picture in terms of biological, psychological and social factors may be required in order to move towards the goal of tailoring treatment. Furthermore, it may be useful to focus on examining different patterns of symptoms that may have greater value in discriminating outcomes.

Funding

The study was funded by the Medical Research Council (grant reference: G0200243).

Acknowledgements

The study was supported by the Mental Health Research Network. We are grateful for the support of the patients who agreed to participate and their general practitioners. We would like to thank the members of the Trial Steering Committee: Ian Anderson (Chair), Tony Johnson, John Geddes and Rodney Elgie and of the Data Monitoring Committee: Linda Gask (Chair), Nick Freemantle and Irwin Nazareth. We thank the people who contributed towards the fieldwork, including the following: Helen Lester, Laura Webber, Morag Turnbull, Louise Paterson, Ben Newton, Alex Smith, Nicola Morris, Leigh Franks, Joy Farrimond, Nathan Filer, Caitlin Jarrett and Angela Hill.

Nicola J. Wiles, BSc, PhD, **Jean Mulligan**, BSc, PhD, School of Social and Community Medicine, University of Bristol, Bristol; **Tim J. Peters**, BSc, MSc, PhD, CStat, FFPH, School of Social and Community Medicine and School of Clinical Sciences, University of Bristol, Bristol; **Philip J. Cowen**, MD, FRCPsych, Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford; **Victoria Mason**, BSc, MSc, PhD, CPsychol, Psychological Sciences, University of Worcester, Worcester; **David Nutt**, DM, FRCP, FRCPsych, FMedSci, Department of Neuropsychopharmacology, Imperial College, London; **Deborah Sharp**, MA, BM, BCh, PhD, FRCGP, **Debbie Tallon**, BSc, MSc, **Laura Thomas**, BA, MPhil, School of Social and Community Medicine, University of Bristol, Bristol; **Michael C. O'Donovan**, PhD, FRCPsych, MRC Centre for Neuropsychiatric Genetics and Genomics, School of Medicine, Cardiff University, Cardiff; **Glyn Lewis**, PhD, FRCPsych, School of Social and Community Medicine, University of Bristol, Bristol, UK

Correspondence: Nicola J. Wiles, Academic Unit of Psychiatry, School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Clifton, Bristol BS8 2BN, UK. Email: nicola.wiles@bristol.ac.uk

First received 23 Dec 2010, final revision 2 Aug 2011, accepted 15 Sep 2011

References

- 1 Prescribing Support Unit. *Prescribing Monitoring Report April 2007–March 2008*. Information Centre for Health and Social Care, 2008.
- 2 McManus P, Mant A, Mitchell PB, Montgomery WS, Marley J, Auland ME. Recent trends in the use of antidepressant drugs in Australia. *Med J Aust* 2000; **173**: 458–61.
- 3 Pincus HA, Tanielian TL, Marcus SC, Olfon M, Zarin DA, Thompson J, et al. Prescribing trends in psychotropic medications. *JAMA* 1998; **279**: 526–31.
- 4 Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry* 2006; **163**: 28–40.
- 5 Trusheim MR, Berndt ER, Douglas FL. Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. *Nat Rev Drug Discov* 2007; **6**: 287–93.
- 6 Freemantle N, Anderson IM, Young P. Predictive value of pharmacological activity for the relative efficacy of antidepressant drugs. Meta-regression analysis. *Br J Psychiatry* 2000; **177**: 292–302.
- 7 Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JPT, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009; **373**: 746–58.
- 8 Eyding D, Leigemann M, Grouven U, Harter M, Kromp M, Kaiser T, et al. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. *BMJ* 2010; **341**: c4737.
- 9 Papakostas GI, Nelson JC, Kasper S, Moller H-J. A meta-analysis of clinical trials comparing reboxetine, a norepinephrine reuptake inhibitor, with selective serotonin reuptake inhibitors for the treatment of major depressive disorder. *Eur Neuropsychopharmacol* 2008; **18**: 122–7.
- 10 Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* 2008; **5**: e45.
- 11 Massana J, Moller H-J, Burrows GD, Montenegro RM. Reboxetine: a double-blind comparison with fluoxetine in major depression. *Int Clin Psychopharmacol* 1999; **14**: 73–80.
- 12 Venditti LN, Arcelus A, Birnbaum H, Greenberg P, Barr CE, Rowland C, et al. The impact of antidepressant use on social functioning: reboxetine versus fluoxetine. *Int Clin Psychopharmacol* 2000; **15**: 279–89.
- 13 Brown WA, Dornseif BE, Wernicke JF. Placebo response in depression: a search for predictors. *Psychiatry Res* 1988; **26**: 259–64.
- 14 Uher R, Maier W, Hauser J, Marušić A, Schmael C, Mors O, et al. Differential efficacy of escitalopram and nortriptyline on dimensional measures of depression. *Br J Psychiatry* 2009; **194**: 252–9.
- 15 Bielski RJ, Friedel RO. Prediction of tricyclic antidepressant response. *Arch Gen Psychiatry* 1976; **33**: 1479–89.
- 16 Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**: c869.
- 17 Lewis G, Mulligan J, Wiles N, Cowen PJ, Craddock N, Ikeda M, et al. Polymorphisms of the 5-HT transporter and response to antidepressants: randomised controlled trial. *Br J Psychiatry* 2011; **198**: 464–71.
- 18 Thomas L, Mulligan J, Mason V, Tallon D, Wiles N, Cowen PJ, et al. Genetic and clinical predictors of treatment response in depression: the GENPOD randomised trial protocol. *Trials* 2008; **9**: 29.
- 19 World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. WHO, 1992.
- 20 Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychol Med* 1992; **22**: 465–86.
- 21 Lewis G. Assessing psychiatric disorder with a human interviewer or a computer. *J Epidemiol Community Health* 1994; **48**: 207–10.
- 22 Beck A, Ward C, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961; **4**: 561–71.
- 23 Beck A, Steer RA, Brown GK. *Beck Depression Inventory (2nd edn)*. The Psychological Corporation, 1996.
- 24 Meyer JH, Wilson AA, Ginovart N, Goulding V, Hussey D, Hood K, et al. Occupancy of serotonin transporters by paroxetine and citalopram during treatment of depression: a [¹¹C]DASB PET imaging study. *Am J Psychiatry* 2001; **158**: 1843–9.
- 25 Hill SA, Taylor MJ, Harmer CJ, Cowen PJ. Acute reboxetine administration increases plasma and salivary cortisol. *J Psychopharmacol (Oxf)* 2003; **17**: 273–5.
- 26 Szabadi E, Bradshaw CM, Boston PF, Langley RW. The human pharmacology of reboxetine. *Hum Psychopharmacol* 1998; **13** (suppl 1): S3–12.
- 27 Jenkinson C, Layte R. Development and testing of the UK SF-12 (Short Form Health Survey). *J Health Serv Res Policy* 1997; **2**: 14–8.
- 28 Zigmond AS, Snaith RP. The Hospital Anxiety And Depression Scale. *Acta Psychiatr Scand* 1983; **67**: 361–70.
- 29 Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; **23**: 56–62.
- 30 Carpenter J, Kenward M. *Guidelines for Handling Missing Data in Social Science Research*. London School of Hygiene and Tropical Medicine, no date (<http://missingdata.lshtm.ac.uk/downloads/guidelines.pdf>).
- 31 Montgomery SA, Pedersen V, Tanghøj P, Rasmussen C, Rioux P. The optimal dosing regimen for citalopram – a meta-analysis of nine placebo-controlled studies. *Int Clin Psychopharmacol* 1994; **9** (suppl 1): 35–40.
- 32 Zhong H, Hansen KB, Boyle NJ, Han K, Muske G, Huang X, et al. An allosteric binding site at the human serotonin transporter mediates the inhibition of escitalopram by R-citalopram: Kinetic studies with the ALI/VFL-SI/TT mutant. *Neurosci Lett* 2009; **462**: 207–12.
- 33 Ward E, King M, Lloyd M, Bower P, Sibbald B, Farrelly S, et al. Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. I: clinical effectiveness. *BMJ* 2000; **321**: 1383–8.
- 34 Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, et al. Therapist-delivered internet psychotherapy for depression in primary care: a randomised controlled trial. *Lancet* 2009; **374**: 628–34.
- 35 Hotopf M, Sharp D, Lewis G. What's in a name? A comparison of four psychiatric assessments. *Soc Psychiatry Psychiatr Epidemiol* 1998; **33**: 27–31.
- 36 DeRubeis RJ, Hollon SD, Amsterdam JD, Shelton RC, Young PR, Salomon RM, et al. Cognitive therapy vs medications in the treatment of moderate to severe depression. *Arch Gen Psychiatry* 2005; **62**: 409–16.
- 37 Blier P, Ward HE, Tremblay P, Laberge L, Hebert C, Bergeron R. Combination of antidepressant medications from treatment initiation for major depressive disorder: a double-blind randomized study. *Am J Psychiatry* 2010; **167**: 281–8.
- 38 Fischer K, Goetghebuer E, Vrijens B, White IR. A structural mean model to allow for noncompliance in a randomized trial comparing 2 active treatments. *Biostatistics* 2010; **12**: 247–57.

