THE LANCET
Psychiatry

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Lewis G, Duffy L, Ades A, et al. The clinical effectiveness of sertraline in primary care and the role of depression severity and duration: a pragmatic, double-blind, placebo-controlled randomised trial. *Lancet Psychiatry* 2019; published online Sept 19. https://doi.org/10.1016/S2215-0366(19)30366-9.

Title: The clinical effectiveness of sertraline in primary care and the role of depression severity and duration: a pragmatic randomised controlled trial.

Running head: Antidepressants in primary care: the PANDA trial.

Supplementary Appendix

Authors

Gemma Lewis, PhD[1]
Larisa Duffy, BSc[1]
Anthony Ades, PhD[2]
Rebekah Amos, BSc[7]
Ricardo Araya, PhD[3]
Sally Brabyn, MSc[8]
Katherine S. Button, PhD[4]
Rachel Churchill, PhD[5]
Catherine Derrick, BSc[2]
Christopher Dowrick, MD[7]
Simon Gilbody, PhD[8]
Christopher Fawsitt, PhD[2]
William Hollingworth, PhD[2]
Vivien Jones, BA[2]
Tony Kendrick, MD[9]
David Kessler, PhD[2]
Daphne Kounali, PhD[2]
Naila Khan, PhD[7]
Paul Lanham, BA[1]
Jodi Pervin, BSc[8]
Tim J. Peters, PhD[2]
Derek Riozzie[7]
George Salaminios, MSc[1]
Laura Thomas, MSc[2]
Nicky J. Welton, PhD[2]
Nicola Wiles, PhD[2]
Rebecca Woodhouse, MSc[8]
Glyn Lewis, PhD[1]

Affiliations

[1]Division of Psychiatry, University College London, 6th Floor Maple House, 149 Tottenham Court Road, London W1T 7NF, UK.
[2]Bristol Medical School, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK.
[3]Institute of Psychiatry, Psychology and Neuroscience, Kings' College London, Denmark Hill, London SE5 8AF, UK.
[4]Department of Psychology, University of Bath, 10 West, Bath BA2 7AT, UK.
[5]Centre for Reviews and Dissemination, University of York, Heslington, York YO10 5DD, UK.
[6]School of Nursing and Health Studies, University of Dundee, Airlie Place, Dundee DD1 4HJ, UK.
[7]Institute of Psychology Health and Society, University of Liverpool, Waterhouse Building Block B, Liverpool L69 3BX, UK.
[8]Department of Health Sciences, University of York, Seebohm Rowntree Building, Heslington, York YO10 5DD, UK.
[9]Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Aldermoor Health Centre, Southampton SO16 5ST, UK.
[10]Department of Sociology, Social Policy and Criminology, University of Liverpool, Eleanor Rathbone Building, Bedford Street South, Liverpool L69 7ZA, UK.

Correspondence to:

Dr Gemma Lewis, UCL Division of Psychiatry, 6th Floor, Maple House, 149 Tottenham Court Road, London W1T 7NF, gemma.lewis@ucl.ac.uk, 0203 108 7183.

Contents

Methods for statistical analyses reported in the Supplement

We described characteristics of the General Practice surgeries by reporting the percentage of practices with each characteristic. We compared the age and gender of the patients we invited to take part in the trial with those who actually participated by reporting the mean and SD for age and the number and percentage for gender (according to 'invited' versus 'participated').

To identify baseline demographic and clinical variables associated with missing data we made a binary variable at each follow-up time-point coded 0 if the PHQ score was not missing and 1 if the PHQ score was missing. We calculated odds ratios and 95% confidence intervals for each characteristic, at each follow-up, using univariable logistic regression models.

Methods for post-hoc (unplanned) analyses

We examined physical health comorbidities according to treatment allocation, in a descriptive analysis. We also compared the characteristics of patients recruited through GP consultation with those recruited through the record searches. We used a cut-off of <=4 for remission on the PHQ-9 and analysed this as a repeated measures secondary outcome. We conducted sub-group analyses to further investigate severity and duration. We also conducted analyses in which the outcome was absolute depression score (e.g. non-log transformed PHQ-9) and calculated interactions between treatment allocation and CIS-R depression severity score and duration. Finally we examined response as an outcome (50% reduction in depression score compared to baseline).

All sub-group analyses were conducted by running regression models stratified by the sub-group variable. We also report p values for interactions between each sub-group variable and

treatment allocation. For sub-group analyses according to the severity of depressive symptoms at baseline we use absolute (not log transformed) PHQ-9 or GAD-7 scores as the outcome, to investigate if mean rather than proportional differences varied by baseline severity.

We calculated logistic multilevel models using response (at least a 50% reduction in symptoms from baseline) across follow-up time-points as a binary outcome. This model was adjusted for PHQ and CISR depression severity scores at baseline and the stratification variables (severity assessed by CISR in three categories, duration in 2 categories and centre).

Results of post-hoc analyses

We found that 272 (42%) participants reported long-standing illness, disability or infirmity. Most prevalent were a mental health problem (21%), asthma or progressive lung disease (19%), diabetes (8%), arthritis (8%) and heart problems (4%). Prevalence of stroke, cancer or kidney disease was under 1% (Table S9). Patients recruited from GP consultations had more severe depressive and anxiety symptoms than those recruited through record searches, but were less likely to have taken antidepressants before (Table S10). We carried out sub-group analyses of the primary outcome and found no evidence that the treatment effect varied by method of recruitment or whether patients had used antidepressants or had depression in the past, Table S11. We examined other sub-groups defined by severity of depressive symptoms or ICD-10 diagnosis of depression. We found no evidence for a treatment effect on our primary outcome in any of these sub-groups, Table S12. There was also no evidence that the treatment effect varied according to depressive symptom duration, Table S13. We conducted similar analyses using GAD-7 scores as the outcome and found no evidence for an influence of severity or duration on the treatment effect, Tables S14 and S15.

We also analysed data using PHQ-9 and GAD-7 scores without log transforming and came to similar conclusions. The adjusted difference in PHQ-9 means between sertraline and placebo at 6 weeks was -0.51 (95% CI -1.33 to 0.31), equivalent to a standardised difference of -0.09 (95% CI -0.23 to 0.05). The adjusted difference in PHQ-9 means at 12 weeks was -1.07 (95% CI -1.96 to -0.19), a standardised difference of -0.19 (95% CI -0.33 to -0.03). We did not find any evidence that the treatment response varied with depression severity (p=0.89) or duration (p=0.73) in this model. The adjusted difference in GAD-7 means between sertraline and placebo at 6 weeks was -1.25 (-1.98 to -0.52), a standardised difference of -0.24 (95% -0.38 to -0.10). At 12 weeks this difference was -1.30 (95% CI -2.07 to -0.53), a standardised difference of -0.24 (95% CI -0.39 to -0.10). We found no evidence that treatment response varied with baseline severity (p=0.72) or duration (p=0.42) in this model.

The odds of response were 1.58 (95% CI 1.05 to 2.37, p=0.028) times higher in those taking sertraline compared to placebo, with a suggestion that this difference increased over time (p=0.062), Table S16. Evidence for an effect of sertraline on GAD-7 response was strong (adjusted odds ratio 2.51, 95% CI 1.58 to 4.00, p<.0001), with an indication that this became larger over time (p=0.094).

Table S1. Characteristics of the 179 General Practice surgeries used for recruitment

Characteristic	Percentage of practices with characteristic
Centre	
Bristol	31%
Liverpool	17%
London	40%
York	12%
Geographical location[a]	
Urban	86%
Rural	14%
List size[b]	
1-4999	12%
5000-9999	41%
10,000-14,999	24%
15,000+	23%
Number of GPs employed	
0-5	37%
6-10	41%
11-15	17%
16+	5%
Number of patients randomised	
0-4	74%
5-12	21%
13-20	4%
21+	1%
Index of Multiple Deprivation[c]	
1-3	20%
4-6	28%
7-10	52%

[a]Based on the 2011 rural-urban classification for output areas in England

[b]Number of patients enrolled in practice

[c]The Index of Multiple Deprivation combines UK national census information from 38 indicators into seven domains of deprivation (income; employment; health and disability; education, skills, and training; barriers to housing and services; living environment; and crime). This results in a deprivation score for each 32,482 'lower super output area' in England, geographical units used for the reporting of neighbourhood level statistics.

Table S2. Comparing the age and gender of the participants who were invited with those who participated.

	Age in years			Gender		
	Number	Mean	SD	Number	Female (n)	Female (%)
Trial participants[a]	653	39.7	15.0	653	384	61
Patients identified as potentially eligible and invited to participate[b]	11,636[c]	37.3	13.8	11,561[c]	7,001	59

[a]Included in the trial

[b]Identified as potentially eligible during the database search and sent an invitation letter. These data were provided by 53 of the 179 GP practices.

[c]A subset of the total number of participants who were identified as eligible and sent an invitation letter (31,645). This subset was comprised from the practices who returned these data.

Table S3. Univariable associations between baseline demographic and clinical variables and missing data on the PHQ-9 at each follow-up (binary outcome; 0=not missing, 1=missing). Odds ratio >1 indicates increased odds of missing data, <1 decreased odds.

Baseline variable	Odds ratio (95% confidence interval) p value at each follow-up		
	2 week (84/653, 13%, missing)	6 week (102/653, 16%, missing)	12 week (128/653, 20%, missing)
Centre (n=653)			
Bristol	**Ref**	**Ref**	**Ref**
Liverpool	0.87 (0.40 to 1.86) 0.715	0.76 (.38 to 1.52) 0.435	1.17 (0.65 to 2.11) 0.597
London	3.60 (2.09 to 6.22) <0.0001	3.36 (2.04 to 5.52) <0.0001	3.45 (2.14 to 5.56) <0.0001
York	0.60 (0.24 to 1.37) 0.227	0.37 (0.16 to 0.87) 0.022	0.68 (0.35 to 1.30) 0.242
Age (n=653)			
18-34	**Ref**	**Ref**	**Ref**
35-54	0.86 (0.53 to 1.40) 0.536	0.77 (0.48 to 1.21) 0.254	0.62 (0.40 to 0.94) 0.024
55-74	0.37 (0.17 to 0.81) 0.013	0.49 (0.25 to 0.94) 0.018	0.32 (0.17 to 0.61) <0.0001
Sex (n=653)			
Female	**Ref**	**Ref**	**Ref**
Male	1.38 (0.86 to 2.23) 0.185	1.10 (0.72 to 1.70) 0.567	1.21 (0.81 to 1.80) 0.344
Ethnic (n=652)			
White	**Ref**	**Ref**	**Ref**
Ethnic Minority	2.78 (1.55 to 4.98) 0.001	3.15 (1.83 to 5.42) <0.0001	2.88 (1.65 to 4.68) <0.0001
Marital status (n=652)			
Married or living as married	**Ref**	**Ref**	**Ref**
Single	1.72 (1.04 to 2.83) 0.035	1.95 (1.22 to 3.12) 0.007	2.08 (1.35 to 3.20) 0.001
Separated, divorced/ widowed	0.63 (0.26 to 1.49) 0.293	0.62 (0.28 to 1.40) 0.252	0.70 (0.34 to 1.43) 0.324
Employment status (n=652)			
In paid employment	**Ref**	**Ref**	**Ref**
Not employed	1.25 (0.78 to 2.01) 0.350	1.48 (0.96 to 2.27) 0.078	1.19 (0.79 to 1.78) 0.403
Financial difficulty (n=652)			
Comfortable or doing alright	Ref	Ref	Ref
Just about getting by	1.74 (1.05 to 2.92) 0.033	2.04 (1.27 to 3.27) 0.003	2.26 (1.4 to 3.48) <0.0001
Difficult or very difficult	2.39 (1.26 to 4.51) 0.007	2.70 (1.49 to 4.89) 0.001	2.88 (1.55 to 5.00) <0.0001
Highest educational qualification (n=652)			
A Level or higher	**Ref**	**Ref**	**Ref**
GCSE, standard grade/other	0.76 (0.43 to 1.33) 0.336	1.27 (0.78 to 2.06) 0.326	1.30 (0.84 to 2.01) 0.245
No formal qualification	1.14 (0.42 to 3.06) 0.797	2.67 (1.21 to 5.88) 0.015	2.31 (1.08 to 4.96) 0.031
CIS-R total score (n=652)			

0 to 11	**Ref**	**Ref**	**Ref**
12 to 20	1.29 (0.61 to 2.74) 0.502	0.81 (0.42 to 1.57) 0.527	0.84 (0.46 to 1.53) 0.572
20 to 49	1.59 (0.82 to 3.09) 0.172	1.17 (0.67 to 2.04) 0.593	1.14 (0.68 to 1.90) 0.628
CIS-R depression duration (n=652)			
Less than 2 years	**Ref**	**Ref**	**Ref**
2 years or more	1.16 (0.72 to 1.88) 0.538	0.83 (0.52 to 1.32) 0.432	0.92 (0.61 to 1.39) 0.683
ICD-10 CIS-R depression diagnosis (n=652)			
No	**Ref**	**Ref**	**Ref**
Yes	1.23 (0.80 to 1.90) 0.335	1.20 (0.75 to 1.91) 0.444	1.23 (0.83 to 1.82) 0.294
ICD-10 CIS-R anxiety diagnosis (n=652)			
No	Ref	Ref	Ref
Yes	1.16 (0.76 to 1.77) 0.486	1.51 (0.95 to 2.39) 0.081	1.18 (0.80 to 1.74) 0.395
Depression in the past (n=652)			
No	**Ref**	**Ref**	**Ref**
Yes	0.83 (0.50 to 1.38) 0.473	0.91 (0.51 to 1.58) 0.714	0.97 (0.60 to 1.57) 0.906
Antidepressant in the past (n=652)			
No	**Ref**	**Ref**	**Ref**
Yes	0.78 (0.51 to 1.20) 0.256	0.63 (0.40 to 0.99) 0.047	0.83 (0.56 to 1.22) 0.338
PHQ total score (n=651)	1.04 (1.00 to 1.08) 0.050	1.02 (0.99 to 1.07) 0.223	1.05 (1.02 to 1.09) 0.002
CIS-R depression severity score (n=652)	1.03 (0.98 to 1.08) 0.205	1.02 (0.98 to 1.07) 0.339	1.04 (0.99 to 1.08) 0.082
BDI-II total score (n=652)	1.03 (1.01 to 1.05) 0.006	1.02 (1.00 to 1.05) 0.032	1.03 (1.01 to 1.05) 0.002
Social support score (n=652)	1.03 (0.98 to 1.09) 0.166	0.97 (0.92 to 1.03) 0.326	0.98 (0.93 to 1.03) 0.334
Number of life events in past 6 months (n=652)	0.70 (0.60 to 0.83) <0.0001	1.25 (1.05 to 1.50) 0.013	1.40 (1.20 to 1.63) <.0001
GAD-7 score (n=652)	1.05 (1.01 to 1.10) 0.011	1.04 (1.00 to 1.09) 0.050	1.06 (1.02 to 1.10) 0.002
SF-12 mental health subscale (n=649)	0.99 (0.97 to 1.01) 0.178	0.98 (0.96 to 1.01) 0.157	0.98 (0.96 to 1.00) 0.050
SF-12 physical health subscale (n=649)	0.98 (0.96 to 1.00) 0.045	1.00 (0.98 to 1.03) 0.774	0.99 (0.97 to 1.01) 0.394

Table S4: Adherence to medication and beliefs about medication according to treatment allocation 2, 6 and 12 weeks after randomisation.

Questionnaire item/s	2 weeks, n (%)		P value[c]	6 weeks, n (%)		P value[c]	12 weeks, n (%)		P value[c]
	Placebo	Sertraline		Placebo	Sertraline		Placebo	Sertraline	
>=80% adherence[a]	262 (97)	241 (96)	0.376	245 (89)	219 (85)	0.149	211 (82)	204 (80)	0.608
Have started taking medication[b]	272 (93)	251 (90)	0.128	278 (98)	259 (97)	0.689	261 (98)	258 (97)	0.987
Currently taking medication[b]	269 (99)	243 (97)	0.188	264 (95)	239 (93)	0.262	222 (85)	204 (80)	0.109
Taken their tablets every day or taken nearly all their tablets[b]	268 (99)	243 (97)	0.190	267 (97)	242 (94)	0.108	243 (94)	240 (94)	0.973
Thought they were taking sertraline	30 (11)	74 (31)	<0.0001	52 (19)	115 (46)	<0.0001	42 (17)	123 (50)	<0.0001

[a]We used a five-item self-report adherence scale developed for the CoBalt study (reference 29 in main manuscript). >=80% adherence is a binary variable, with a score of 0 (range 0–4) indicating at least 80% adherence.

[b]Additional adherence items we included in the questionnaires

[c]Calculated using Pearson's chi-squared test

Table S5. Serious adverse events

Allocation	Brief description of event	SAE	Seriousness [a]	Related to IMP	Outcome [b]
Sertraline	Suicidal ideation	Yes	6	Possibly	1
Placebo	Hospitalised for physical illness	Yes	3	Not related	1
Sertraline	Non-cardiac chest pain	Yes	3	Unlikely	1

Abbreviations: IMP - Investigational medicinal product; SAE – serious adverse event

[a]Seriousness: 1=Resulted in Death, 2=life Threatening, 3=required inpatient or prolonged existing hospitalisation, 4=resulted in persistent or significant disability/incapacity, 5=resulted in congenital anomaly/birth defect, 6= Important Medical Event.

[b]Outcome: 1= Resolved, 2 = Resolved with sequelae, 3 = Unresolved, 4= Worsening, 5 = Fatal, 6= not assessable.

Table S6. Physical symptoms that are potential SSRI side-effects[a] according to treatment allocation 2, 6 and 12 weeks after randomisation.

	Mean (SD) number of symptoms		P	Mean (SD) frequency of symptoms		P
	Placebo	Sertraline		Placebo	Sertraline	
2 weeks	7.68 (4.85)	7.98 (4.54)	0.4513	39.38 (9.06)	40.09 (9.07)	0.3504
6 weeks	7.10 (5.00)	7.41 (4.63)	0.4546	38.45 (8.83)	39.06 (8.61)	0.4175
12 weeks	6.77 (5.27)	6.78 (4.86)	0.9863	48.77 (17.98)	49.09 (17.32)	0.8360

[a] Recorded with a modified Toronto scale.

Table S7. Repeated measures analyses of continuous secondary outcomes at 2, 6, 12 weeks, adjusted for variables not balanced at baseline.

Outcome	Sertraline		Placebo		Adjusted proportional difference[a] (95% CI)	P value
PHQ-9 (n=547)	n	Mean (SD)	n	Mean (SD)		
2 weeks	277	9.94 (5.83)	292	10.32 (5.55)	.96 (.87 to 1.07)	
6 weeks	266	7.98 (5.63)	284	8.76 (5.86)	.95 (.86 to 1.06)	
12 weeks	262	6.90 (5.83)	263	8.02 (6.12)	.87 (.78 to .97)	
Average over time					.93 (.86 to 1.01)	.075
Group by time interaction						.093
BDI-II (n=540)						
2 weeks	273	18.77 (11.08)	286	19.10 (11.17)	.98 (.89 to 1.10)	
6 weeks	266	14.82 (10.44)	285	15.91 (10.74)	.95 (.85 to 1.07)	
12 weeks	256	12.44 (10.96)	259	14.78 (11.70)	.84 (.74 to .94)	
Average over time					.92 (.84 to 1.02)	.10
Group by time interaction						.015
GAD-7 (n=546)						
2 weeks	277	7.55 (5.49)	291	8.16 (5.26)	.91 (.81 to 1.02)	
6 weeks	264	5.55 (5.19)	284	6.96 (5.24)	.78 (.70 to .88)	
12 weeks	263	4.95 (5.30)	263	6.27 (5.28)	.76 (.68 to .86)	
Average over time					.82 (.74 to .90)	<.0001
Group by time interaction						.008
Outcome	Sertraline		Placebo		Adjusted mean difference (95% CI)	P value
SF-12 Mental Health (n=538)	n	Mean (SD)	n	Mean (SD)		
2 weeks	275	37.32 (11.47)	291	35.37 (11.36)	1.67 (-.01 to 3.44)	
6 weeks	254	41.95 (12.35)	277	38.67 (11.91)	2.97 (1.24 to 4.70)	
12 weeks	263	42.70 (12.91)	264	39.71 (11.87)	2.91 (1.18 to 4.64)	
Average over time					2.49 (1.21 to 3.77)	<.0001
Group by time interaction						.22
SF-12 Physical Health (n=538)						
2 weeks	275	51.92 (9.18)	291	52.40 (6.64)	-.69 (-1.74 to .36)	
6 weeks	245	51.98 (8.39)	277	51.76 (9.90)	-.35 (-1.42 to .73)	
12 weeks	263	51.92 (8.53)	264	52.50 (9.99)	-.88 (-1.96 to .20)	
Average over time					-.64 (-1.47 to .18)	.13
Group by time interaction						.78

[a]These models use log transformed scores as the outcome. Adjusted proportional differences can be interpreted as the difference in scores between randomised groups expressed as a proportion (or percentage). Models are adjusted for baseline measure of each outcome (continuous), baseline CIS-R depression severity score and stratification variables (baseline total CIS-R score in three categories; duration of depressive episode in two categories; site) and variables not balanced at baseline (sex, ICD-10 depression diagnosis, marital status).

Table S8. Repeated measures analyses of binary secondary outcomes at 2, 6, 12 weeks, adjusted for variables not balanced at baseline.

Outcome	Sertraline		Placebo		Adjusted[a] odds ratio (95% CI)	P value
PHQ-9 remission[b] (n=547)	n	n (%)	n	n (%)		
2 weeks	277	145 (52)	292	136 (47)	1.42 (.80 to 2.51)	
6 weeks	267	169 (63)	285	164 (58)	1.36 (.76 to 2.43)	
12 weeks	262	190 (73)	263	170 (65)	1.83 (.99 to 3.42)	
Average over time					1.50 (.98 to 2.32)	.063
Group by time interaction						.49
BDI-II remission[b] (n=541)						
2 weeks	273	58 (21)	286	58 (20)	1.06 (.54 to 2.08)	
6 weeks	266	94 (35)	285	91 (32)	1.32 (.73 to 2.44)	
12 weeks	256	131 (51)	259	102 (39)	2.74 (1.49 to 5.05)	
Average over time					1.63 (1.04 to 2.56)	.34
Group by time interaction						.014
Feeling better (n=546)						
2 weeks	279	110 (39)	292	89 (30)	1.66 (1.07 to 2.57)	
6 weeks	267	157 (59)	285	132 (46)	1.92 (1.25 to 2.94)	
12 weeks	264	156 (59)	265	112 (42)	2.46 (1.57 to 3.82)	
Average over time					1.98 (1.47 to 2.66)	<.0001
Group by time interaction						.16

[a]All multi-level models adjusted for baseline measure of each outcome (continuous), baseline CIS-R depression severity score, stratification variables (baseline total CIS-R score in three categories; duration of depressive episode in two categories; site) and variables not balanced at baseline (sex, ICD-10 depression diagnosis, marital status).
[b]Remission defined as a score <=10.

Table S9. Physical health problems reported at baseline in the sample overall and according to treatment allocation

Physical health problem[a]	Overall (N=653)	Sertraline (n=324)	Placebo (n=329)
Any	272 (42%)	140 (43%)	132 (40%)
Diabetes	23 (8%)	11 (48%)	12 (52%)
Asthma or COPD[b]	52 (19%)	29 (56%)	23 (44%)
Arthritis	21 (8%)	12 (57%)	9 (43%)
Heart disease or heart problem	11 (4%)	4 (36%)	7 (64%)
Stroke	2 (0.7%)	1 (50%)	1 (50%)
Cancer	3 (1%)	2 (67%)	1 (33%)
Kidney disease	4 (1%)	1 (25%)	3 (75%)
None of the above	98 (36%)	54 (55%)	44 (45%)

[a]Assessed with the CIS-R

[b]Chronic Obstructive Pulmonary Disease (COPD) includes progressive lung diseases: emphysema, chronic bronchitis, and refractory (non-reversible) asthma.

Table S10. Baseline comparison of patients recruited through the GP record search and those recruited during GP consultation.

Characteristic	Recruitment method	
	Record search (n=466)	Consultation (n=187)
Site[a]		
Bristol	133 (29%)	132 (71%)
Liverpool	114 (25%)	2 (1%)
London	90 (19%)	52 (28%)
York	129 (28%)	1 (0.5%)
CIS-R total score[a]		
0 to 11	107 (23%)	22 (12%)
12 to 20	128 (27%)	45 (24%)
20 to 49	231 (50%)	120 (64%)
CIS-R depression duration[a]		
Less than 2 years	293 (63%)	146 (78%)
2 years or more	173 (37%)	41 (22%)
Age		
18-34	188 (40%)	78 (42%)
35-54	181 (39%)	78 (42%)
55-74	97 (21%)	31 (17%)
Sex		
Female	286 (61%)	98 (52%)
Male	180 (39%)	89 (48%)
ICD-10 CIS-R depression diagnosis[b]		
Yes	232 (50%)	123 (66%)
No	233 (50%)	64 (34%)
ICD-10 CIS-R anxiety diagnosis[b]		
Yes	192 (41%)	107 (57%)
No	273 (59%)	80 (43%)
Ethnic group[b]		
White	423 (91%)	156 (83%)
Ethnic minority	42 (9%)	31 (17%)
Marital status[b]		
Married or living as married	172 (37%)	83 (44%)
Single	213 (46%)	83 (44%)
Separated, divorced or widowed	80 (17%)	21 (11%)
Employment status[b]		
In paid employment	304 (65%)	129 (69%)
Not employed	161 (35%)	58 (31%)
Financial difficulty[b]		
Living comfortably or doing alright	267 (57%)	97 (52%)
Just about getting by	144 (31%)	60 (32%)
Finding it difficult or very difficult	54 (12%)	30 (16%)
Highest educational qualification[b]		
A Level or higher	331 (71%)	119 (64%)
GCSE, standard grade or other	114 (25%)	55 (29%)
No formal qualification	20 (4%)	13 (7%)
Depression in the past[b]		
Yes	403 (87%)	119 (64%)

No	62 (13%)	68 (20%)
Antidepressant in the past[b]		
Yes	319 (69%)	72 (39%)
No	146 (31%)	115 (62%)
PHQ-9 total score (range 0-27)	11.31 (5.76)	13.71 (5.55)
CIS-R total score (range 0-64)	19.91 (9.97)	24.56 (9.83)
CIS-R depression severity score (range 0-21)[c]	10.09 (4.91)	11.80 (4.66)
BDI-II total score (range 0-63)	24.01 (10.54)	23.87 (10.07)
GAD-7 score (range 0-21)	8.70 (5.08)	11.25 (5.32)
Social support score (range 1-24)	12.57 (3.81)	12.84 (3.85)
SF-12 mental health subscale (range 0-100)	35.56 (11.20)	29.76 (10.16)
SF-12 physical health subscale (range 0-100)	52.27 (9.45)	51.58 (10.32)
Number of life events in past 6 months	1.11 (1.12)	1.50 (1.32)
Number of physical symptoms in past 2 weeks	9.67 (5.31)	11.05 (5.59)
Frequency of physical symptoms (range 55-112)[d]	42.94 (10.19)	46.32 (12.35)

Data are n (%) or mean (SD).
CIS-R data are missing for one person.
Ranges for continuous scales are possible rather than actual ranges
PHQ-9=Patient Health Questionnaire, 9-item version
CIS-R=Clinical Interview Schedule Revised
BDI-II=Beck Depression Inventory, second edition
GAD-7=Generalised Anxiety Disorder Assessment, 7-item version
SF-12= Short-Form Health Survey
[a]The total CIS-R score assesses severity of symptoms of common mental disorder. Total CIS-R score in three categories was a stratification variable at randomisation: 0-11 (minimal symptoms); 12-19 (moderate to severe symptoms); 20+ (severe symptoms)
[b]A CIS-R diagnosis uses the criteria and threshold required to meet an ICD-10 clinical diagnosis of depression or anxiety. CIS-R data missing for one person
[c]The CIS-R depression severity score (range 0-21) assesses the severity of depressive symptoms
[d] How often during the past two weeks the patient experienced each symptom: 1 (not at all); 2 (several days); 3 (more than half the days); 4 (nearly every day)

Table S11. Post-hoc sub-group analyses of primary outcome (log transformed PHQ9 scores at 6 weeks). Means are for non-log transformed PHQ9 scores at 6 weeks.

Recruitment	Placebo		Sertraline		Comparison
	n	Mean (SD)	n	Mean (SD)	Adjusted proportional difference (95% CI)*
Record search	202	8.76 (5.93)	203	7.95 (5.69)	0.97 (0.85 to 1.10)
GP consultation	83	8.73 (5.69)	64	8.20 (5.51)	0.90 (0.73 to 1.11)
P value for interaction: 0.592					
Depression in the past	Placebo		Sertraline		Comparison
	n	Mean (SD)	n	Mean (SD)	Adjusted proportional difference (95% CI)*
Depressed in past	227	8.93 (6.07)	217	8.14 (5.61)	0.96 (0.84 to 1.08)
Not depressed in past	58	8.07 (4.89)	49	7.29 (5.74)	0.92 (0.72 to 1.19)
P value for interaction: 0.86					
Antidepressants in past	Placebo		Sertraline		Comparison
	n	Mean (SD)	n	Mean (SD)	Adjusted proportional difference (95% CI)*
Yes	175	8.98 (5.99)	160	7.97 (5.54)	0.96 (0.84 to 1,08)
No	110	8.38 (5.63)	106	8.00 (5.79)	0.99 (0.83 to 1.17)
P value for interaction: 0.21					

*Primary analysis model: Adjusted for baseline PHQ-9 score, continuous baseline CIS-R depression score and the stratification variables (baseline total CIS-R score in three categories; duration of depressive episode in two categories and site).

Table S12. Post-hoc sub-group analyses of PHQ-9 scores at 6 weeks, according to severity of depressive symptoms at baseline. Models use absolute (non-log transformed) PHQ-9 scores as outcome, to examine whether absolute differences differ according to baseline severity.

Baseline measure of depression severity	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
No ICD-10 depression diagnosis	125 (44)	6.76 (5.20)	129 (49)	6.20 (4.89)	-.44 (-1.53 to .65)
ICD-10 depression diagnosis	159 (56)	10.34 (5.89)	137 (52)	9.66 (5.78)	-.72 (-1.96 to .51)
P value for interaction: 0.792					
	Placebo		Sertraline		Mean difference (95% CI)*
CIS-R total score	n (%)	Mean (SD)	n (%)	Mean (SD)	
0 to 11	52 (18)	4.40 (4.12)	56 (21)	4.48 (3.90)	-.14 (-1.499 to 1.22)
12 to 20	79 (28)	7.70 (5.54)	72 (27)	6.81 (4.73)	-.66 (-2.27 to .95)
20 to 49	153 (54)	10.80 (5.59)	138 (52)	10.01 (5.81)	-.56 (-1.80 to .68)
P value for interaction: 0.733					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
CIS-R total score<28	203 (71)	7.89 (5.40)	198 (74)	6.91 (5.22)	-.85 (-1.76 to .05)
CIS-R total score>=28	81 (29)	10.96 (6.39)	68 (26)	11.10 (5.65)	.56 (-1.25 to 2.37)
P value for interaction: 0.153					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
CIS-R depression score<15	215 (76)	7.74 (5.24)	212 (80)	6.98 (5.20)	-.66 (-1.54 to .22)
CIS-R depression score>=15	69 (24)	11.96 (6.55)	54 (20)	11.91 (5.60)	-.05 (-2.14 to 2.04)
P value for interaction: 0.397					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
PHQ9 score <10	101 (36)	5.16 (4.11)	114 (43)	5.27 (4.28)	-.75 (-1.94 to .44)
PHQ9 score >=10	183 (64)	10.75 (5.74)	152 (57)	10.01 (5.68)	-.06 (-1.11 to .98)
P value for interaction: 0.440					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
PHQ9 score <20	253 (89)	8.44 (5.59)	242 (91)	7.50 (5.25)	-.77 (-1.59 to .06)
PHQ9 score >=20	31 (11)	11.42 (7.28)	24 (9)	12.83 (7.06)	1.86 (-1.69 to 5.41)
P value for interaction: 0.071					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
BDI-II score <10	16 (6)	2.44 (2.69)	20 (8)	5.00 (4.35)	.88 (-2.15 to 3.91)
BDI-II score >=10	268 (94)	9.14 (5.78)	246 (92)	8.22 (5.66)	-.67 (-1.53 to .20)
P value for interaction: 0.197					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
BDI-II score <29	199 (70)	7.50 (5.36)	187 (70)	6.44 (4.95)	-.77 (-1.69 to .15)
BDI-II score >=29	85 (30)	11.72 (5.95)	79 (30)	11.62 (5.50)	.19 (-1.51 to 1.88)
P value for interaction: 0.314					

*The placebo group was the reference category and negative mean differences indicate lower scores in the sertraline compared to placebo group. All models adjusted for baseline PHQ9 score, continuous baseline CIS-R depression score and stratification variables (baseline total CIS-R score in three categories; duration of depressive episode in two categories; centre).

All interaction terms and between treatment allocation and the severity variable used to create the sub-group.

Table S13. Post-hoc sub-group analyses of primary outcome (log transformed PHQ9 scores at 6 weeks), stratified according to the duration of the depressive episode at baseline.

Baseline measure of depression duration	Placebo		Sertraline		Adjusted proportional difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
Less than two years	187 (66)	8.15 (5.84)	179 (67)	7.72 (5.58)	0.97 (0.84 to 1.11)
Two years or more	97 (34)	9.95 (5.75)	87 (33)	8.53 (5.73)	0.90 (0.75 to 1.09)
P value for interaction: 0.690					
Baseline measure of depression duration	Placebo		Sertraline		Adjusted proportional difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
Less than 6 months	64 (12)	7.66 (5.15)	54 (12)	6.96 (5.78)	0.94 (0.74 to 1.20)
6 months to 1 year	41 (23)	9.80 (5.91)	48 (20)	9.46 (5.27)	1.00 (0.74 to 1.34)
Between 1 and 2 years	48 (14)	9.44 (6.40)	44 (18)	9.50 (5.41)	0.90 (0.70 to 1.16)
Between 2 and 5 years	48 (17)	10.08 (5.76)	50 (17)	8.58 (5.33)	0.88 (0.69 to 1.12)
More than 5 years	49 (17)	9.82 (5.80)	37 (14)	8.46 (6.31)	0.95 (0.70 to 1.28)
P value for interaction: 0.734					

*All models were adjusted for baseline PHQ9 score, continuous baseline CIS-R depression score and stratification variables (baseline total CIS-R score in three categories; duration of depressive episode in two categories; centre).

All interaction terms are between treatment allocation and the duration variable used to create the sub-group.

Table S14. Post-hoc sub-group analyses of GAD7 scores at 6 weeks, according to severity of depressive and anxiety symptoms at baseline. Models use absolute (non-log transformed) GAD-7 scores as the outcome, to examine whether mean differences vary according to baseline severity.

Baseline measure of depression severity	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
No ICD-10 depression diagnosis	124 (44)	5.25 (4.59)	127 (49)	4.36 (4.59)	-0.87 (-1.86 to .13)
ICD-10 depression diagnosis	159 (56)	8.33 (5.34)	136 (51)	6.60 (5.47)	-1.68 (-2.76 to -0.16)
P value for interaction: 0.27					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
No ICD-10 GAD diagnosis	151 (53)	5.42 (4.67)	150 (57)	4.09 (4.19)	-1.31 (-2.18 to -.42)
ICD-10 GAD diagnosis	133 (47)	8.71 (5.33)	113 (43)	7.42 (5.74)	-1.10 (-2.33 to .14)
P value for interaction: 0.89					
	Placebo		Sertraline		Mean difference (95% CI)*
CIS-R total score	n (%)	Mean (SD)	n (%)	Mean (SD)	
0 to 11	52 (18)	2.91 (3.22)	56 (21)	2.36 (2.94)	-.79 (-1.88 to .31)
12 to 20	79 (28)	5.37 (4.50)	69 (26)	4.24 (3.71)	-1.14 (-2.46 to .18)
20 to 49	152 (54)	9.28 (5.06)	138 (52)	7.43 (5.69)	-1.47 (-2.62 to -.32)
P value for interaction: 0.39					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
CIS-R total score<28	202 (71)	5.82 (4.65)	195 (74)	4.23 (4.26)	-1.52 (-2.30 to -.73)
CIS-R total score>=28	81 (29)	9.88 (5.54)	68 (26)	9.21 (5.79)	-.35 (-2.06 to 1.35)
P value for interaction: 0.23					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
CIS-R depression score<15	214 (53)	6.04 (4.67)	209 (47)	4.57 (4.51)	-1.41 (-2.19 to -.64)
CIS-R depression score>=15	69 (47)	9.88 (5.86)	54 (53)	9.15 (6.01)	-.63 (-2.55 to 1.39)
P value for interaction: 0.78					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
PHQ9 score <10	101(36)	4.22 (3.92)	112 (43)	3.26 (3.65)	-1.33 (-2.40 to -.26)
PHQ9 score >=10	182 (64)	8.51 (5.27)	151 (57)	7.19 (5.51)	-1.00 (-1.92 to -.07)
P value for interaction: 0.34					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
PHQ9 score <20	252 (89)	6.61 (5.00)	239 (91)	5.04 (4.77)	-1.50 (-2.24 to -.76)
PHQ9 score >=20	31 (11)	9.97 (6.22)	24 (9)	10.25 (6.21)	1.71 (-1.34 to 4.77)
P value for interaction: 0.047					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
GAD-7 score <8	118 (42)	4.07 (3.58)	118 (45)	3.32 (3.92)	-.82 (-1.73 to .09)
GAD-7 score >=8	166 (58)	9.02 (5.27)	146 (55)	7.36 (5.40)	-1.55 (-2.65 to -.45)
P value for interaction: 0.31					
	Placebo		Sertraline		Mean difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
GAD-7 score <14	221 (78)	5.67 (4.36)	205 (78)	4.40 (4.26)	-1.09 (-1.84 to -.35)
GAD-7 score >=14	63 (22)	11.51 (5.58)	58 (22)	9.48 (6.15)	-1.50 (-3.55 to .56)
P value for interaction: 0.42					

*The placebo group was the reference category and negative mean differences indicate lower scores in the sertraline compared to placebo group. All models adjusted for baseline PHQ9 score, continuous baseline CIS-R depression score and stratification variables (baseline total CIS-R score in three categories; duration of depressive episode in two categories; centre).

All interaction terms and between treatment allocation and the severity variable used to create the sub-group.

Table S15. Post-hoc sub-group analyses of GAD-7 scores at 6 weeks (log transformed), according to the duration of anxiety symptoms at baseline.

Baseline measure of anxiety duration	Placebo		Sertraline		Adjusted proportional difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
Less than two years	194 (68)	6.12 (4.93)	183 (69)	4.73 (4.59)	.79 (.69 to .93)
Two years or more	90 (32)	8.77 (5.33)	81 (31)	7.40 (5.98)	.75 (.60 to .95)
P value for interaction: 0.74					
Baseline measure of anxiety duration	Placebo		Sertraline		Adjusted proportional difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
Less than 6 months	69 (24)	6.80 (4.86)	50 (19)	5.18 (4.86)	.79 (.59 to 1.16)
6 months to 1 year	25 (9)	6.52 (4.46)	31 (12)	5.87 (4.34)	1.04 (.72 to 1.00)
Between 1 and 2 years	33 (12)	6.79 (5.28)	44 (17)	5.34 (4.96)	.77 (.54 to 1.15)
Between 2 and 5 years	54 (19)	9.06 (5.30)	41 (16)	7.46 (5.75)	.72 (.52 to 1.51)
More than 5 years	36 (13)	8.36 (5.73)	40 (15)	7.33 (6.27)	.82 (.59 to 1.05)
P value for interaction: 0.95					
Baseline measure of depression duration	Placebo		Sertraline		Adjusted proportional difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
Less than two years	194 (68)	6.12 (4.93)	183 (69)	4.73 (4.59)	.77 (.66 to .91)
Two years or more	90 (32)	8.77 (5.33)	81 (31)	7.40 (5.98)	.78 (.63 to .97)
P value for interaction: 0.42					
Baseline measure of depression duration	Placebo		Sertraline		Adjusted proportional difference (95% CI)*
	n (%)	Mean (SD)	n (%)	Mean (SD)	
Less than 6 months	69 (24)	6.80 (4.86)	50 (19)	5.18 (4.86)	.69 (.51 to .91)
6 months to 1 year	25 (9)	6.52 (4.46)	31 (12)	5.87 (4.34)	1.08 (.78 to 1.51)
Between 1 and 2 years	33 (12)	6.79 (5.28)	44 (17)	5.34 (4.96)	.58 (.44 to .79)
Between 2 and 5 years	54 (19)	9.06 (5.30)	41 (16)	7.46 (5.75)	.76 (.58 to .98)
More than 5 years	36 (13)	8.36 (5.73)	40 (15)	7.33 (6.27)	.72 (.50 to 1.04)
P value for interaction: 0.96					

*All models were adjusted for baseline PHQ9 score, continuous baseline CIS-R depression score and stratification variables (baseline total CIS-R score in three categories; duration of depressive episode in two categories; centre).

All interaction terms are between treatment allocation and the duration variable used to create the sub-group.

Table S16. Repeated measures analyses of binary response outcome at 2, 6, 12 weeks.

Outcome	Sertraline		Placebo		Adjusted[a] odds ratio (95% CI)	P value
PHQ-9 response[b] (n=547)	n	n (%)	n	n (%)		
2 weeks	277	145 (52)	292	136 (47)	1.12 (.60 to 2.07)	
6 weeks	267	169 (63)	285	164 (58)	1.47 (.86 to 2.51)	
12 weeks	262	190 (73)	263	170 (65)	2.16 (1.26 to 3.72)	
Average over time					1.58 (1.05 to 2.37)	.028
Group by time interaction						.062
BDI-II response[b] (n=541)						
2 weeks	273	58 (21)	286	58 (20)	0.81 (0.41 to 1.58)	
6 weeks	266	94 (35)	285	91 (32)	1.33 (0.72 to 2.45)	
12 weeks	256	131 (51)	259	102 (39)	2.35 (1.26 to 4.39)	
Average over time					1.42 (0.88 to 2.23)	.15
Group by time interaction						.0005
GAD-7 response[b] (n=546)						
2 weeks	278	83 (30)	291	72 (25)	1.44 (0.78 to 2.69)	
6 weeks	265	135 (51)	284	89 (31)	3.97 (2.14 to 7.32)	
12 weeks	264	152 (58)	263	112 (42)	2.68 (1.46 to 4.93)	
Average over time					2.51 (1.58 to 4.00)	<.0001
Group by time interaction						.094

[a]All multi-level models adjusted for baseline measure of each outcome (continuous), baseline CIS-R depression severity score, stratification variables (baseline total CIS-R score in three categories; duration of depressive episode in two categories; site) and variables not balanced at baseline (sex, ICD-10 depression diagnosis, marital status).
[b]Defined as a 50% reduction in symptoms or greater, from baseline.

Table S17. Repeated measures analyses of PHQ-9 remission, with remission defined as scoring 0-4.

Outcome	Sertraline		Placebo		Adjusted[a] odds ratio (95% CI)	P value
PHQ-9 remission	n	n (%)	n	n (%)		
2 weeks	277	56 (20)	292	44 (15)	1.63 (.78 to 3.44)	
6 weeks	267	92 (34)	285	80 (28)	1.50 (.78 to 2.89)	
12 weeks	262	117 (45)	263	92 (35)	2.16 (1.13 to 4.13)	
Average over time					1.63 (1.07 to 2.47)	.022
Group by time interaction						.44

[a]Adjusted for baseline measure of each outcome (continuous), baseline CIS-R depression severity score, and stratification variables (baseline total CIS-R score in three categories; duration of depressive episode in two categories; site.

Changes to protocol version 3 (protocol version 3 was dated 30/05/2014 and was the last version of the protocol before the study started 01/01/2015):

1. To capture patient-rated change, we added the measure of self-reported improvement as a secondary outcome in protocol version 4 dated 05/03/2015 and approval of this change was received from the ethics committee (NRES) 15/05/2015.
2. It was apparent towards the later stages of designing the RCT and in formulating the detailed analysis plan (uploaded before any analyses were performed to: http://discovery.ucl.ac.uk/10041458/ and approved by the Trial Steering Committee), that we would have insufficient statistical power to estimate plausible interaction effects. Our power calculation and primary analysis (as stated in the analysis plan: http://discovery.ucl.ac.uk/10041458/) were therefore based on a primary aim to examine the clinical effectiveness of sertraline versus placebo. This change was made in protocol version 4, dated 05/03/2015.
3. In line with the change above (point 2), we changed the primary analysis to a linear regression of log transformed PHQ-9 scores at 6 weeks. Interactions between severity and duration at baseline and treatment response were planned as exploratory. This is documented in the detailed analysis plan (uploaded before any analyses were performed to: http://discovery.ucl.ac.uk/10041458/ and approved by the Trial Steering Committee)
4. Due to a poor response rate from GP mail-outs, the recruitment process was modified to include a further telephone call to non-responders. This change was submitted in protocol version 5 dated 16/11/2015 and approval was received from NRES 01/12/2015.
5. Due to a release of SmPC v7 (the information about the drug that the manufacturing company releases that includes all the safety and adverse effects) that mentions an increased QT interval associated with Sertraline and after discussions with our Sponsor, we amended the protocol. This change was submitted in protocol version 6 dated 11/02/2016 and approval was received from NRES 04/04/2016.
6. We updated the GP referral sheet and GP eligibility confirmation to reflect the change referenced in point 4 above. This change was submitted in protocol version 6 dated 11/02/2016 and approval was received from NRES 04/04/2016.
7. The sertraline patient information leaflet was replaced with the updated version, issued by Bristol Labs Ltd. This change was submitted in protocol version 6 dated 11/02/2016 and approval was received from NRES 04/04/2016.
8. Minor amendments were made to include nurse prescribers, change the procedure and contact details for reporting Pharmacovigilance and add 'GP practices' to the insurance section. This change was submitted in protocol version 6.1 dated 22/04/2016 and approval was received from the trial sponsor, 13/07/2016. This was a minor change which required approval from the sponsor rather than the ethics committee.
9. Attrition was higher than expected and a minor amendment was submitted to recruit more participants than originally intended. This change to the protocol was submitted in protocol version 6.1 dated 22/04/2016 and approval was received from the trial sponsor, dated 13/07/2016.
10. Results from analyses of the EQ5D and emotional processing secondary outcomes will be reported in a separate paper. The EQ5D will form part of the economics analysis and the emotional processing tasks will be analysed using complex computational modelling.
11. Qualitative analyses of the PANDA RCT that were planned to aid recruitment were not conducted because recruitment rates were higher than expected.