

Research paper

An exploratory parallel-group randomised controlled trial of antenatal Guided Self-Help (plus usual care) versus usual care alone for pregnant women with depression: DAWN trial



K. Trevillion^{a,1,*}, E.G. Ryan^{a,b,1}, A. Pickles^a, M. Heslin^a, S. Byford^a, S. Nath^a, D. Bick^c, J. Milgrom^d, R. Mycroft^a, J. Domoney^a, C. Pariante^a, M.S. Hunter^{a,2}, L.M. Howard^{a,2}

^a Institute of Psychiatry, Psychology & Neuroscience at King's College London, United Kingdom

^b Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, Coventry CV4 7AL, United Kingdom

^c Department of Women and Children's Health, School of Life Course Sciences, Faculty of Life Sciences and Medicine, King's College London, United Kingdom

^d University of Melbourne and Parent-Infant Research Institute, Melbourne, Australia

ARTICLE INFO

Keywords:

Randomised Controlled trial

Guided Self-Help

Clinical Trial

Depression

ABSTRACT

Background: Depression is a common antenatal mental disorder associated with significant maternal morbidity and adverse fetal outcomes. However, there is a lack of research on the effectiveness or cost-effectiveness of psychological interventions for antenatal depression.

Methods: A parallel-group, exploratory randomised controlled trial across five hospitals. The trial compared Guided Self-Help, modified for pregnancy, plus usual care with usual care alone for pregnant women meeting DSM-IV criteria for mild-moderate depression. The trial objectives were to establish recruitment/follow-up rates, compliance and acceptability, and to provide preliminary evidence of intervention efficacy and cost-effectiveness. The primary outcome of depressive symptoms was assessed by blinded researchers using the Edinburgh Postnatal Depression Scale at 14-weeks post-randomisation.

Results: 620 women were screened, 114 women were eligible and 53 (46.5%) were randomised. 26 women received Guided Self-Help – 18 (69%) attending ≥ 4 sessions - and 27 usual care; $n = 3$ women were lost to follow-up (follow-up rate for primary outcome 92%). Women receiving Guided Self-Help reported fewer depressive symptoms at follow-up than women receiving usual care (adjusted effect size -0.64 (95%CI: -1.30 , 0.06) $p = 0.07$). There were no trial-related adverse events. The cost-effectiveness acceptability curve showed the probability of Guided Self-Help being cost-effective compared with usual care ranged from 10 to 50% with a willingness-to-pay range from £0 to £50,000.

Conclusions and Limitations: Despite intense efforts we did not meet our anticipated recruitment target. However, high levels of acceptability, a lack of adverse events and a trend towards improvements in symptoms of depression post-treatment indicates this intervention is suitable for talking therapy services.

1. Introduction

Antenatal mental disorders are common (Howard et al., 2018), with depression affecting 11% of women in early pregnancy (Howard et al., 2018), and up to 18% of women across the pregnancy period (Howard et al., 2014b). Antenatal depression is associated with an increased risk of preterm delivery (Stein et al., 2014), postnatal maternal psychopathology (Bick and Howard, 2010; Howard et al., 2014b;

Milgrom et al., 2008a,b), and subsequent behavioural/emotional problems in children (Stein et al., 2014). Timely antenatal mental health interventions are therefore necessary to alleviate these adverse effects.

International guidelines advocate a lower threshold for psychological therapies and a higher threshold for psychotropic medication for mental health problems during pregnancy, due to potential risks to the fetus (Howard et al., 2014a). Yet, there is a paucity of research investigating the effectiveness and cost-effectiveness of antenatal

* Corresponding author at: King's College London, Section of Women's Mental Health, PO31 Institute of Psychiatry, Psychology & Neuroscience, De Crespigny Park, London SE5 8AF, United Kingdom.

E-mail address: kylee.trevillion@kcl.ac.uk (K. Trevillion).

¹ Joint first authors.

² Joint last authors.

Abbreviations	
RCT	Randomised Controlled Trial
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
CBT	Cognitive Behavioural Therapy
PWP(s)	Psychological Wellbeing Practitioner(s)
GSH	Guided Self Help
EPDS	Edinburgh Postnatal Depression Scale
PHQ-9	Patient Health Questionnaire-9
GAD-7	Generalized Anxiety Disorder scale-7
PBQ	Postpartum Bonding Questionnaire
SPS	Social Provisions Scale
AUDIT-C	the Alcohol Use Disorders Identification Test-C
CAS	Composite Abuse Scale
SF-36	36-item Short Form Health Survey
EQ-5D-5L	EuroQol measure of health-related quality of life
QALYs	Quality-Adjusted Life Years
CI	Confidence Interval
IQR	Inter-Quartile Range
DNAs	Did Not Attend

interventions for one of the most common disorders - mild to moderate depression (Sockol et al., 2011). We are aware of only one individualised randomised controlled trial (RCT) for pregnant women with diagnosed mild to moderate depression, a pilot RCT of an eight-session antenatal Cognitive Behavioural Therapy (CBT), delivered by psychologists, to 54 Australian women; substantial reductions in depression and anxiety were observed during pregnancy, compared to the control group, with effects maintained at nine months postpartum (Milgrom et al., 2014).

The National Institute for Health and Care Excellence (NICE), an organisation providing evidence-based guidance to health services in the UK, recommends a stepped-care approach to treating perinatal mental disorders; starting with a low-cost intervention for mild to moderate depression - Guided Self-Help (GSH) - delivered by National Health Service (NHS) Psychological Wellbeing Practitioners (PWPs). GSH is a standardised psychological treatment that individuals can work through semi-independently using step-by-step or modular instructions. Recent systematic reviews and meta-analyses of RCTs of GSH for adults with depression report evidence of effectiveness post-treatment (Coull and Morris, 2011; Cuijpers et al., 2010). In the perinatal context, Self-Help interventions have been successfully delivered within Australian antenatal settings (Milgrom et al., 2011). What remains unclear is whether such interventions could be successfully delivered within psychological therapy services as well as demonstrating cost-effectiveness in reducing antenatal depression. We therefore developed an exploratory RCT to test a new form of GSH for antenatal depression in women attending NHS maternity services in the UK, which could be delivered by PWPs in primary care psychology services.

There were four main aims of the trial:

- To establish that the trial procedures worked (and to fine tune where necessary) so that a Phase III trial could follow;
- To evaluate if antenatal GSH was beneficial in improving depressive symptoms for women with antenatal depression;
- To evaluate if antenatal GSH had the added benefit of improving other outcomes, including post-treatment and post-delivery psychological symptoms, post-delivery bonding and quality of life;
- To explore if antenatal GSH was cost-effective compared to usual care.

We hypothesised that women with mild or moderate antenatal depression treated with GSH would have significantly lower depressive symptom scores on the Edinburgh Postnatal Depression Scale (EPDS) at 14 weeks post-randomisation compared to women with mild or moderate antenatal depression receiving usual care.

2. Methods

The *Depression: a trial of antenatal Guided Self-Help for women (DAWN)* trial was registered on 08/08/14 (prior to recruitment) on ISRCTN.com (<http://www.isrctn.com/ISRCTN83768230>). The protocol has been published (Trevillion et al., 2016). CONSORT reporting

guidelines for pilot and feasibility trials (Eldridge et al., 2016), the TiDIER checklist (Hoffmann et al., 2014) and Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist (Husereau et al., 2013) have been completed (see Supplementary Tables 1–3).

This was a multi-centre Phase II exploratory randomised controlled trial with two parallel groups and a primary endpoint of EPDS score at 14 weeks post-randomisation.

2.1. Oversight

The trial and protocol were approved by the Camberwell St Giles Research Ethics Committee, London (reference: 14.LO.0597). A Trial Management Group (set up to monitor the day-to-day running of the trial) met approximately every eight weeks during the trial. Independent oversight was provided by a Programme Steering Committee (PSC). An independent Data Monitoring and Ethics Committee (IDMEC) reviewed the progress of the trial, the accruing data and monitored the safety of participants, and an independent Trial Steering Committee (TSC) provided advice and oversight on the trial conduct/progress. The PSC and IDMEC met annually and the TSC biannually during the trial. Analyses were conducted according to a pre-specified Statistical and Health Economics Analysis Plan which were approved by the TSC and IDMEC (v2, approved 29/06/2017). The trial PWPs received weekly supervision meetings from one of the authors (RM).

2.2. Randomisation and blinding

A central randomisation system, provided by the UK Clinical Research Collaboration registered at King's College London Clinical Trials Unit, allocated participants to either Guided Self-Help plus usual care (GSH) or to usual care alone (i.e. treatment as usual). Randomisation, stratified by type of depression (3 levels: mild depression, moderate depression, or mixed anxiety and depression), was applied using computer-generated block-randomisation of varying sizes, with a 1:1 allocation. Block sizes were not disclosed, to ensure concealment. Trial researchers responsible for collection of outcome data and trial statisticians were blind to treatment allocation. Effectiveness of allocation concealment (among the researchers responsible for collection of outcome data) was assessed; researchers reported which arm they believed participants to be in.

2.3. Study setting

The trial was conducted in five large National Health Service (NHS) maternity units within South East London; these services serve ethnically and socially diverse populations. The NHS is a publicly-funded healthcare system in the UK, free at the point of use, to every legal resident in the UK.

2.4. Eligibility criteria

Adult women, aged ≥ 16 years, who were pregnant (not exceeding 26 weeks gestation) and who met criteria for DSM-IV depression on the Structured Clinical Interview (i.e. mild or moderate major depressive disorder, or mixed anxiety and depressive disorder) were eligible. Women were excluded if they were: receiving CBT or any other psychological therapy; taking antidepressants; suffering from psychosis, a current eating disorder, borderline personality disorder or a current post-traumatic stress disorder; reporting current suicidality; receiving care from secondary mental health services; unable to complete questionnaires or follow the trial workbook in English; unable to provide informed consent.

2.5. Recruitment

Recruitment occurred between 5th January 2015 and 30th June 2016; follow-up data collection ended on 10th April 2017. Women were recruited to the trial in one of three ways: (1) via their participation in a related study on well-being in pregnancy (Howard et al., 2018) (REC reference: 14/LO/0075); (2) via midwives who considered a woman suitable for the trial; (3) through self-referral, via advertised study posters.

Baseline trial researchers recruited and consented participants. They provided women with a minimum of 24 hours to consider their participation and obtained written informed consent from everyone. These researchers communicated the group allocation to participants and, for women allocated to GSH, arranged an appointment with the PWP to start the intervention; a choice of locations were offered for appointments (e.g. clinical trial facility/clinical room at study sites or women's homes) (see Trevillion et al., 2016 full details on recruitment methods).

2.6. Intervention

2.6.1. Guided Self-Help (GSH) interventions for depression

GSH is delivered as part of the Improving Access to Psychological Therapies (IAPT) programme (Clark, 2011) within NHS-commissioned services in England. IAPT is delivered within every local health area in England and it provides psychological treatments for people with depression and anxiety. IAPT treatments conform to stepped-care clinical guidelines produced by NICE (Clark et al., 2018). Recent data from the IAPT programme shows that over 537,000 people are treated each year, with the majority of treatments offered comprising CBT (NHS Digital, 2016). PWPs deliver evidence-based low intensity CBT interventions within the IAPT programme, particularly treatments based on GSH manuals.

2.6.2. Modified Guided Self-Help, with usual care, (GSH) for antenatal depression

A GSH workbook was developed specifically for this trial (available from corresponding author on request). The workbook is divided into six chapters, including psychoeducation on antenatal depression; managing relationships; planning for parenthood; health and lifestyle factors. Homework tasks are included at the end of each chapter, and at various points throughout the workbook (see Trevillion et al., 2016 for further details on the workbook).

Four IAPT PWPs were trained to deliver the modified GSH intervention; PWPs are experienced in delivering treatments based on GSH manuals, and in supporting participants' use of GSH manuals. The trial GSH intervention comprised an initial session with the PWP followed by up to eight 30-min sessions with the PWP. In addition, the PWPs conducted a check-in session with women at six to eight weeks post-delivery (i.e. before the post-delivery research outcomes were collected); this post-delivery session did not form part of the therapeutic programme. The initial session was delivered face-to-face; subsequent sessions were either delivered face-to-face or by telephone, depending

on women's preferences. Based on our pilot work, a minimum dose of four sessions was determined necessary to impact on depressive symptoms.

2.6.3. Usual care

We recorded use of all health and social care services over the follow-up period, as part of the economic evaluation.

2.7. Trial measures

Participant data were collected at baseline, 14 weeks post-randomisation and 3 months post-delivery. Key measures included: Edinburgh Postnatal Depression Scale (EPDS) (Cox and Holden, 2003); Patient Health Questionnaire-9 (PHQ-9) (Kroenke et al., 2001); Generalized Anxiety Disorder scale-7 (GAD-7) (Spitzer et al., 2006) and Postpartum Bonding Questionnaire (PBQ) (Brockington et al., 2001). Other measures included the Social Provisions Scale (SPS) (Cutrona and Russell, 1987), Alcohol Use Disorders Identification Test-C (AUDIT-C) (Babor et al., 2008), questions on smoking behaviours (developed by the study team), the Composite Abuse Scale - short version (CAS), which measures experiences of domestic violence (Hegarty et al., 2005), and questions relating to timing of delivery (see Trevillion et al., 2016) for full details on all trial measures).

A fidelity rating scale was developed to rate PWPs' adherence to the GSH components. At completion of the trial a randomly sampled selection of 20 recordings were evaluated for fidelity by RM and an independent assessor (a clinical psychologist not involved in the trial); recordings were stratified by therapist and time (first versus second half of the trial). Randomly generated numbers were used to select sessions from each group, which represented initial, mid and last sessions. In addition, the interpersonal effectiveness item from the Cognitive Therapy Scale Revised (Blackburn et al., 2001) was used as an indicator of positive therapeutic alliance.

Process evaluation data were collected by PWPs on the mode of delivery of GSH sessions (i.e. face-to-face or telephone) and levels of compliance (i.e. number and length of sessions offered, number of sessions attended/not attended or cancelled). PWPs also collected the same standardised measures, at every session, as they did within their IAPT service: PHQ-9 (Kroenke et al., 2001), GAD-7 (Spitzer et al., 2006), the Work & Social Adjustment Scale and Phobia Scales (Improving Access to Psychological Therapies, 2011) (see Supplementary Table 4 and Supplementary Fig. 1 for further details).

2.8. Outcomes

2.8.1. Primary outcome

Depressive symptoms on the Edinburgh Postnatal Depression Scale (EPDS) at 14 weeks post-randomisation.

2.8.2. Key secondary outcomes

At 14 weeks post-randomisation

- 1 Proportion meeting Participant Health Questionnaire- 9 (PHQ-9) criteria for depression (i.e. score of ≥ 10).
- 2 Proportion meeting Generalised Anxiety Disorder-7 (GAD-7) criteria for anxiety (i.e. score of ≥ 8).

At three months post-delivery

- 1 Edinburgh Postnatal Depression Scale scores.
- 2 Parenting stress, as measured by the Postpartum Bonding Questionnaire (PBQ).

3. Analysis

Although this was an exploratory RCT, we carried out a sample size

calculation using Stata's *sampsi* procedure to assess what preliminary evidence we could generate on efficacy. Using a two-tailed significance of 5% with a correlation of 0.4 between baseline and outcome EPDS score, a two-arm parallel-group design with 52 women in each arm gives 79% power to detect a difference of 0.5 SD using ANCOVA. Recruiting 110 women, we would have had 66% power to detect a difference in the caseness rate of recovery of 35% in the treatment as usual arm versus 60% in the GSH arm.

3.1. Quantitative data analysis

Recruitment rate, loss to follow-up and withdrawals are described using a CONSORT diagram. Patient demographics, baseline characteristics, primary and secondary outcomes are presented using appropriate descriptive summaries: mean and standard deviation for normally distributed measures; medians and quartiles for skewed distributions; numbers and percentages for discrete outcomes. The primary analysis was performed based on the intention to treat (ITT) principle. EPDS total scores at 14 weeks post-randomisation and 3 months post-delivery were jointly modelled using seemingly unrelated regressions (implemented using the Stata *sem* procedure) to account for any correlation between the two measures. This involved performing an ANCOVA on the EPDS score at 14 weeks post-randomisation with treatment arm and depression severity as factors and the baseline EPDS score as a covariate. Effect size was calculated using the standard deviation of the baseline EPDS over both groups; 95% confidence intervals (CIs) were calculated using bootstrapping (1000 samples used). A similar model was used for the EPDS at 3 months post-delivery. The piecewise nature of the model minimises the contamination of the 3-month post-delivery estimates by the 14-week post-randomisation scores. Maximum likelihood with missing values (MLMV) was used for estimation.

Due to the small sample size [see results section for details], planned supplementary analyses to estimate complier-average causal effects for a binary measure of compliance to treatment using the instrumental variable method were not undertaken. Exact logistic regressions were used to model the difference in proportions meeting the PHQ-9 criteria for depression (i.e., a score ≥ 10) and the GAD-7 criteria for anxiety (i.e., a score ≥ 8), as well as alcohol intake ("yes drinking", "not drinking") on the AUDIT-C questionnaire. These were analysed at 14 weeks post-randomisation and 3 months post-delivery, with depression severity and treatment arm included as factors, and total baseline scores included as a covariate. Due to the small numbers of participants reporting smoking behaviours, domestic violence, or a premature birth (< 37 weeks) planned logistic regressions were not undertaken for these outcomes. Planned analyses of infant height and weight were not undertaken as these data were not available. The social provisions scale (SPS) total score was analysed using an ANCOVA model that included the baseline score. Parenting stress, as measured by the PBQ, was assessed using a two-way ANOVA of the total score. SPS and PBQ analyses included treatment arm and depression severity as factors. Due to small numbers, individuals with mixed anxiety and depression were combined with the mild depression individuals in all analyses that contained the depression severity stratifier (i.e. 2 levels: mild/mixed anxiety depression, moderate depression).

For baseline measures used as a covariate in the main modelling, participant mean item score was used to impute missing items (for up to 30% missing items) in forming a total score (pro-rating). This was deemed to be satisfactory as the amount of missing data was small (approximately 1-2% of participants). No imputation was performed for participants who had $> 30\%$ missing data in a questionnaire. Missing data in post-randomisation outcome variables were accounted for under the Missing-at-Random assumption of maximum likelihood. All statistical analyses were performed in Stata (v14).

3.2. Economic evaluation

A cost-effectiveness analysis was conducted at the 3-month post-delivery follow-up point taking an NHS/Personal Social Services perspective preferred by NICE (National Institute for Health and Care Excellence, 2008). Quality-Adjusted Life Years (QALYs) were calculated using the SF-6D measure of health-related quality of life, derived from the 36 item Short Form Health Survey (SF-36) (Ware, 2000), which was completed by participants at baseline, and at the 14-week post-randomisation and 3-month post-delivery follow-up points.

Data on the number of contacts participants in the GSH arm received from PWPs was recorded using a proforma, completed by the PWPs, which included the number and duration of sessions for each participant. To account for the ratio of direct face-to-face to indirect non-face-to-face time (preparation, administration, supervision etc.) a proforma was completed by the PWPs to record time spent on different activities in a typical week; this was used to estimate additional time associated with each face-to-face contact. Data on the use of all other health and social care services were collected using the Adult Service Use Schedule (AD-SUS) adapted for this study to cover all hospital and community-based health and social care services for both participants and their index baby (post-birth) (Howard, 2011; Howard et al., 2010). The AD-SUS was completed in interview at both follow-up points and covered the period from baseline to the 14-week post-randomisation interview and from the 14-week post-randomisation interview to the 3-month post-delivery interview, respectively. Total costs were calculated by applying unit costs to resource use at the individual level. Nationally applicable unit costs were applied to all services (see online supplement for additional details). All costs are reported in pounds sterling at 2015/2016 prices. Discounting was not relevant as the follow-up did not exceed 12 months.

Differences in mean costs and outcomes were obtained by non-parametric bootstrap regressions (10,000 repetitions, bias-corrected) to account for non-normally distributed data commonly found in economic data (Thompson and Barber, 2000). To provide more relevant treatment-effect estimates (Assmann et al., 2000), regressions to calculate mean differences in costs included baseline variables which could influence cost over the follow-up, including depression stratifier (defined as mild depression or mixed depression and anxiety versus moderate depression) and the baseline variable of interest, where available. The primary economic analysis excluded patients with missing cost or outcome data.

Cost-effectiveness was explored with the net benefit approach. We calculated area under the curve values for QALYs with linear interpolation between assessments (Manca et al., 2005). We used cost-effectiveness acceptability curves to explore uncertainty around costs and cost-effectiveness due to sampling variation and the maximum cost-effectiveness ratio that a decision maker would deem acceptable (Fenwick and Byford, 2005). The curves were created from bootstrapped costs and effects to calculate the probability of each treatment being the optimum choice, subject to a range of possible maximum values that a decision maker might be willing to pay for an increase in QALYs.

A secondary analysis was performed substituting the SF-6D with the 5-dimension, 5-level version of the EuroQol measure of health-related quality of life (EQ-5D-5L) (Herdman et al., 2011). Sensitivity analyses included: a narrow perspective including mental health services only in order to control for the fact that many services used in the postpartum period may be unrelated to mental health and the intervention under investigation; replacing the proportion of direct to indirect time spent by PWPs with published estimates (1:1; (Curtis and Burns, 2015)); and removing outliers and influential observations (see Supplementary Information 3 for full details on the economic evaluation methods).

4. Results

4.1. Aim one: to establish that the trial procedures worked

See Table 1 for the demographic details of the $n = 53$ participants and Table 2 for the baseline clinical characteristics of participants. Overall the two groups had similar characteristics, except for annual income; the usual care arm had considerably higher annual incomes than the intervention arm (annual income $\geq £46,000$ of 70% versus 24%, respectively).

4.2. Recruitment and retention

We recruited $n = 53$ women: 26 participants were randomly allocated to GSH and 27 randomly allocated to usual care. Three participants were lost to follow-up at 14 weeks post-randomisation (two from the GSH arm and one from the usual care only arm); four participants were lost to follow-up at 3 months post-delivery (two from each arm). No participants withdrew from the trial (See Fig. 1 for full details).

Despite intensive efforts, we did not reach our recruitment target of $n = 110$ participants. We implemented several different strategies to enhance recruitment, including: trial researchers working evenings and weekends to fit in with women's work/childcare commitments; trial researchers and PWPs conducting home visits; the trial team attending expectant parent workshops and antenatal staff meetings to advertise the study; extending recruitment by three months. The trial team also sought advice and guidance from their lived experience advisory group on ways to enhance recruitment. At the beginning of the trial, findings of local audits and a recent study (Burns et al., 2013) indicated that very few pregnant women were referred for psychological therapy in routine practice. However, during the trial substantial national policies and developments for perinatal IAPT services occurred. Local IAPT services subsequently became more proactive in facilitating access by pregnant women and midwives made more referrals to IAPT; these policy and practice changes are likely to have reduced our ability to recruit women. We saw recruitment tail off around the end of December 2015 and, after a few more months of intense efforts, we took the decision to close recruitment on 30th June 2016.

4.3. Levels of compliance with the intervention

18 participants (69%) attended at least the minimum number of sessions (≥ 4 sessions) and eight (31%) did not. All eight 'non-completers' remained in the study and their reasons for withdrawal from treatment included: inconvenience of the intervention alongside existing commitments ($n = 3$); intervention no longer needed as symptoms resolved ($n = 2$); disengagement ($n = 2$). The 18 participants who attended ≥ 4 sessions also attended the additional post-delivery session; only one of the non-completers (who did not withdraw from treatment) attended this session.

The overall median number of sessions attended was 6.5 (IQR 3–8; range 0–8). For completers the median number of sessions attended was 7 (IQR 6–8; range 4–8) for non-completers the median number of sessions attended was 1.5 (IQR 0–3; range 0–3). Three participants did not attend any GSH sessions.

There were 231 scheduled appointments, of which 157 (68%) were attended, 65 (28%) were cancelled, and 9 (4%) were not attended (DNAs). Of the attended sessions, 115 (73%) were face-to-face and 42 (27%) were by telephone. In 64% of sessions relevant chapter reading of the workbook had been completed and in 41% of sessions all homework had been completed.

4.4. Treatment fidelity and therapeutic alliance

Fidelity ratings were high; a median score of 100% (IQR: 100%–100%; range: 67%–100%) was achieved across 16 randomly

selected sessions. PWPs demonstrated competence with regards to interpersonal effectiveness, achieving a median score of 3 (IQR: 3–4; mean 3.4).

4.5. Trial adverse events

Nine adverse events were reported, five in the usual care arm and four in the GSH arm. None of these adverse events were related to

Table 1
Demographic details of $n = 53$ DAWN trial participants.

	Treatment arm: Guided Self-Help group	Comparison arm: Usual care group	Overall
Age group ($n = 53$)			
<25	3 (11.54%)	2 (7.41%)	5 (9.43%)
25–29	5 (19.23%)	3 (11.11%)	8 (15.09%)
30–39	18 (69.23%)	18 (66.67%)	36 (67.92%)
40+	0 (0%)	4 (14.81%)	4 (7.55%)
Ethnicity ($n = 53$)			
White	18 (69.23%)	17 (62.96%)	35 (66.04%)
Black	7 (26.92%)	7 (25.93%)	14 (26.42%)
Asian	0 (0%)	1 (3.70%)	1 (1.89%)
Mixed/Other	1 (3.85%)	2 (7.41%)	3 (5.66%)
Gestational age at baseline visit ($n = 49$)			
Mean: 10 weeks ($sd = 1.76$)	Mean: 11 weeks ($sd = 2.19$)	Mean: 10.6 weeks ($sd = 2.06$)	Mean: 10.6 weeks ($sd = 2.06$)
Min: 7; Max: 15	Min: 8; Max: 17	Min: 7; Max: 17	Min: 7; Max: 17
Employment status ($n = 53$)			
Working	23 (88.46%)	21 (77.78%)	44 (83.02%)
Student	1 (3.85%)	0 (0%)	1 (1.89%)
Unemployed	2 (7.69%)	2 (7.41%)	4 (7.55%)
Homemaker	0 (0%)	3 (11.11%)	3 (5.66%)
Not working due to illness/Other	0 (0%)	1 (3.70%)	1 (1.89%)
Income ($n = 45$)			
< £15,000	6 (24%)	1 (5%)	7 (15.56%)
£15,000–£30,999	5 (20%)	3 (15%)	8 (17.78%)
£31,000–£45,999	8 (32%)	2 (10%)	10 (22.22%)
£46,000–£60,999	0 (0%)	7 (35%)	7 (15.56%)
£61,000 or more	6 (24%)	7 (35%)	13 (28.89%)
Living situation ($n = 53$)			
Alone	5 (19.23%)	5 (18.52%)	10 (18.87%)
Spouse/partner	19 (73.08%)	18 (66.67%)	37 (69.81%)
Parents/family	2 (7.69%)	2 (7.41%)	4 (7.55%)
Other	0 (0%)	2 (7.41%)	2 (3.77%)
Relationship status ($n = 53$)			
Single	5 (19.23%)	4 (14.81%)	9 (16.98%)
Partner not cohabiting	3 (11.54%)	5 (18.52%)	8 (15.09%)
Cohabiting/ married	18 (69.23%)	18 (66.67%)	36 (67.92%)
Immigration status ($n = 53$)			
UK National	19 (73.08%)	17 (62.96%)	36 (67.92%)
European Economic Area citizen	4 (15.38%)	3 (11.11%)	7 (13.21%)
Indefinite leave to remain/Exceptional leave to remain	2 (7.69%)	4 (14.81%)	6 (11.32%)
Temporary admission/ Awaiting initial decision	1 (3.85%)	3 (11.11%)	4 (7.55%)
Education ($n = 53$)			
None/only school qualifications	4 (15.38%)	4 (14.81%)	8 (15.09%)
Training/Higer Certificate/Diploma	6 (23.08%)	8 (29.63%)	14 (26.42%)
Degree/Postgraduate	16 (61.54%)	15 (55.56%)	31 (58.49%)
Other children ($n = 53$)			
None	6 (23.08%)	13 (48.15%)	19 (35.85%)
1	4 (15.38%)	3 (11.11%)	7 (13.21%)
≥ 2			

Table 2

Descriptive statistics for key variables at baseline, 14 weeks post-randomisation and 3 months post-delivery for each trial arm.

	Baseline		14 weeks post-randomisation		3 months post-delivery	
	Guided Self Help (plus usual care) group	Usual care	Guided Self Help (plus usual care) group	Usual care	Guided Self Help (plus usual care) group	Usual care
Primary outcome						
EPDS	Median: 15 (11–18) Min: 5; Max: 25	Median: 15 (11–17) Min: 4; Max: 21	Median: 8 (4.5–15) Minimum: 0 Maximum: 25	Median: 12 (9–16) Minimum: 4, Maximum: 22	Median: 6.5 (3–11) Minimum: 0 Maximum: 17	Median: 9 (5–13) Minimum: 1 Maximum: 21
Secondary outcomes						
PHQ-9 depression (score ≥ 10)						
Yes	12 (46.15%)	15 (57.69%)	7 (30.43%)	11 (42.31%)	4 (16.67%)	4 (16%)
No	14 (53.85%)	11 (42.31%)	16 (69.57%)	15 (57.69%)	20 (83.33%)	21 (84%)
GAD-7 anxiety (score ≥ 8)						
Yes	13 (52%)	16 (59.26%)	6 (25%)	11 (42.31%)	5 (21.74%)	12 (48%)
No	12 (48%)	11 (40.74%)	18 (75%)	15 (57.69%)	18 (78.26%)	13 (52%)
Total Postpartum bonding questionnaire score	NA	NA	NA	NA	Median: 6 (5–15) Minimum: 1 Maximum: 32	Median: 7 (4–9) Minimum: 1 Maximum: 22
Other outcomes						
Smoking (nowadays)						
Yes	2 (7.69%)	0 (0%)	1 (4.35%)	0 (0%)	3 (12.5%)	0 (0%)
No	24 (92.31%)	27 (100%)	22 (95.65%)	26 (100%)	21 (87.5%)	24 (100%)
Drinking*						
Yes	5 (20%)	2 (7.69%)	9 (39.13%)	9 (34.62%)	17 (70.83%)	16 (64%)
No	20 (80%)	24 (92.31%)	14 (60.87%)	17 (65.38%)	7 (29.17%)	9 (36%)
Domestic violence (CAS score)**						
0:	20 (90.91%)	24 (100%)	16 (84.21%)	22 (95.65%)	18 (90%)	18 (100%)
1:	2 (9.09%)	0 (0%)	2 (10.53%)	0 (0%)	1 (5%)	0 (0%)
≥2:	0 (0%)	0 (0%)	1 (5.26%)	1 (4.35%)	1 (5%)	0 (0%)
Social provisions scale (total)	Median: 79.5 (71–86) Min: 54; Max: 93	Median: 82.5 Min: 59; Max: 93	Median: 75 (68–82) Min: 46; Max: 92	Median: 82 (74–88) Min: 65; Max: 96	Median: 80 (72.5–90) (74–88) Min: 46; Max: 96	Median: 88 (84–91) Min: 65; Max: 96

* For baseline, “yes” was an AUDIT baseline score ≥ 8; at follow-up, “yes” was an AUDIT-C score ≥ 1.

** At baseline N = 46. 6 participants indicated that they were not in a relationship since finding out that they were pregnant, and 3 patients had missing data on the CAS score. At 14 weeks N = 42. 7 participants indicated that they were not in a relationship since the baseline interview and 4 participants had missing data. At 3 months post-delivery N = 38. 11 participants indicated that they were not in a relationship since the last interview and 4 participants had missing data.

mental health or to the intervention: very preterm birth, severe asthma attack, an incident of domestic violence, post-elective caesarean complication, three postpartum haemorrhages, pre-eclampsia and neonatal blood transfusion.

4.6. Success of blinding

The blinded follow-up researchers correctly guessed the allocation of 50% of the GSH arm at 14 weeks post-randomisation and 32% at 3 months post-delivery. For the usual care arm, correct guesses were made for 70% of participants at 14 weeks post-randomisation and 68% at 3 months post-delivery.

4.7. Aim two: to evaluate if antenatal GSH was beneficial in improving depressive symptoms for women with antenatal depression

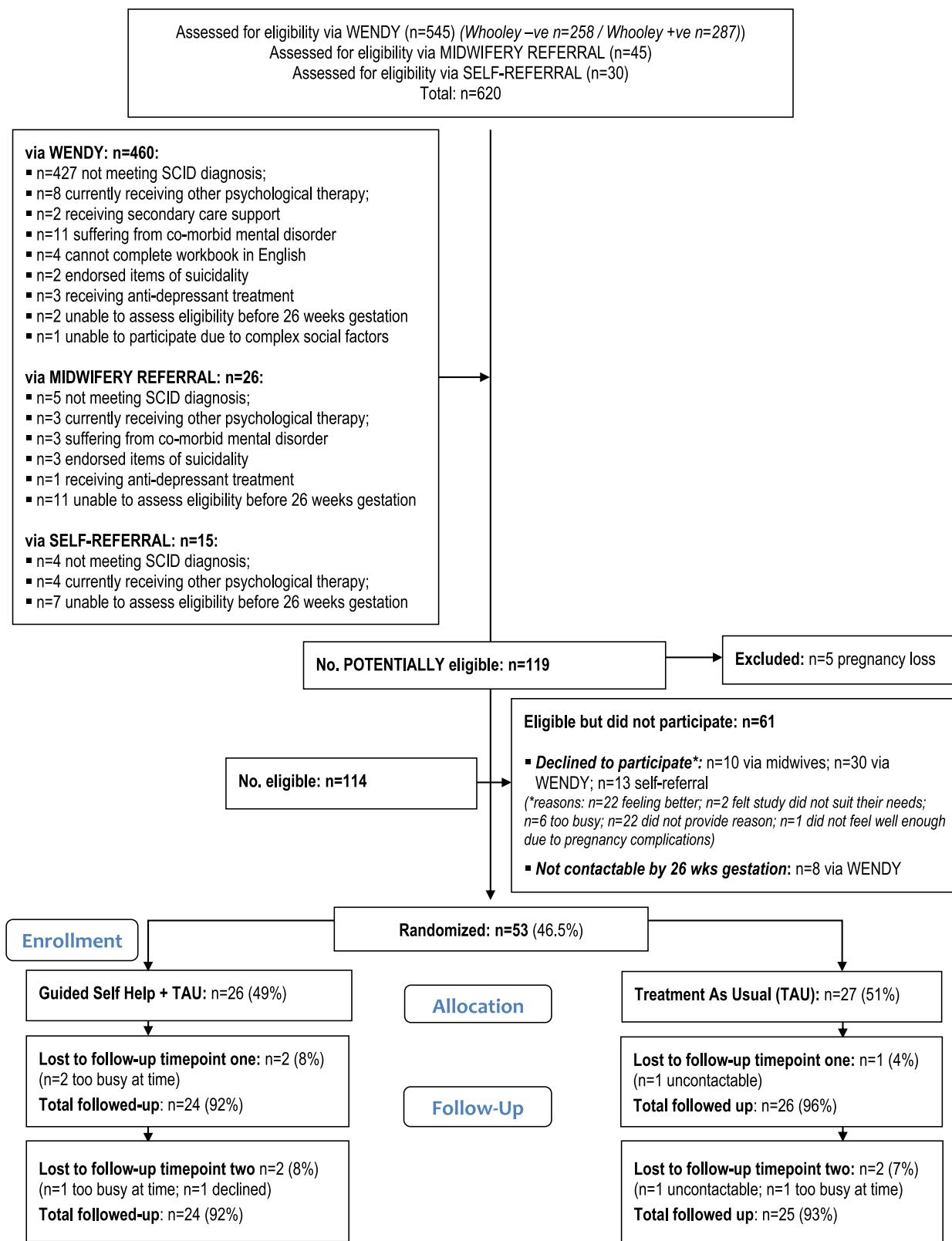
The mean EPDS score at 14 weeks post-randomisation was 9.50 ($sd = 6.35$) for the GSH arm, 12.27 ($sd = 4.79$) for the usual care arm, and 10.94 ($sd = 5.71$) overall. The mean EPDS score at 3 months post-delivery was 7.00 ($sd = 4.86$) in the GSH arm, 9.36 ($sd = 5.33$) in the usual care arm, and 8.20 ($sd = 5.19$) overall.

On the EPDS, an effect size of -0.64 (95%CI: -1.30, 0.06, $p = 0.07$) was observed at 14-weeks post randomisation. At 3 months post-delivery an effect size of -0.39 (95%CI: -1.03, 0.25, $p = 0.24$) was observed. Though the pattern of effect estimates was in the direction favouring GSH, these were individually non-significant (See Fig. 2 and Table 2 for full details).

4.8. Aim three: to evaluate if antenatal GSH had the added benefit of improving other outcomes

At 14 weeks post-randomisation, women in the GSH arm had lower odds of meeting PHQ-9 criteria for depression (AOR: 0.49, 95% CI: 0.09, 2.21; $p = 0.46$) and lower odds of meeting GAD-7 criteria for anxiety (AOR 0.48, 95% CI: 0.08, 2.44; $p = 0.50$) than women in the usual care arm. At 3 months post-delivery, women in the GSH arm had lower odds of meeting GAD-7 criteria for anxiety (AOR 0.37, 95% CI: 0.07, 1.71; $p = 0.25$) and higher odds of meeting PHQ-9 criteria for depression (AOR 1.47, 95% CI: 0.21, 11.71; $p = 0.96$) than women in the usual care arm. These findings were all statistically non-significant (see Fig. 3 and Table 2 for full details). In relation to parenting stress (measured using the PBQ), the GSH arm reported higher mean parenting stress levels than the usual care arm (effect size 0.42 (95%CI: -0.15, 0.97, $p = 0.14$) but the effect estimates were non-significant (See Fig. 2 and Table 2 for full details).

On the Social Provisions Scale, at 14 weeks post-randomisation and 3 months post-delivery the GSH arm reported lower levels of perceived social support than the usual care arm ((effect size of -0.48 (95% CI: -0.89, -0.11, $p = 0.02$ at 14 weeks post-randomisation) / (effect size of -0.40 (95% CI: -0.85, 0.02, $p = 0.07$ at 3 months post-delivery)) but the observed effects were non-significant. Most other outcomes were comparable across both arms at each time-point (see Table 2 for full details).

**Fig. 1.** CONSORT diagram.

4.9. Aim four: to explore if antenatal GSH was cost-effective compared to usual care

Eighty-one percent (43/53) of participants had complete data for inclusion in the case economic analyses based on SF-6D derived QALYs,

with the GSH arm having slightly fewer complete cases than the usual care group (77% versus 85%).

The mean cost of delivering GSH was £418 (£214 SD) per participant. Total costs and SF-6D-based QALYs were similar in the two groups at 3 months post-delivery (see Table 3). Results based on the secondary

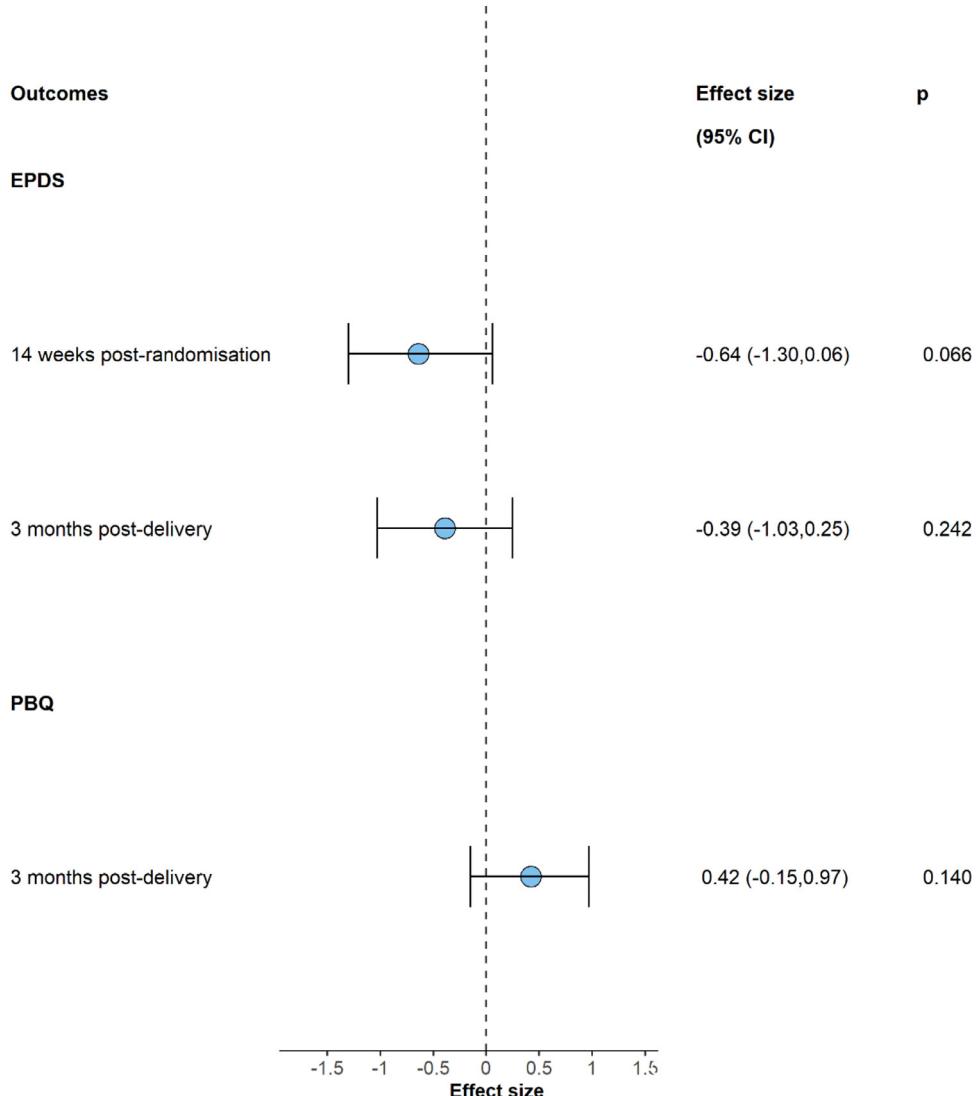


Fig. 2. Effect size plot of the Edinburgh Postnatal Depression Scale (EPDS) and Postpartum Bonding questionnaire (PBQ) at each follow up time point. Key: In each comparison, Guided Self-Help is compared to usual care alone (reference arm). A positive effect size corresponds to the Guided Self-Help arm having a larger mean; a negative effect size corresponds to the usual care arm having a larger mean. For all measures, an increase in score corresponds to a worse outcome

analysis using EQ-5D-5L-based QALYs, and results of the sensitivity analyses did not alter the significance of these results. Full details on the economic evaluation results are available as supplementary information.

The incremental cost-effectiveness ratio based on QALYs calculated from the SF-6D was £7200 per QALY (GSH was cheaper but less effective on average than usual care alone). The cost-effectiveness acceptability curve based on SF-6D QALYs shows that the probability of GSH being cost-effective compared with usual care is around 50% at the NICE preferred willingness-to-pay threshold of £20 000–£30 000 per QALY (see Fig. 4). When based on QALYs calculated from the EQ-5D-5L, the results were similar although the probability of GSH being cost-effective compared with usual care dropped to around 40% at the NICE preferred willingness-to-pay threshold. When examining the cost-effectiveness based on sensitivity analyses, the probability of GSH being cost-effective compared with usual care at the NICE preferred willingness-to-pay threshold was around 20% for the narrow mental health care perspective, around 50% for the adjusted PWP indirect time analysis, and around 10% for the analysis with outliers and influential observations removed (see supplementary information 3).

5. Discussion

This RCT is one of the first to provide evidence on the efficacy and cost-effectiveness of antenatal GSH in improving depressive symptoms among pregnant women with DSM-IV mild to moderate depression. Our trial procedures were feasible (aim 1): pregnant women were successfully randomised and retained in both arms; we achieved high levels of compliance and treatment fidelity, and no trial-related adverse events were reported. Due to changes in NHS policy, we did not meet our anticipated recruitment targets. GSH signalled reductions in depressive symptoms post-treatment (aim 2) but it had mixed effects in improving other key outcomes for women (aim 3). Finally, economic analyses indicate little difference in the costs and outcomes between the arms (aim 4), meaning the cost-effectiveness of GSH versus usual care remains uncertain.

This RCT suggests that GSH can be successfully delivered to depressed women in early pregnancy and this form of treatment does not cause harm. There was a trend towards improvements in symptoms of depression post-treatment among women who received GSH. These findings, which are important targets for early intervention, suggest this intervention is suitable for use in talking therapy services. Despite this, we conclude that it is not feasible to upscale the intervention to a full

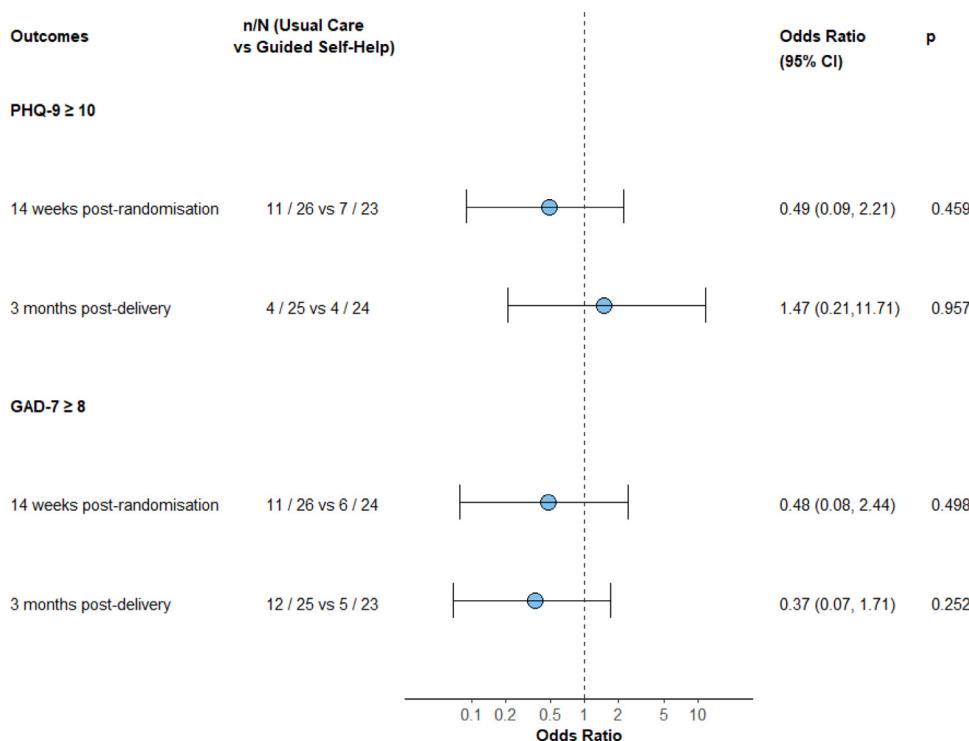


Fig. 3. Plot of adjusted odds ratios between the trial arms for the PHQ-9 and GAD-7 measures at each follow-up time point.

Key: In each comparison, Guided Self-Help is compared to usual care alone; an odds ratio greater than one corresponds to patients randomised to Guided Self-Help being more likely than patients randomised to usual care alone to have the condition (i.e. depression on PHQ-9 or anxiety on GAD-7); an odds ratio less than one corresponds to patients randomised Guided Self-Help being less likely than patients randomised to usual care alone to have the condition.

effectiveness RCT in the UK. This is because substantial national developments were made during the running of the RCT regarding perinatal IAPT services, including development of new community perinatal service provisions. Alongside this an NHS England policy was established, with new quality standards implemented for pregnant women and additional funds for perinatal mental health. These developments changed local service practices; IAPT services became more proactive in facilitating access by pregnant women, and midwives made more referrals to IAPT. It is likely that these changes reduced our ability to recruit to the trial; as a result, we ended the trial earlier than anticipated. Nevertheless, we had minimal loss-to-follow-up across both arms.

Among women randomised to GSH, the majority received the minimum dose of four sessions (69%). These findings are similar to a recent Australian RCT of an eight-session antenatal CBT intervention for pregnant women (Milgrom et al., 2014). We found that women receiving the intervention reported higher session attendance rates (i.e. 68%) and lower "did not attend" (DNA) rates (i.e. 4%) compared to standard talking-therapy services (Binnie and Boden, 2016; Richards and Borglin, 2011). Women allocated to GSH elected to receive both face-to-face and telephone-based sessions, with high levels of therapeutic alliance reported overall. These results provide support for research findings that levels of therapeutic alliance do not significantly differ if CBT-based interventions are delivered face-to-face versus over the phone (Stiles-Shields et al., 2014).

The primary economic analysis suggests little difference in the costs and outcomes between the arms. However, sensitivity analyses taking a narrow mental health care perspective and the removal of outliers and influential observations suggest these results are sensitive to perspective and outliers. In this study, these two sensitivity analyses overlapped significantly as the outliers were due to complications with child-birth and thus were removed from both sensitivity analyses. The cost-effectiveness planes also demonstrate large variation driven by the small sample size. Therefore, the cost-effectiveness of GSH versus usual care remains uncertain.

6. Limitations

Despite intense efforts we did not meet our anticipated recruitment target. Although blinded researchers only correctly identified the allocation of half of the participants receiving GSH, they correctly identified the allocation of 70% of participants who received usual care. The latter finding may have resulted in possible rater bias for some usual care arm cases. Resource use data were collected using self-report, making it subject to participant recall bias. However, this approach was necessary due to these data being unavailable from another single source. Further, there is evidence that self-reported resource use is reliable even in populations who have cognitive deficits (Calsyn et al., 1993; Goldberg et al., 2002).

7. Conclusions

The absence of trial-related adverse events suggests that this GSH intervention is safe to deliver to pregnant women meeting diagnostic criteria for mild to moderate depression. In addition, observed trends in relation to improvements in women's depressive symptoms post-treatment suggests GSH is suitable for use in psychological therapy services. Future trials of this intervention would not be feasible, due to national policy developments, but challenges regarding participation may be overcome if this treatment were embedded as part of routine care within psychological therapy services.

Authors' contributions

LMH, MSHunter, KT, EGR, MHeslin, SB, JM, AP, CP, DB, SN, JD and RM have made substantial contributions to conception and design, acquisition of data, analysis and/or interpretation of data; LMH, EGR, KT, MHeslin, RM and JD have been involved in drafting the manuscript; AP, DB, SN, JM, SB and CP revised it critically for important intellectual content. SB and MHeslin designed and conducted all economic analyses and wrote the economic sections of the paper. EGR performed the statistical analyses and these were overseen by AP. All authors contributed to the writing of this manuscript, including reviewing and editing the manuscript content, and have given final approval of the

Table 3
Total costs and outcomes over the follow-up period.

	Treatment arm: Guided Self-Help group Valid n	Mean (SD)	Comparison arm: Usual care Valid n	Mean (SD)	Unadjusted mean difference	95% C.I. ^a	Adjusted mean difference ^a	95% C.I. ^a	p-value	p-value ^a
COSTS										
Main analysis										
Total baseline to 3 months post-delivery costs	21	8251 (2909)	25	8332 (6732)	-80	-2976 to 2816	0.957	-72	-3045 to 2901	0.962
Sensitivity analyses										
Mental health care perspective	21	860 (1138)	25	560 (290)	300	-393 to 993	0.396	310	-372 to 993	0.373
Alternative direct to indirect time ratio for PWP	21	8100 (2931)	25	8332 (6732)	-231	-3131 to 2669	0.876	-223	-3200 to 2754	0.883
Outliers and influential observations removed	20	7863 (2360)	22	6820 (2185)	1043	-315 to 2402	0.132	1055	-337 to 2448	0.138
OUTCOMES										
Main analysis										
SF-6D based QALY	20	0.50 (0.08)	23	0.49 (0.08)	0.01	-0.04 to 0.06	0.734	-0.01	-0.06 to 0.04	0.827
Secondary analysis										
EQ-5D-5L based QALY	21	0.63 (0.15)	23	0.60 (0.13)	0.03	-0.06 to 0.11	0.526	-0.01	-0.08 to 0.07	0.857
Non-parametric bootstrap regressions (10,000 repetitions, bias-corrected).										

^a Adjusted for depression stratifier, trial arm, and baseline variable of interest.

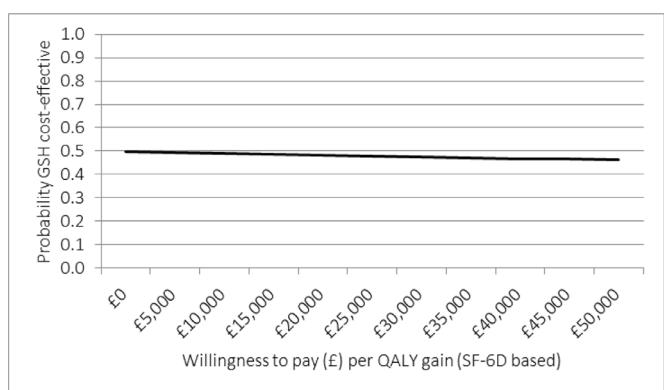


Fig. 4. Cost-effectiveness acceptability curves for GSH versus treatment as usual at 3 months post-delivery from the health- and social-care perspective for SF-6D based QALYs.

version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Role of funding sources

This paper summarises independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research (PGfAR) Programme (Grant Reference Number: RP-PG-1210-12002) and the National Institute for Health Research (NIHR) / Wellcome Trust King's Clinical Research Facility. The study team acknowledges the study delivery support given by the south London NIHR Clinical Research Networks and King's Clinical Trials Unit. This study represents independent research supported by the NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. Louise Howard is also supported by an NIHR Research Professorship (NIHR-RP-R3-12-011) and Debra Bick is also supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care South London at King's College Hospital NHS Foundation Trust. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. The funding source and sponsor had no role in the design, execution or dissemination of the results, other than arranging peer review.

Declaration of Competing Interest

Professor Louise Howard chaired the NICE CG192 guidelines development group on antenatal and postnatal mental health in 2012–2014.

Acknowledgements

We would like to acknowledge the members of the Trial Steering Committee (Prof Catherine Creswell (chair), Vivette Glover, Juliette O'Donnell and Christine MacArthur) and the Independent Data Monitoring Committee (Roch Cantwell, Liz McDonald, Marian Knight and Stephen Bremner).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2019.10.013.

References

- Assmann, S.F., Pocock, S.J., Enos, L.E., Kasten, L.E., 2000. Subgroup analysis and other (mis) uses of baseline data in clinical trials. *Lancet* 355, 1064–1069.
- Babor, F., Higgins-Biddle, J.C., Saunders, J.B., Monterio, M.G., 2008. Audit - The Alcohol Use Disorders Identification Test - Guidelines for use in primary care, Second ed. The World Health Organisation, Geneva.
- Bick, D., Howard, L.M., 2010. When should women be screened for postnatal depression? *Expert Rev. Neurother.* 10, 151–154.
- Binnie, J., Boden, Z., 2016. Non-attendance at psychological therapy appointments. *Ment. Health Rev.* 21, 231–248.
- Blackburn, I.-M., James, I.A., Milne, D.L., Baker, C., Standart, S., Garland, A., Reichelt, F.K., 2001. The revised cognitive therapy scale (CTS-R): psychometric properties. *Behav. Cogn. Psychother.* 29, 431–446.
- Brockington, I.F., Oates, J., George, S., Turner, D., Vostanis, P., Sullivan, M., Loh, C., Murdoch, C., 2001. A screening questionnaire for mother-infant bonding disorders. *Arch. Womens Ment. Health* 3, 133–140.
- Burns, A., O'Mahen, H., Baxter, H., Bennert, K., Wiles, N., Ramchandani, P., Turner, K., Sharp, D., Thorn, J., Noble, S., Evans, J., 2013. A pilot randomized controlled trial of cognitive behavioural therapy for antenatal depression. *BMC Psychiatry* 13, 33.
- Calsyn, R.J., Allen, G., Morse, G.A., Smith, R., Tempelhoff, B., 1993. Can you trust self-report data provided by homeless mentally ill individuals? *Evalu. Rev.* 17, 353–366.
- Clark, D.M., 2011. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *Int. Rev. Psychiatry* 23, 318–327.
- Clark, D.M., Canvin, L., Green, J., Layard, R., Pilling, S., Janecka, M., 2018. Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. *Lancet* 391, 679–686.
- Coull, G., Morris, P., 2011. The clinical effectiveness of CBT-based guided self-help interventions for anxiety and depressive disorders: a systematic review. *Psychol. Med.* 41, 2239.
- Cox, J., Holden, J., 2003. Perinatal Mental Health: A Guide to the Edinburgh Postnatal Depression Scale (EPDS). Royal College of Psychiatrists.
- Cuijpers, P., Donker, T., van Straten, A., Li, J., Andersson, G., 2010. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychol. Med.* 40, 1943–1957.
- Curtis, L., Burns, A., 2015. Unit Costs of Health and Social Care. Canterbury: Personal Social Services Research Unit. University of Kent 2015.
- Cutrona, C.E., Russell, D.W., 1987. The provisions of social relationships and adaptation to stress. *Adv. Pers. Relatsh.* 1, 37–67.
- Eldridge, S.M., Chan, C.L., Campbell, M.J., Bond, C.M., Hopewell, S., Thabane, L., Lancaster, G.A., 2016. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Pilot Feasibility Stud.* 2, 64.
- Fenwick, E., Byford, S., 2005. A guide to cost-effectiveness acceptability curves. *Br. J. Psychiatry* 187.
- Goldberg, R.W., Seybold, D.C., Lehman, A., 2002. Reliable self-report of health service use by individuals with serious mental illness. *Psychiatr. Serv.* 53, 879–881.
- Hegarty, K., Bush, R., Sheehan, M., 2005. The Composite abuse scale: further development and assessment of reliability and validity of a multidimensional partner abuse measure in clinical settings. *Violence Victims* 529–547.
- Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G., Badia, X., 2011. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual. Life Res.* 20, 1727–1736.
- Hoffmann, T.C., Glasziou, P.P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D.G., Barbour, V., Macdonald, H., Johnston, M., 2014. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 348, g1687.
- Howard, L.M., 2011. Effectiveness and Cost-Effectiveness of Perinatal Psychiatry Services. NIHR Programme Development Grant, London Final Report NIHR.
- Howard, L.M., Flach, C., Leese, M., Byford, S., Killaspy, H., Cole, L., Lawlor, C., Betts, J., Cutting, P., McNicholas, S., Sharac, J., Johnson, S., 2010. The effectiveness and cost effectiveness of admissions to women's crisis houses compared with traditional psychiatric wards - a pilot patient preference randomized controlled trial. *Br. J. Psychiatry* 197, s32–s40.
- Howard, L.M., Megnin-Viggars, O., Symington, I.a., Pilling, S., 2014a. Antenatal and postnatal mental health: summary of updated NICE guidance. *BMJ* 349, g7394.
- Howard, L.M., Molyneaux, E., Dennis, C.-L., Rochat, T., Stein, A., Milgrom, J., 2014b. Non-psychotic mental disorders in the perinatal period. *Lancet* 384, 1775–1788.
- Howard, L.M., Ryan, E.G., Trevillion, K., Anderson, F., Bick, D., Bye, A., Byford, S., O'Connor, S., Sands, P., Demilew, J., 2018. Accuracy of the Whooley questions and the Edinburgh Postnatal Depression Scale in identifying depression and other mental disorders in early pregnancy. *Br. J. Psychiatry* 212, 50–56.
- Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., Augustovski, F., Briggs, A.H., Mauskopf, J., Loder, E., 2013. Consolidated health economic evaluation reporting standards (CHEERS) statement. *Cost Eff. Resour. Alloc.* 11, 6.
- Improving Access to Psychological Therapies, 2011. The IAPT Data Handbook: Guidance on Recording and Monitoring Outcomes to Support Local Evidence-Based Practice. IAPT Programme, London.
- Kroenke, K., Spitzer, R.L., Williams, J.B.W., 2001. The PHQ-9. *J. Gen. Internal Med.* 16, 606–613.
- Manca, A., Hawkins, N., Sculpher, M.J., 2005. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ.* 14, 487–496.
- Milgrom, J., Gemmill, A.W., Bilszta, J.L., Hayes, B., Barnett, B., Brooks, J., Buist, A., 2008a. Antenatal risk factors for postnatal depression: a large prospective study. *J. Affect. Disord.* 108, 147–157.
- Milgrom, J., Gemmill, A.W., Bilszta, J.L., Hayes, B., Barnett, B., Brooks, J., Erickson, J., Ellwood, D., Buist, A., 2008b. Antenatal risk factors for postnatal depression: a large prospective study. *J. Affect. Disord.* 108, 147–157.
- Milgrom, J., Holt, C., Ross, J., Erickson, J., Gemmill, A., 2014. Feasibility study and pilot randomised trial of an antenatal depression treatment with infant follow-up. *Arch. Womens' Ment. Health* 1–14.
- Milgrom, J., Schembri, C., Erickson, J., Ross, J., Gemmill, A.W., 2011. Towards parenthood: an antenatal intervention to reduce depression, anxiety and parenting difficulties. *J. Affect. Disord.* 130, 385–394.
- National Institute for Health and Care Excellence, 2008. Guide to the Methods of Technology Appraisal. <http://www.nice.org.uk/aboutnice/howtework/devnitech/technologyappraisalprocessguides/guidetothemethodsoftechnologyappraisal.jsp>.
- NHS Digital, 2016. Psychological Therapies: Annual Report on the use of IAPT Services in England (2015–16). NHS Digital, Leeds.
- Richards, D.A., Borglin, G., 2011. Implementation of psychological therapies for anxiety and depression in routine practice: two year prospective cohort study. *J. Affect. Disord.* 133, 51–60.
- Sockol, L.E., Epperson, C.N., Barber, J.P., 2011. A meta-analysis of treatments for perinatal depression. *Clin. Psychol. Rev.* 31, 839–849.
- Spitzer, R.L., Kroenke, K., Williams, J.B., Lowe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch. Internal Med.* 166, 1092–1097.
- Stein, A., Pearson, R.M., Goodman, S.H., Rapo, E., Rahman, A., McCallum, M., Howard, L.M., Pariante, C.M., 2014. Effects of perinatal mental disorders on the fetus and child. *Lancet* 384, 1800–1819.
- Stiles-Shields, C., Kwasny, M.J., Cai, X., Mohr, D.C., 2014. Therapeutic alliance in face-to-face and telephone-administered cognitive behavioral therapy. *J. Consult. Clin. Psychol.* 82, 349.
- Thompson, S.G., Barber, J.A., 2000. How should cost data in pragmatic randomised trials be analysed? *Br. Med. J.* 320, 1197–1200.
- Trevillion, K., Domoney, J., Pickles, A., Bick, D., Byford, S., Heslin, M., Milgrom, J., Mycroft, R., Pariante, C.M., Ryan, E., Hunter, M., Howard, L.M., 2016. Depression: an exploratory parallel-group randomised controlled trial of Antenatal guided self help for Women (DAWN): study protocol for a randomised controlled trial. *Trials* 17.
- Ware Jr, J.E., 2000. SF-36 health survey update. *Spine* 25, 3130–3139.