Universal prevention of depression in women postnatally: cluster randomized trial evidence in primary care

T. S. Brugha[1]*, **C. J. Morrell**[2], **P. Slade**[3] and **S. J. Walters**[4]

[1] *Clinical Division of Psychiatry, Department of Health Sciences, University of Leicester, Leicester, UK*
[2] *School of Nursing, Faculty of Medicine and Health Sciences, Midwifery and Physiotherapy, University of Nottingham, Nottingham, UK*
[3] *Clinical Psychology Unit, Department of Psychology, University of Sheffield, Sheffield, UK*
[4] *Medical Statistics Group, School of Health and Related Research, University of Sheffield, Sheffield, UK*

Background. To test whether receiving care from a health visitor (HV) trained in identification and psychological intervention methods prevents depression 6–18 months postnatally in women who are not depressed 6 weeks postnatally.

Method. The study was a prospective cluster trial, randomized by GP practice, with follow-up for 18 months in 101 primary care teams in the Trent area of England. The participants were women scoring <12 on the postal Edinburgh Postnatal Depression Scale (EPDS) at 6 weeks postnatally (1474 intervention and 767 control women). Intervention HVs ($n=89$, 63 clusters) were trained in identifying depressive symptoms using the EPDS and face-to-face clinical assessment and in providing psychologically orientated sessions based on cognitive behavioral or person-centered principles. The control group comprised HVs ($n=49$, 37 clusters) providing care as usual (CAU). The primary outcome measure was the proportion of women scoring ⩾12 on the EPDS at 6 months postnatally. Secondary outcomes were mean EPDS score, Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) score, State–Trait Anxiety Inventory (STAI), 12-item Short Form Health Survey (SF-12) and Parenting Stress Index Short Form (PSI-SF) scores at 6, 12 and 18 months.

Results. After adjusting for individual-level covariates, living alone, previous postnatal depression (PND), the presence of one or more adverse life events and the 6-week EPDS score, the odds ratio (OR) for EPDS ⩾12 at 6 months was 0.71 [95% confidence interval (CI) 0.53–0.97, $p=0.031$] for the intervention group (IG) women compared with the control (CAU) group women. Two subgroups were formed by baseline severity: a 'subthreshold' subgroup with a 6-week EPDS score of 6–11 ($n=999$) and a 'lowest severity' subgroup with a 6-week EPDS score of 0–5 ($n=1242$). There was no difference in psychological effectiveness by subgroup (interaction term: $z=-0.28$, $p=0.782$).

Conclusions. This study provides new evidence of a universal, enduring preventive effect for depression in women who screen negative for depression postnatally.

Introduction

The need for effective population-wide prevention approaches to depression, including mixed depression and anxiety (Tyrer, 2001; Das-Munshi *et al.* 2008), has been highlighted by the demonstration of undiminished population prevalence rates despite at least a twofold increase in the usage of evidence-based treatments (Brugha *et al.* 2004; Kessler *et al.* 2005). Evidence of prevention of depression has not been reported in a trial in adults but may be feasible; in a recent meta-analysis synthesizing the results of 19 small prevention trials (Cuijpers *et al.* 2008), the mean incidence rate ratio (IRR) of developing a depressive disorder (including postpartum depression) was 0.78 [95% confidence interval (CI) 0.65–0.93], indicating a reduction of the incidence by 22% in experimental compared to control groups.

Prevention can be considered as avoiding future illness in those already at risk although disease free or, more strictly, as avoiding future illness for all the

* Address for correspondence: Professor T. S. Brugha, M.D. (N.U.I.), F.R.C. Psych., Department of Health Sciences, University of Leicester, Leicester General Hospital, Leicester LE5 4PW, UK.
(Email: tsb@leicester.ac.uk)

population, not just those at immediate risk (Rose, 2001), hence the term universal prevention (Gordon, 1983; Mrazek & Haggerty, 1994). But how might this be achieved? A highly significant predictor of depression within adulthood is 'subthreshold' depression (Judd *et al.* 2002; Cuijpers & Smit, 2004). Depression prevention through the extension of cognitive behavioral therapy (CBT) to at-risk adolescents who had symptoms that fell short of the threshold for case depression (Clarke *et al.* 2001; Garber *et al.* 2009) may be effective because subthreshold levels are highly predictive of depression (Cuijpers *et al.* 2005). It is not known whether such skilful psychological approaches could also benefit a much wider range of people at risk of developing depression if delivered universally across an entire population to adults at all risk levels.

In the UK all infants and all mothers, following childbirth, receive individual care from a specialist community nurse, known as a 'Health Visitor' (HV). In addition to supporting infant care, the HV has a role in maternal mental health that should involve establishing a relationship with the mother and the use of interpersonal and communication skills (Morrell *et al.* 2009*b*). However, such practitioners are given little more than basic mental health knowledge. Could training HVs in psychological skills benefit women under their care who were not at immediate risk? In this paper we report on the benefits for all non-depressed postnatal women randomized to receive support from primary care nurses with specialized training in psychologically informed approaches to health care, in a large pragmatic cluster randomized controlled and economic evaluation trial.

We have already reported on postnatal women in randomly allocated intervention group (IG) clusters cared for by an HV who had received additional training in postnatal mental health assessment and in one of two psychologically informed approaches that were compared to usual HV care (Morrell *et al.* 2009*a*). Clinical assessment and psychologically informed approaches were designed to be offered by intervention group HVs to women scoring positive on the Edinburgh Postnatal Depression Scale (EPDS) (Morrell *et al.* 2009*b*) at 8 weeks following childbirth. The EPDS is a widely used self-report measure to identify women at risk of postnatal depression (PND) with good evidence of sensitivity and specificity (Gibson *et al.* 2009; Hewitt *et al.* 2009) and a score ranging from 0 to 30, consisting of eight depression and two anxiety items. We compared IG care with postnatal care as usual (CAU). We showed that these scoring positive IG women were less likely than CAU women to have later depression, indicated by scoring $\geqslant 12$ on a postal EPDS at 6 months postnatally [odds

ratio (OR) 0.62, 95% CI 0.04–0.97, unadjusted $p = 0.036$] (Morrell *et al.* 2009*a*).

The trial was also designed to assess effectiveness according to severity of depression at baseline (www.controlled-trials.com/ISRCTN92195776). Because outcome data were obtained on all randomized and consenting women, we were able to estimate any effects on women with baseline symptom scores below the 6 weeks postnatal EPDS threshold of $\geqslant 12$ (Cox *et al.* 1987). Thus, the present paper reports new, unpublished analyses of outcomes from the original treatment trial, principally for women who scored < 12 on the postal EPDS at 6 weeks postnatally, using the same trial primary outcome: EPDS $\geqslant 12$ at 6 months. First, effects on outcomes at 6, 12 and 18 months postnatally were examined for women with an EPDS < 12 at 6 weeks. Second, to test whether additional training of HVs was of benefit to women across a lower range of EPDS-negative scores, we compared the effectiveness in women at the lowest EPDS severity levels (EPDS < 6 at 6 weeks) with subthreshold scoring women (6-week EPDS scores from 6 to 11). Third, we examined differences in HV care, including the use of randomly allocated face-to-face administered depression assessment, because such additional practitioner involvement could explain any observed effect on later risk of depression.

Method

The main pragmatic cluster randomized trial design and methods, on which the present analysis is based, are reported in detail elsewhere (Morrell *et al.* 2009*a*). The present study involved assessing outcomes for women in 101 primary care practices (clusters) (in one region of England, serving a population of 5.1 million), termed 'EPDS-negative', who had a postal EPDS score < 12 at 6 weeks postnatally. Following recruitment of participating practices, an independent statistician generated the random allocation sequence using a computer randomization program (Morrell *et al.* 2009*a*). Thus, HVs registered to consenting practices were randomized either to continue to provide postnatal CAU or to an experimental IG in which all HVs received additional training in systematic assessment of depressive symptoms, establishing warm, therapeutic relationships, and in one of two distinct experimental psychologically informed approaches. Eligible antenatal women were then recruited if they were registered with participating (and therefore randomized) GP practices, aged $\geqslant 18$ years and had no severe or enduring mental health problems. HVs were asked to send these women a research information leaflet and a consent form at 36 weeks of pregnancy or 4 weeks antenatally. Women returned their signed

forms to the HVs. Altogether, 4084 (53.4%) of 7649 eligible antenatal women provided written informed consent during 2003–2004. Consenting women were sent a questionnaire at 6 weeks and 6, 12 and 18 months postnatally from the research office. Women who scored ⩾12 on their 6-week postal EPDS were regarded as EPDS positive ($n=595$ women in 86 clusters) and were excluded from all of the present outcomes analyses. Half of all women in the intervention clusters (including EPDS-negative women) were randomly allocated to complete a face-to-face EPDS 6 weeks postnatally to coincide with usual HV care during an existing HV contact. The intervention cluster HVs, if their clinical assessment indicated that any mother might benefit, could also offer the psychologically informed sessions to which their practice had been randomized.

Identical to the treatment trial protocol (Morrell *et al.* 2009*b*), the pretrial determined primary outcome was the proportion of women scoring ⩾12 on the EPDS at 6 months postnatally. We assumed an intracluster coefficient (ICC) of 0.006 (Morrell *et al.* 2009*b*) and an average cluster of practice size of 25 EPDS-negative women. Thus, 2489 women (1659 IG and 830 CAU) were required to have 80% power at the 5% two-sided level of significance, to detect a 3.5% absolute difference (10% *v.* 6.5%) in the proportions of 6-week EPDS-negative IG and CAU women scoring ⩾12 on the EPDS at 6 months postnatally. Clusters were allocated for the trial to one of two HV psychologically based approaches (experimental IG) or HV CAU, in a ratio of 1:1:1. The clusters were stratified by the number of expected births per cluster per year, into three groups (<70, 70–100, >100).

The cluster level intervention comprised the package of HV training in assessment of postnatal women and identification of depressive symptoms using the EPDS and additional clinical assessment skills and establishing warm, therapeutic relationships. They were trained to deliver one of two distinct psychologically informed approaches [a cognitive-behavioral approach (CBA) and a person-centered approach (PCA)], delivered at the individual level, with the provision of additional supervisory support for the HVs (Morrell *et al.* 2009*b*). The HV-provided psychologically informed session was a one-hour visit, once a week, for a maximum of 8 weeks, commencing around 8 weeks postnatally. Sessions were offered to women who scored ⩾12 on the postal EPDS at 6 weeks and also on a face-to-face EPDS administered at 8 weeks postnatally.

IG HVs received a day of training in the identification of depression using the EPDS and clinical assessment. The group training in the pretrial preparatory phase lasted 8 days in total (including four half-days), focusing on the development of cognitive behavioral or person-centered intervention skills delivered in a group training format. A common element of the training, preceding specific approach-based skills, concerned development of empathic relationships. The training package was followed by two half-day reflective practice sessions and access to regular supervision sessions. Two experienced psychotherapist practitioner trainers delivered the training, which followed a prepared manualized form and was agreed by a training reference group composed of experienced psychotherapy trainers. HVs were asked to tape record their intervention sessions (with the women's consent) so that session treatment fidelity could be checked. From the recorded sessions available it was clear that high levels of adherence were achieved. HVs and women were also asked to complete the Agnew Relationship Measure to measure the therapeutic relationship between them. Further details of the training are available in Morrell *et al.* (2009*b*) and in findings to be reported separately.

HVs were blind to all research outcomes. Secondary outcomes were measured in all women who were followed up at 6, 12 and 18 months postnatally by postal questionnaires (including the EPDS at 6 and 12 months only). These included the 12-item Short Form Health Survey (SF-12) and the SF-6D (Ware *et al.* 1995), the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) scales (Barkham *et al.* 2001) and the State–Trait Anxiety Inventory (STAI; Spielberger *et al.* 1970). We used the CORE-OM as a measure of global distress as used by psychological therapy services to distinguish between clinical and general populations. We measured postnatal anxiety symptoms using the STAI (Spielberger *et al.* 1970). To achieve the required sample size for the 6-month primary outcome, the recruitment phase was extended. However, the follow-up phase was not similarly extended; therefore, all the women for whom a 6-month outcome was available did not have a baby who was 1 year or 18 months old within the follow-up phase. Therefore, fewer questionnaires were administered at these time points. We requested funds to extend the recruitment phase but none were provided.

Ethical approval was received from Trent Region Multi-centre Research Ethics Committee.

Statistical analysis

The primary comparison reported for this analysis was between eligible women in randomized IG practices *versus* those in CAU randomized practices, with an EPDS score <12 at 6 weeks postnatally in addition to a valid 6-month EPDS outcome score ($n=2241$

EPDS-negative women in 101 clusters). All analyses were by intention to treat with a p value <0.05 regarded as statistically significant. A marginal generalized linear model (GLM), with coefficients estimated using generalized estimating equations (GEEs), with robust standard errors and an exchangeable autocorrelation matrix in Stata version 11 (StataCorp, 2009) was used to analyze outcomes and allow for the clustered nature of the data, with associated 95% CIs. We fitted simple unadjusted models, and also adjusted outcome comparisons for individual level covariates: living alone, history of PND, presence of stressful life events (Brugha *et al.* 1985) and 6-week EPDS score, as the most widely identified significant predictors of PND (O'Hara & Swain, 1996). To test whether any intervention effect was confined to those scoring immediately below the EPDS scoring cut-off only, we performed an intervention by baseline (6-week) EPDS score statistical interaction test. This compared women grouped to baseline EPDS scores of 6–11 with women with a baseline EPDS score <6. We compared 6-month outcomes according to whether women were in practices randomized for HVs to provide a face-to-face assessment at 6 weeks plus postal assessment compared to postal EPDS only. Evidence for diminished and differential health benefit over time, between the groups, was tested for in a longitudinal model by looking for evidence of an IG by time interaction. Thus, longitudinal modeling of 6, 12 and 18 months secondary outcomes, also using a marginal GLM with coefficients estimated by GEEs, was performed to examine time effects, group differences and possible time by group interactions on the secondary outcomes. The longitudinal models adjust for clustering of individual responses over time and not by practice. For the primary outcomes, the binary EPDS scores at 6 and 12 months, we used the multiple imputation procedure in STATA version 11 (StataCorp, 2009) to impute missing EPDS outcomes based on 100 imputations using a logistic regression model with 6-week EPDS score, lives alone, history of PND and life events as predictors. The trial is reported according to the CONSORT guidelines for cluster trials (Campbell *et al.* 2004).

Results

Figure 1 shows the trial profile. The 6-week questionnaire return rate was 83.7% (3419/4084), of whom 17.4% of women (595/3419) had an EPDS score of $\geqslant 12$ and 82.6% (2824/3419) were EPDS-negative. At 6 months, 79.4% (2241/2824) of the EPDS-negative women had returned both 6-week and 6-month questionnaires, 767 in the control (CAU) group and 1474 in the IG. Table 1 shows the baseline and 6-week

characteristics of these 2241 new mothers. The observed ICC for the primary outcome, the proportion of women with an EPDS score of $\geqslant 12$ at 6 months, in the 100 clusters was 0.001 (95% CI 0.000–0.015).

As reported previously (Morrell *et al.* 2009*a*), Table 2 shows that, at 6 months following childbirth, 83 out of 767 (10.8%) control (CAU) women and 113 of 1474 IG women (7.7%) scored $\geqslant 12$ on the EPDS at 6 months, that is an absolute difference of 3.1% (95% CI 0.4–5.5) or a relative difference of 0.68 (95% CI 0.50–0.93, $p=0.016$). Of the covariates (living alone, previous PND, presence of adverse life events in the past 6 months and 6-week EPDS score), only living alone was not a significant predictor of the EPDS at 6 months. After adjusting for individual-level covariates, the OR for a score of $\geqslant 12$ at 6 months was 0.71 (95% CI, 0.53–0.97, $p=0.031$) for IG women compared with CAU women. Multiple imputation computing for missing 6-month EPDS scores increased the sample size from 2241 to 2786, and the OR using the imputed data, with adjustment for covariates, was similar to the observed data (0.76, 95% CI 0.56–1.03, $p=0.073$).

The EPDS-negative women were divided into two subgroups: a 'subthreshold' subgroup scoring between 6 and 11 on the 6-week EPDS ($n=999$) and a 'lowest severity' subgroup scoring 0–5 on the 6-week EPDS ($n=1242$) based on the approximate median score for the sample. The choice of a cut-point between 5 and 6 on the EPDS was supported by the finding that 55% (1242/2241) of EPDS-negative women scored 0–5. Those in the subthreshold subgroup (score 6–11) did not benefit significantly more from IG membership as defined by the 6-week EPDS score (0–5, compared to 6–11); a significant intervention by subgroup interaction was not found ($z=-0.28$, $p=0.782$) (Table 2). The ORs for the effect of HV training in these two EPDS score subgroups scarcely differed (0.76 and 0.69 respectively; Table 2). The results of a similar comparison of low- and high-risk women also showing no difference in effectiveness are available on request.

Among the secondary outcomes of this study, at 6 months the mean EPDS score in EPDS-negative women was 5.4 (s.d.=4.5) for CAU women and 4.8 (s.d.=4.2) for IG women, a difference of -0.6 (95% CI -0.9 to -0.1, $p=0.007$), which was statistically significant ($p=0.007$). This difference remained statistically significant after adjusting for the same 6-week and baseline covariates ($p=0.013$) (Table 3). A significant difference ($p=0.020$) in the 6-month CORE-OM total score favoring the IG women (mean=0.38) over CAU women at 6 months was also found (mean= 0.43) with other outcome comparisons (SF-12 and State Anxiety) being not significant. A comparison of outcomes for IG women in the two randomly allocated
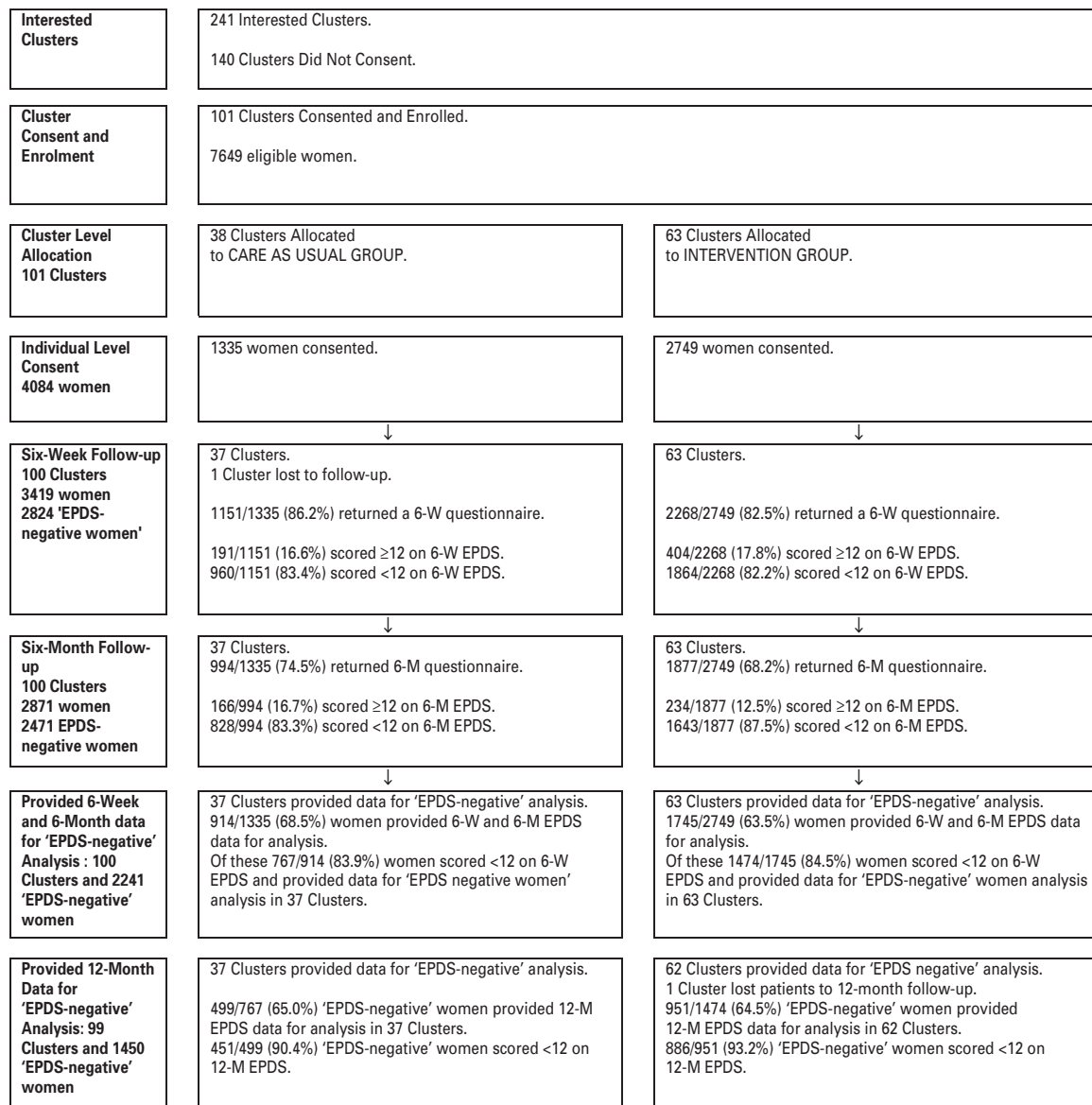
Interested Clusters	241 Interested Clusters. 140 Clusters Did Not Consent.	
Cluster Consent and Enrolment	101 Clusters Consented and Enrolled. 7649 eligible women.	
Cluster Level Allocation 101 Clusters	38 Clusters Allocated to CARE AS USUAL GROUP.	63 Clusters Allocated to INTERVENTION GROUP.
Individual Level Consent 4084 women	1335 women consented.	2749 women consented.
Six-Week Follow-up 100 Clusters 3419 women 2824 'EPDS-negative women'	37 Clusters. 1 Cluster lost to follow-up. 1151/1335 (86.2%) returned a 6-W questionnaire. 191/1151 (16.6%) scored ≥12 on 6-W EPDS. 960/1151 (83.4%) scored <12 on 6-W EPDS.	63 Clusters. 2268/2749 (82.5%) returned a 6-W questionnaire. 404/2268 (17.8%) scored ≥12 on 6-W EPDS. 1864/2268 (82.2%) scored <12 on 6-W EPDS.
Six-Month Follow-up 100 Clusters 2871 women 2471 EPDS-negative women	37 Clusters. 994/1335 (74.5%) returned 6-M questionnaire. 166/994 (16.7%) scored ≥12 on 6-M EPDS. 828/994 (83.3%) scored <12 on 6-M EPDS.	63 Clusters. 1877/2749 (68.2%) returned 6-M questionnaire. 234/1877 (12.5%) scored ≥12 on 6-M EPDS. 1643/1877 (87.5%) scored <12 on 6-M EPDS.
Provided 6-Week and 6-Month data for 'EPDS-negative' Analysis : 100 Clusters and 2241 'EPDS-negative' women	37 Clusters provided data for 'EPDS-negative' analysis. 914/1335 (68.5%) women provided 6-W and 6-M EPDS data for analysis. Of these 767/914 (83.9%) women scored <12 on 6-W EPDS and provided data for 'EPDS negative women' analysis in 37 Clusters.	63 Clusters provided data for 'EPDS-negative' analysis. 1745/2749 (63.5%) women provided 6-W and 6-M EPDS data for analysis. Of these 1474/1745 (84.5%) women scored <12 on 6-W EPDS and provided data for 'EPDS-negative' women analysis in 63 Clusters.
Provided 12-Month Data for 'EPDS-negative' Analysis: 99 Clusters and 1450 'EPDS-negative' women	37 Clusters provided data for 'EPDS-negative' analysis. 499/767 (65.0%) 'EPDS-negative' women provided 12-M EPDS data for analysis in 37 Clusters. 451/499 (90.4%) 'EPDS-negative' women scored <12 on 12-M EPDS.	62 Clusters provided data for 'EPDS negative' analysis. 1 Cluster lost patients to 12-month follow-up. 951/1474 (64.5%) 'EPDS-negative' women provided 12-M EPDS data for analysis in 62 Clusters. 886/951 (93.2%) 'EPDS-negative' women scored <12 on 12-M EPDS.

Fig. 1. Trial profile of clusters and participants in the intervention group (IG) and the care as usual (CAU) group for all women and Edinburgh Postnatal Depression Scale (EPDS) negative women.

training groups (CBA or PCA) also showed no difference in effectiveness: compared to the control (CAU) EPDS score, at 6 months the mean EPDS score in the CBA group was 4.7 ($n=708$, S.D$=4.2$); and in the PCA group it was 4.9 ($n=766$, S.D$=4.3$).

EPDS-negative women in the face-to-face clinical assessment group ($n=746$) had a lower mean EPDS score at 6 months than women who only received a postal EPDS ($n=728$; difference 0.68, 95% CI 0.25–1.20, $p=0.002$); this was of borderline statistical significance following adjustment ($p=0.054$). When comparing the binary 6-month EPDS outcomes (EPDS score $\geqslant 12$), the differences were not statistically significant (unadjusted $p=0.126$; adjusted $p=0.301$).

Because of trial funding limitations, fewer postal questionnaires were sent to women and thus returned at 12 and 18 months ($n=807$) than at 6 months ($n=2659$). In addition to the 6-week and 6-month returns, 1450 women also completed and returned a postal follow-up at 12 months and 806 at 18 months. At the 12-month follow-up, 9.6% (48/499) of women in the CAU group and 6.8% (65/951) of women in the IG had an EPDS score $\geqslant 12$ and an unadjusted OR of 0.68 (95% CI 0.44–1.05, $p=0.079$). After adjusting for covariates, the OR was 0.74 (95% CI 0.48–1.14, $p=0.167$). Using a multiple imputation method for computing missing 12-month EPDS scores increased the sample size from 1450 to 2786, and the OR using

Table 1. *Baseline and 6 weeks characteristics of 'EPDS-negative women' by group (n = 2241)*

	Care as usual (CAU) (n = 767)	Intervention group (IG) (n = 1474)
Age (years), *n* (mean), S.D.	766 (32.2), 5.0	1474 (31.4), 5.0
Age when first child born (years), *n* (mean), S.D.	744 (29.2), 5.2	1449 (28.2), 5.2
Number of other children, *n* (mean), S.D.	767 (0.7), 0.8	1474 (0.7), 0.8
EPDS score at 6 weeks, *n* (mean), S.D.	767 (5.2), 3.2	1474 (5.0), 3.1
SF-12 PCS score at 6 weeks, *n* (mean), S.D.	745 (50.9), 8.1	1454 (51.6), 7.7
SF-12 MCS score at 6 weeks, *n* (mean), S.D.	745 (45.3), 7.1	1454 (45.4), 7.0
CORE-OM score at 6 weeks, *n* (mean), S.D.	760 (0.4), 0.3	1466 (0.4), 0.3
Lives alone with child/ren, *n* (%)		
No	735 (96.8)	1430 (97.3)
Yes	24 (3.2)	39 (2.7)
Total	759 (100.0)	1469 (100.0)
Previous history of PND, *n* (%)		
No/first child	709 (92.8)	1364 (92.7)
Yes	55 (7.2)	107 (7.3)
Total	764 (100.0)	1471 (100.0)
Previous history of major life events, *n* (%)		
No	466 (61.2)	903 (61.6)
Yes	295 (38.8)	563 (38.4)
Total	761 (100.0)	1466 (100.0)
Number of other children, *n* (%)		
0	370 (48.2)	661 (44.8)
1	305 (39.8)	601 (40.8)
2	71 (9.3)	166 (11.3)
3	13 (1.7)	33 (2.2)
4	5 (0.7)	3 (0.2)
5	3 (0.4)	10 (0.7)
Total	767 (100.0)	1474 (100.0)
Ethnicity, *n* (%)		
Non-white	29 (3.8)	45 (3.1)
White	738 (96.2)	1429 (96.9)
Total	767 (100.0)	1474 (100.0)
English first language, *n* (%)		
No	26 (3.4)	41 (2.8)
Yes	741 (96.6)	1433 (97.2)
Total	767 (100.0)	1474 (100.0)

EPDS, Edinburgh Postnatal Depression Scale; SF-12, 12-item Short Form Health Survey; PCS, physical component summary scale; MCS, mental component summary scale; CORE-OM, Clinical Outcomes in Routine Evaluation – Outcome Measure; PND, postnatal depression, S.D., standard deviation.

Better health is represented by a lower score in CORE-OM and EPDS and a higher score in the SF-12 and SF-6D.

the imputed data, with adjustment for covariates, was 0.89 (95% CI 0.62–1.28).

Group effects and time effects were found for the EPDS (continuous) outcome. However, after adjusting for baseline covariates, no significant group by time statistical interactions were found on any of the outcome measures [EPDS binary, EPDS continuous, CORE-OM, SF-12 mental component summary scale (MCS), SF-12 physical component summary scale (PCS), State-Anxiety]. Group effects, but not time effects, were found for the CORE-OM and EPDS binary outcomes. A significant time effect only was observed for the two SF-12 outcomes. No group or time effects were observed for the State Anxiety outcome.

Table 2. *Primary outcome: proportion of EPDS-negative women with EPDS score* ⩾12 *at 6 months, by intervention group (IG) or care as usual (CAU) group, comparing EPDS subthreshold and lowest EPDS score subsamples defined according to EPDS score at 6 weeks (n = 2241)*

EPDS score at 6 weeks	6-month EPDS score	CAU n (%)	IG n (%)	OR$_{IG/CAU}$*	95% CI for OR
Lowest score (EPDS ⩽5)	<12	386 (94.8)	802 (96.0)		
	⩾12	21 (5.2)	33 (4.0)	0.76	0.42–1.36
		407 (100.0)	835 (100.0)		
Subthreshold (5<EPDS<12)	<12	298 (82.8)	559 (87.5)		
	⩾12	62 (17.2)	80 (12.5%)	0.68	0.48–0.97
		360 (100.0)	639 (100)		
EPDS<12	<12	684 (89.2)	1361 (92.3)		
	⩾12	83 (10.8)	113 (7.7)	0.68	0.50–0.93
		767 (100.0)	1474 (100.0)		

EPDS, Edinburgh Postnatal Depression Scale; OR, odds ratio; CI, confidence interval.

p value for interaction term between lowest score (EPDS ⩽5) and subthreshold (5<EPDS<12) groups = 0.782.

* *p* value for interaction term adjusted for lives alone (yes/no), history of postnatal depression (PND) (yes/no), any life events (yes/no) = 0.749.

Discussion

Our analysis has shown that women who scored below the EPDS threshold at 6 weeks after childbirth were less likely to score above the EPDS threshold at 6 months after childbirth if their HV had undergone training in identifying depressive symptoms, developing therapeutic relationships and providing psychological approaches to managing depression. Furthermore, this effect applied equally across women at all initial levels of risk, not just those close to the above threshold, thus suggesting a universal effect. This effect did not diminish over later assessments at 12 and 18 months following childbirth. How plausible is this effect and can it be explained?

We found that the training was effective in the separately reported treatment trial in EPDS positive (EPDS score ⩾12) women (Morrell *et al*. 2009*a*), although 61% agreed to, and only 28% of all EPDS positive IG women providing outcome data had completed, HV sessions. In the present study, approximately 1% of IG EPDS-negative women were offered and completed at least one session. So we need to ask why and how was HV training universally effective most notably in the EPDS-negative women? Cross-contamination effects are likely to have been minimized by clustering HVs. One previous primary care trial, using a similarly powered clustered design, in which midwives were trained in a protocol to focus on specific physical and mental health needs of mothers after childbirth, also reported significantly lower EPDS scores at 4 months follow-up in women in IG practices (MacArthur *et al*. 2002). However, there were insufficient baseline data available to identify women with greater depression severity to determine whether effectiveness was confined to a subgroup of likely cases of depression or was preventive across the whole range of baseline severity.

What seems to matter is the difference between being registered with a practice that has adopted the experimental *versus* the CAU policy. This involves differences in practice environment and culture. More specifically, receiving training and developing the skill to assess mental health and provide psychologically informed sessions is likely to affect the practitioners themselves through increased confidence. It creates a focus of care on the mother's psychological well-being and not just on the physical welfare of the child. Women may benefit from the HVs' enhanced communication skills and knowing that emotional issues are open for discussion if needed, without the possibly stigmatizing requirement to refer to a different profession or service. We were able to show evidence in the women who were below threshold at 6 weeks of a positive effect of being randomly assigned to receive a face-to-face EPDS administration and clinical assessment, when compared with a research office postally administered EPDS questionnaire only. As reported separately, empirical support for change in HV activity was shown in the resource use logs of these lower-risk women (Brugha *et al*. 2009). As the EPDS score increased, in the IGs, a clear increase in the

Table 3. *Six-month secondary outcomes for the 'EPDS-negative sample' (n = 2241), care as usual (CAU) versus intervention group (IG), unadjusted and adjusted*

	CAU			IG			Unadjusted			Adjusted[a]		
	n	Mean	S.D.	n	Mean	S.D.	Difference	95% CI	p	Difference	95% CI	p
EPDS at 6 months	767	5.4	4.5	1474	4.8	4.2	−0.6	−1.1 to −0.2	0.007	−0.5	−0.9 to −0.1	0.013
SF-12 PCS at 6 months	743	54.5	6.3	1431	55.0	5.8	0.5	−0.1 to 1.0	0.093	0.3	−0.2 to 0.9	0.198
SF-12 MCS at 6 months	743	49.5	9.1	1431	50.1	8.7	0.7	−0.3 to 1.7	0.165	0.6	−0.3 to 1.5	0.174
CORE-OM total score at 6 months	760	0.43	0.43	1467	0.38	0.39	−0.05	−0.10 to −0.1	0.020	−0.04	−0.08 to −0.00	0.035
State Anxiety at 6 months	722	32.3	10.3	1380	31.6	9.9	−0.7	−1.9 to 0.5	0.231	−0.7	−1.8 to 0.4	0.236

EPDS, Edinburgh Postnatal Depression Scale; SF-12, 12-item Short Form Health Survey; PCS, physical component summary scale; MCS, mental component summary scale; CORE-OM, Clinical Outcomes in Routine Evaluation – Outcome Measure; S.D., standard deviation; CI, confidence interval.

Better health represented by lower score in CORE-OM, EPDS, STATE Anxiety. Better health represented by higher score in SF-12.

[a] Adjusted for 6-week EPDS score, lives alone, history of postnatal depression (PND) and any life events.

number of visits was seen (Brugha *et al.* 2009), which seemed to be due to a greater recognition of the need for professional input relating to PND by IG teams; whereas concurrently, there were fewer visits by intervention HVs relating to the mothers with lower EPDS scores, presumably because visits are judged to be unnecessary.

The question remains unanswered: how did intervention cluster women benefit without formal psychological therapy sessions? The purpose of training was also to establish a high-quality, warm, therapeutic relationship with the women (Morrell *et al.* 2009 *b*). Concepts of social support have been suggested to explain the common underlying mechanisms in the effectiveness of cognitive behavioral, interpersonal and other therapies (Brugha, 1995). Tailoring to need and availability of support from a trained professional could potentially influence perception of available support, which may be the key parameter, rather than objective support. The value of tailoring and flexible care has certainly been suggested by MacArthur *et al.* (2002), but in that study women were provided care from midwives and followed only to 4 months postnatally.

Psychological interventions focusing on interpersonal issues during pregnancy have also shown some promise (Spinelli & Endicott, 2003; Zlotnick *et al.* 2006). This result contrasts with two previous ineffective prevention trials conducted in the same population by the present authors (T.S.B., C.J.M.) that employed external (research) therapists (Brugha *et al.* 2000; Morrell *et al.* 2000). Thus, individually delivered psychosocial interventions by professionals may be more effective postnatally (Dennis, 2005) when integrated into routine visits. For example, one other randomized trial of enhanced and personalized midwife-managed care in the early postnatal period also reported better EPDS outcomes at 7 weeks postnatally (Shields *et al.* 1997). The similarity of both sets of research findings runs counter to the suggestion that our result is a chance, fortuitous finding.

Study limitations need to be considered given the potential public health significance of the results. A potential limitation of the study is that the EPDS subgroup by IG interaction tests have limited power to detect anything other than large interactions (Montgomery *et al.* 2003). The CI for the OR for the interaction effect at 6 months after adjustment for covariates (0.44–1.87) was relatively wide, so we cannot exclude such an effect, although the point estimate of the OR of 0.91 suggests this is unlikely. However, the interaction was not of primary interest and the trial was adequately powered to detect a reasonable target intervention effect, equivalent to an OR of 0.63. A full clinical interview was not performed on all CAU and

IG EPDS-negative women and our primary outcome was thus not clinically assessed depression caseness. However, the EPDS is a well-established, reliable indicator of likelihood of clinical depression (Gaynes *et al*. 2005). Bias may arise with cluster randomization because individual recruitment occurs after randomization (Puffer *et al*. 2005). According to Table 1, the profiles of IG and CAU women were remarkably balanced but, nevertheless, our analysis adjusted for key predictors of depression outcome. A planned replication and extended generalization of findings will be needed in other populations.

Rose (1992) argued that more people would benefit from a universal intervention. In our treatment trial (Morrell *et al*. 2009*b*) we found that 31 (11.4%) of 271 EPDS positive women benefited from intervention. In this prevention analysis, in EPDS-negative women, 46 (3.1%) of 1474 women benefited. Accordingly, the new intervention policy benefited far more women in the prevention (EPDS-negative) than in the treatment (EPDS positive) groups. Our finding that mental health benefits did not diminish over 18 months of follow-up contrasts with the short-term effectiveness of depression interventions in general and may reflect the ongoing extended contact that HVs often have with postnatal women. Additional analyses from this trial showed the probability that the intervention is cost-effective at various 'threshold values' of a quality-adjusted life year (QALY). In the range of QALY values between £20 000 and £30 000, the probability of the intervention being cost-effective is over 99% (Brugha *et al*. 2009). Thus, costs were lower and outcomes better in IG compared to CAU women, a concept often referred to by economists as 'dominance'.

Until now, governments and policy makers have held back from investing significantly in depression prevention programs, arguably because of lack of convincing evidence for policy change and cost. To date, there have been no large-scale trials testing whether universal prevention effects have occurred across a whole population. There is now new evidence for clinically significant, useful and persistent reductions in the prevalence of depression in a key part of the population, women following childbirth, which requires further independent evaluation.

Acknowledgments

Declaration of Interest

None.

References

Barkham M, Margison F, Leach C, Lucock M, Mellor-Clark J, Evans C, Benson L, Connell J, Audin K, McGrath G (2001). Service profiling and outcomes benchmarking using the CORE-OM: towards practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology* **69**, 184–196.

Brugha T, Bebbington P, Tennant C, Hurry J (1985). The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. *Psychological Medicine* **15**, 189–194.

Brugha TS (1995). Social support and psychiatric disorder: recommendations for clinical practice and research. In *Social Support and Psychiatric Disorder: Research Findings and Guidelines for Clinical Practice*, 1st edn (ed. T. S. Brugha), pp. 295–334. Cambridge University Press: Cambridge.

Brugha TS, Bebbington PE, Singleton N, Melzer D, Jenkins R, Lewis G, Farrell M, Bhugra D, Lee A, Meltzer H (2004). Trends in service use and treatment for mental disorders in adults throughout Great Britain. *British Journal of Psychiatry* **185**, 378–384.

Brugha TS, Morrell CJ, Slade P, Dixon S (2009). Prevention of postnatal depression: cost-effectiveness of health visitor training; results from a cluster randomised controlled trial in primary care. Ninth Workshop on Costs and Assessment in Psychiatry. Quality and Outcomes in Mental Health Policy and Economics. *Journal of Mental Health Policy and Economics* **12** (Suppl. 1), S8.

Brugha TS, Wheatley S, Taub NA, Culverwell A, Friedman T, Kirwan PH, Jones DR, Shapiro DA (2000). Pragmatic randomized trial of antenatal intervention to prevent postnatal depression by reducing psychosocial risk factors. *Psychological Medicine* **30**, 1273–1281.

Campbell MK, Elbourne DR, Altman DG (2004). CONSORT statement: extension to cluster randomised trials. *British Medical Journal* (*Clinical Research Edition*) **328**, 702–708.

Clarke GN, Hornbrook M, Lynch F, Polen M, Gale J, Beardslee W, O'Connor E, Seeley J (2001). A randomized trial of a group cognitive intervention for preventing depression in adolescent offspring of depressed parents. *Archives of General Psychiatry* **58**, 1127–1134.

Cox JL, Holden JM, Sagovsky R (1987). Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry* **150**, 782–786.

Cuijpers P, Smit F (2004). Subthreshold depression as a risk indicator for major depressive disorder: a systematic review of prospective studies. *Acta Psychiatrica Scandinavica* **109**, 325–331.

Cuijpers P, van Straten A, Smit F (2005). Preventing the incidence of new cases of mental disorders: a meta-analytic review. *Journal of Nervous and Mental Disease* **193**, 119–125.

Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A (2008). Preventing the onset of depressive

disorders: a meta-analytic review of psychological interventions. *American Journal of Psychiatry* **165**, 1272–1280.

Das-Munshi J, Goldberg D, Bebbington PE, Bhugra DK, Brugha TS, Dewey ME, Jenkins R, Stewart R, Prince M (2008). Public health significance of mixed anxiety and depression: beyond current classification. *British Journal of Psychiatry* **192**, 171–177.

Dennis CL (2005). Psychosocial and psychological interventions for prevention of postnatal depression: systematic review. *British Medical Journal (Clinical Research Edition)* **331**, 15.

Garber J, Clarke GN, Weersing VR, Beardslee WR, Brent DA, Gladstone TR, DeBar LL, Lynch FL, D'Angelo E, Hollon SD, Shamseddeen W, Iyengar S (2009). Prevention of depression in at-risk adolescents: a randomized controlled trial. *Journal of the American Medical Association* **301**, 2215–2224.

Gaynes BN, Gavin N, Meltzer-Brody S, Lohr KN, Swinson T, Gartlehner G, Brody S, Miller WC (2005). *Perinatal Depression: Prevalence, Screening Accuracy, and Screening Outcomes*. Evidence Report/Technology Assessment No. 119, pp. 1–8 (Rockville, MD: Agency for Healthcare Research and Quality).

Gibson J, Kenzie-McHarg K, Shakespeare J, Price J, Gray R (2009). A systematic review of studies validating the Edinburgh Postnatal Depression Scale in antepartum and postpartum women. *Acta Psychiatrica Scandinavica* **119**, 350–364.

Gordon Jr. RS (1983). An operational classification of disease prevention. *Public Health Reports* **98**, 107–109.

Hewitt C, Gilbody S, Brealey S, Paulden M, Palmer S, Mann R, Green J, Morrell J, Barkham M, Light K, Richards D (2009). Methods to identify postnatal depression in primary care: an integrated evidence synthesis and value of information analysis. *Health Technology Assessment* **13**, 1–230.

Judd LL, Schettler PJ, Akiskal HS (2002). The prevalence, clinical relevance, and public health significance of subthreshold depressions. *Psychiatric Clinics of North America* **25**, 685–698.

Kessler RC, Demler O, Frank RG, Olfson M, Pincus HA, Walters EE, Wang P, Wells KB, Zaslavsky AM (2005). Prevalence and treatment of mental disorders, 1990 to 2003. *New England Journal of Medicine* **352**, 2515–2523.

MacArthur C, Winter HR, Bick DE, Knowles H, Lilford R, Henderson C, Lancashire RJ, Braunholtz DA, Gee H (2002). Effects of redesigned community postnatal care on women's health 4 months after birth: a cluster randomised controlled trial. *Lancet* **359**, 378–385.

Montgomery AA, Peters TJ, Little P (2003). Design, analysis and presentation of factorial randomised controlled trials. *BMC Medical Research Methodology* **3**, 26.

Morrell CJ, Slade P, Warner R, Paley G, Dixon S, Walters SJ, Brugha T, Barkham M, Parry GJ, Nicholl J (2009a). Clinical effectiveness of health visitor training in psychologically informed approaches for depression in postnatal women: pragmatic cluster randomised trial in primary care. *British Medical Journal (Clinical Research Edition)* **338**, a3045.

Morrell CJ, Spiby H, Stewart P, Walters S, Morgan A (2000). Costs and benefits of community postnatal support workers: a randomised controlled trial. *Health Technology Assessment* **4**, 1–88.

Morrell CJ, Warner R, Slade P, Dixon S, Walters S, Paley G, Brugha T (2009b). Psychological interventions for postnatal depression: cluster randomised trial and economic evaluation. The PONDER trial. *Health Technology Assessment* **13**, 1–176.

Mrazek PJ, Haggerty RJ (1994). *Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research*. National Academy of Medicine: Washington, DC.

O'Hara MW, Swain AM (1996). Rates and risk of postpartum depression: a meta-analysis. *International Review of Psychiatry* **8**, 37–54.

Puffer S, Torgerson DJ, Watson J (2005). Cluster randomized controlled trials. *Journal of Evaluation in Clinical Practice* **11**, 479–483.

Rose G (2001). Sick individuals and sick populations. *International Journal of Epidemiology* **30**, 427–432.

Rose GA (1992). *The Strategy of Preventive Medicine*. Oxford University Press: Oxford.

Shields N, Reid M, Cheyne H, Holmes A, McGinley M, Turnbull D, Smith LN (1997). Impact of midwife-managed care in the postnatal period: an exploration of psychosocial outcomes. *Journal of Reproductive and Infant Psychology* **15**, 91–108.

Spielberger C, Gorsuch R, Lushene R (1970). *Manual for the State-Trait Anxiety Inventory (Self-Evaluation Questionnaire)*. Consulting Psychologists Press: Palo Alto, CA.

Spinelli MG, Endicott J (2003). Controlled clinical trial of interpersonal psychotherapy versus parenting education program for depressed pregnant women. *American Journal of Psychiatry* **160**, 555–562.

StataCorp (2009). STATA Statistical Software, Release 11.0. Stata Corporation: College Station, TX.

Tyrer P (2001). The case for cothymia: mixed anxiety and depression as a single diagnosis. *British Journal of Psychiatry* **179**, 191–193.

Ware JE, Kosinski M, Gandek B (1995). *How to Score the SF-12 Physical and Mental Health Summary Scales*. Health Institute: Boston.

Zlotnick C, Miller IW, Pearlstein T, Howard M, Sweeney P (2006). A preventive intervention for pregnant women on public assistance at risk for postpartum depression. *American Journal of Psychiatry* **163**, 1443–1445.