Appendices

Appendix A: Baseline characteristics by eligible women who did and did not give baseline consent.

	Gave baseline consent		
	No (n=53)	Yes (n=280)	p-value [1]
Recruitment site (n [%]) GMC hospital Asilo hospital Bicholim PHC Sanquelim PHC	 31 (58%) 20 (38%) 2 (4%) 0	 141 (50%) 125 (45%) 13 (5%) 1 (<1%)	0·66
Age, years (mean [SD]) Missing	26.2 (4.5) 2 (4%)	25.2 (4.6) 0	0·14
Marital status (n [%]) Married Divorced or separated Missing	 51 (96%) 0 2 (4%)	 279 (100%) 1 (<1%) 0	1·00
Occupation (n [%]) Does not work Manual worker Non-manual worker Missing	 43 (81%) 4 (8%) 4 (8%) 2 (4%)	 237 (85%) 30 (11%) 13 (5%) 0	0·47
Education status (n [%]) No formal education Completed primary Completed secondary Completed higher secondary Graduate/above Missing	 5 (9%) 16 (30%) 19 (36%) 7 (13%) 4 (8%) 2 (4%)	 34 (12%) 120 (43%) 90 (32%) 25 (9%) 11 (4%) 0	0·31
Chronicity of depression category (n [%]) <12 weeks ≥12 weeks	 29 (55%) 24 (45%)	 173 (62%) 107 (38%)	0·33
Median PHQ-9 score (SD; IQR)	12 (3·6; 11-14)	12 (3·4; 11-15)	0·32
PHQ-9 score category (n [%]) 10-14 (moderate) 15-19 (moderately severe) 20-27 (severe)	 43 (81%) 6 (11%) 4 (8%)	 197 (40%) 67 (24%) 16 (6%)	0·11
PHQ question 10 (n [%]) [2] Not difficult at all Somewhat difficult Very difficult Extremely difficult	 15 (28%) 28 (53%) 4 (8%) 6 (11%)	 43 (15%) 165 (59%) 27 (10%) 45 (16%)	0·17

Note: A subset of variables are shown, since not all data were systematically collected in the women who refused baseline consent. [1] P-values were calculated using t-tests for continuous variables where the mean is reported, Wilcoxon rank-

sum tests for continuous variables where the median is reported, X^2 tests for categorical variables where there are at least 5 observations in each category, and Fisher's exact tests for categorical variables where there are less than 5 observations in any category. Women with missing data for a given variable were omitted for the test on that variable. [2] Question: If you checked off any problems (PHQ questions 1-9), how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people? MSPSS=Multidimensional Scale of Perceived Social Support. PHQ-9=Patient Health Questionnaire 9. SD=Standard Deviation. IQR=Interquartile range. GMC=Goa Medical College. PHC=Primary Health Care.

Appendix B: Baseline characteristics of women who do versus do not have 6 month outcome data for the primary endpoint analysis.

	Have 6 month outcome data for the primary endpoint analysis		
	No (n=29) [1]	Yes (n=251)	p-value [2]
Recruitment site (n [%]) GMC hospital Asilo hospital Bicholim PHC Sanquelim PHC	 16 (55%) 12 (41%) 1 (3%) 0	 125 (50%) 113 (45%) 12 (5%) 1 (<1%)	0·90
Residence (n [%]) Rural Urban	 24 (83%) 5 (17%)	 182 (73%) 69 (27%)	0·24
Age (years) (mean [SD])	25·1 (3·8)	25·2 (4·7)	0·90
Marital status (n [%]) Married Divorced/separated	 29 (100%) 0	 250 (100%) 1 (<1%)	1·00
Occupation (of TP) (n [%]) Does not work Manual worker Non-manual worker	 21 (72%) 6 (21%) 2 (7%)	 216 (86%) 24 (10%) 11 (4%)	0·11
Education status (n [%]) No formal education Completed primary Completed secondary Completed higher secondary Graduate/above	 2 (7%) 9 (31%) 12 (41%) 5 (17%) 1 (3%)	 32 (13%) 111 (44%) 78 (31%) 20 (8%) 10 (4%)	0·25
Patient's expectation of usefulness of counselling (n [%]) Not useful A little useful Somewhat useful Moderately useful Very useful	 0 9 (31%) 4 (14%) 5 (17%) 11 (38%)	 1 (<1%) 46 (18%) 50 (20%) 53 (21%) 101 (40%)	0·51
Chronicity of depression category (n [%]) <12 weeks ≥12 weeks	 19 (66%) 10 (34%)	 154 (61%) 97 (39%)	0·66
Median PHQ-9 score [IQR]	12 (10-15)	12 (11-15)	0·58
PHQ-9 score category (n [%]) 10-14 (moderate) 15-19 (moderately severe) 20-27 (severe)	 21 (72%) 7 (24%) 1 (3%)	 176 (70%) 60 (24%) 15 (6%)	1·00
PHQ question 10 (n [%]) [3] Not difficult at all	 4 (13%)	 39 (16%)	0·34

	Have 6 month outcome data for the primary endpoint analysis		
	No (n=29) [1]	Yes (n=251)	p-value [2]
Somewhat difficult	14 (48%)	151 (60%)	
Very difficult	3 (10%)	24 (10%)	
Extremely difficult	8 (28%)	37 (15%)	
MSPSS score (mean [SD])	5·5 (0·9)	5·3 (1·1)	0·38
Parity (n [%])			0·79
Primiparous	13 (45%)	106 (42%)	
Multiparous	16 (55%)	145 (58%)	
Previous non-live birth (n [%])			0·70
None	28 (97%)	233 (93%)	
One/more	1 (3%)	18 (7%)	
Domestic violence, any (n [%])			0·78
No	26 (90%)	217 (86%)	
Yes	3 (10%)	34 (14%)	
Time between screening and birth of child, months (mean [SD])	3·6 (1.46)	3·9 (1·66)	0·47
Missing [4]	11 (38%)	0	

[1] Mainly lost for the 6 month visit; see Figure 1. [2] P-values were calculated using t-tests for continuous variables where the mean is reported, Wilcoxon rank-sum tests for continuous variables where the median is reported, X^2 tests for categorical variables where there are at least 5 observations in each category, and Fisher's exact tests for categorical variables where there are less than 5 observations in any category. Women with missing data for a given variable were omitted for the test on that variable. [3] Question: If you checked off any problems (PHQ questions 1-9), how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people? [4] Women who did not attend any follow up visits have missing information relating to the date of birth of the child. MSPSS=Multidimensional Scale of Perceived Social Support. PHQ-9=Patient Health Questionnaire 9. SD=Standard Deviation. IQR=Interquartile range. GMC=Goa Medical College. PHC=Primary Health Care.

	Participants seen within the protocol-defined visit window (6-8 months)?		
	No (N=19) [1]	Yes (N=232)	p-value [2]
Recruitment site (n [%])			0·57
GMC hospital	8 (42%)	117 (50%)	
Asilo hospital	11 (58%)	102 (44%)	
Bicholim PHC	0	12 (5%)	
Sanquelim PHC	0	1 (<1%)	
Residence (n [%])			0·60
Rural	15 (79%)	167 (72%)	
Urban	4 (21%)	65 (28%)	
Age (years) (mean [SD])	24·8 (4·7)	25·2 (4·7)	0·74
Marital status (n [%])			1·00
Married	19 (100%)	231 (100%)	
Divorced/separated	0	1 (<1%)	
Occupation (n [%])			1·00
Does not work	17 (89%)	199 (86%)	
Manual worker	2 (11%)	22 (9%)	
Non-manual worker	0	11 (5%)	
Education status (n [%])			0·62
None	3 (16%)	29 (13%)	
Primary	8 (42%)	103 (44%)	
Secondary	8 (42%)	70 (30%)	
Higher secondary	0	20 (9%)	
Graduate/above	0	10 (4%)	
Patient's expectation of usefulness of counselling (n [%])			0·61
Not useful	0	1 (<1%)	
A little useful	4 (21%)	42 (18%)	
Somewhat useful	3 (16%)	47 (20%)	
Moderately useful	2 (11%)	51 (22%)	
Very useful	10 (53%)	91 (39%)	
Chronicity of depression, weeks (median [IQR])	8 (4-16)	8 (3-16)	0·41
Chronicity of depression category (n [%])			0·42
<12 weeks	10 (53%)	144 (62%)	
≥12 weeks	9 (47%)	88 (38%)	
Median PHQ-9 score [IQR]	12 (11-15)	12 (11-15)	0·80
PHQ-9 score category (n [%])			0·92
10-14 (moderate)	13 (68%)	163 (70%)	
15-19 (moderately severe)	5 (26%)	55 (24%)	
20-27 (severe)	1 (5%)	14 (6%)	

	Participants seen within the protocol-defined visit window (6-8 months)?		
	No (N=19) [1]	Yes (N=232)	p-value [2]
PHQ question 10 (n [%]) [3]			0·06
Not difficult at all	3 (16%)	36 (16%)	
Somewhat difficult	7 (37%)	144 (62%)	
Very difficult	4 (21%)	20 (9%)	
Extremely difficult	5 (26%)	32 (14%)	
MSPSS score (mean [SD])	5·3 (0·7)	5·3 (1·2)	0·89
Parity (n [%])			0·34
Primiparous	10 (53%)	96 (41%)	
Multiparous	9 (47%)	146 (59%)	
Previous non-live birth (n [%])			0·15
None	16 (84%)	217 (94%)	
One/more	3 (16%)	15 (6%)	
Domestic violence, any (n [%])			0·48
No	18 (95%)	199 (86%)	
Yes	1 (5%)	33 (14%)	
Time between screening and birth of child, months (mean [SD])	3·3 (1·97)	4·0 (1·63)	0·11

[1] All seen at >5.5 months; mainly just one or two days early. [2] P-values were calculated using t-tests for continuous variables where the mean is reported, Wilcoxon rank-sum tests for continuous variables where the median is reported, X^2 tests for categorical variables where there are at least 5 observations in each category, and Fisher's exact tests for categorical variables where there are less than 5 observations in any category. Women with missing data for a given variable were omitted for the test on that variable. [3] If you checked off any problems (PHQ questions 1-9), how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people? MSPSS=Multidimensional Scale of Perceived Social Support. PHQ-9=Patient Health Questionnaire 9. SD=Standard Deviation. IQR=Interquartile range. GMC=Goa Medical College. PHC=Primary Health Care.

Appendix D: Remission (PHQ-9<5) at 6 months by potential effect modifiers.

	N		N with remission (%)		PR (95% CI) [1]	OR (95% CI) [1]	P value for effect modification
	THPP plus EUC	EUC alone	THPP plus EUC	EUC alone			
Age, years							
18-24	61	72	50 (82%)	43 (60%)	1·33 (1·04,1·62)	2·91 (1·28,6·62)	0·15
≥25	61	57	39 (64%)	34 (60%)	1·10 (0·77,1·42)	1·27 (0·59,2·74)	
Chronicity of depression							
<12 weeks	73	81	59 (81%)	47 (58%)	1·39 (1·08,1·70)	3·04 (1·43,6·44)	0·04
≥12 weeks	49	48	30 (61%)	30 (63%)	0·96 (0·67,1·26)	0·90 (0·39,2·11)	
Baseline PHQ-9 severity							
10-14	86	90	63 (73%)	56 (62%)	1·18 (0·93,1·43)	1·69 (0·87,3·26)	0·72
≥15	36	39	26 (72%)	21 (54%)	1·30 (0·85,1·75)	2·11 (0·79,5·60)	
Treatment expectations							
None/somewhat	42	55	29 (69%)	30 (55%)	1·24 (0·85,1·62)	1·80 (0·76,4·25)	0·96
Moderate/very useful	80	74	60 (75%)	47 (64%)	1·20 (0·94,1·46)	1·82 (0·90,3·69)	

[1] By logistic regression, adjusted for recruitment centre, baseline PHQ-9 score, residence, treatment expectations, education and chronicity (see Methods). SD=Standard Deviation, PR=Prevalence Ratio, PHQ=Patient Health Questionnaire, THPP=Thinking Healthy Programme Peer-delivered, EUC=Enhanced Usual Care, CI=confidence interval.

Appendix E: PHQ-9 score at 6 months by potential effect modifiers.

	N		Mean (SD)		SMD (95% CI) [1]	Mean difference (95% CI) [1]	P value for effect modification
	THPP plus EUC	EUC alone	THPP plus EUC	EUC alone			
Age, years							
18-24	61	72	2·7 (4·5)	4·9 (5·7)	-0·40 (-0·73,-0·06)	-1·91 (-3·60,-0·22)	0·09
≥25	61	57	4·2 (4·4)	3·9 (4·2)	0·04 (-0·32,0·40)	0·21 (-1·58,1·99)	
Chronicity of depression							
<12 weeks	73	81	3·0 (4·6)	4·6 (5·1)	-0·29 (-0·60,0·03)	-1·40 (-2·98,0·18)	0·30
≥12 weeks	49	48	4·1 (4·4)	4·3 (5·2)	-0·01 (-0·41,0·39)	-0·06 (-2·04,1·92)	
Baseline PHQ-9 severity							
10-14	86	90	3·7 (5·0)	4·3 (5·2)	-0·11 (-0·41,0·18)	-0·56 (-2·03,0·91)	0·43
≥15	36	39	3·0 (3·2)	4·8 (5·0)	-0·34 (-0·79,0·11)	-1·64 (-3·90,0·62)	
Treatment expectations							
None/somewhat	42	55	3·6 (4·0)	5·3 (5·7)	-0·30 (-0·70,0·10)	-1·46 (-3·46,0·54)	0·47
Moderate /very useful	80	74	3·4 (4·7)	3·9 (4·6)	-0·11 (-0·42,0·21)	-0·52 (-2·08,1·03)	

[1] By linear regression, adjusted for recruitment centre, baseline PHQ-9 score, residence, treatment expectations, education and chronicity (see Methods). SD=Standard Deviation, SMD=Standardised Mean Difference, PHQ=Patient Health Questionnaire, THPP=Thinking Healthy Programme Peer-delivered, EUC=Enhanced Usual Care, CI=confidence interval.

Appendix F: PHQ-9 score and remission at 6 months with data restricted to sensitivity window.

	THPP plus EUC (n=122) [1]	EUC alone (n=129) [1]	SMD or PR (95% CI) [2]	Mean difference or OR (95% CI) [2]	p-value
PHQ-9 score at 6 months, mean (SD) [2]	3·5 (4·5)	4·5 (5·1)	SMD=-0·18 (-0·43,0·07)	AMD=-0·88 (-2·11,0·35)	0·16
Restricted to sensitivity window [3]	3·4 (4·5)	4·5 (5·2)	SMD=-0·21 (-0·46,0·04)	AMD=-1·01 (-2·27,0·25)	0·12
Remission (PHQ-9<5) at 6 months, number (%)	89 (73%)	77 (60%)	PR=1·21 (1·01,1·45)	OR=1·81 (1·04,3·13)	0·04
Restricted to sensitivity window [3]	87 (74%)	74 (59%)	PR=-0·21 (-0·47,0·05)	OR=1·92 (1·11,3·41)	0·02

[1] Results are mean (standard deviation) or number (%). [2] By linear or logistic regression, adjusted for recruitment centre, baseline PHQ-9 score, residence, treatment expectations, education and chronicity (see Methods). [3] Sensitivity window of 5.5 to 7 months, rather than 5 to 8 months as used in the main analyses; N=125 and 118 in in EUC alone and THPP plus EUC groups, respectively. PHQ=Patient Health Questionnaire, THPP=Thinking Healthy Programme Peer-delivered, EUC=Enhanced Usual Care, SMD=standardised mean difference, PR=prevalence ratio, OR=odds ratio, CI=confidence interval.

Appendix G: Alternative models for PHQ-9 score at 6 months.

Adjusted for	Mean difference (95% CI)	P value
Nothing	-1·01 (-2·21,0·19)	0·10
recruitment hospital	-1·02 (-2·22,0·18)	0·10
residence	-1·02 (-2·22,0·19)	0·10
recruitment hospital, residence	-1·02 (-2·22,0·18)	0·10
recruitment hospital, residence, baseline PHQ-9 score	-1·02 (-2·23,0·20)	0·10
recruitment hospital, residence, baseline PHQ-9 score, treatment expectations	-0·91 (-2·13,0·31)	0·15
recruitment hospital, residence, baseline PHQ-9 score, treatment expectations, education	-0·87 (-2·10,0·36)	0·16
recruitment hospital, residence, baseline PHQ-9 score, treatment expectations, education, chronicity [1]	-0·88 (-2·11,0·35)	0·16

Adjusted for	Mean difference (95% CI)	P value
As above and in addition adjusted for therapist (using GEE) [2]	-0·88 (-1·78,0·03)	0·06

	Mean difference (95% CI)	P value
Multiple imputation for missing outcome data [3]	-0·91 (-2·19,0·37)	0·16

	Ratio of mean PHQ-9 score (95% CI) [4]	P value
Poisson (with robust standard errors)	0·79 (0·59, 1·07)	0·13
Negative binomial	0·78 (0·55, 1·09)	0·14

[1] Main results reported. P values for the adjustment variables were: recruitment hospital p=0.70, residence p=0.83, baseline PHQ-9 score p=0.98, treatment expectations p=0.35, education p=0.29, chronicity, p=0.71. [2] Some women had more than one therapist; the most common therapist for each woman was used for analysis (there was one woman with 5 sessions each with 2 different therapists; the first therapist was used). There were 6 women in the intervention group who did not receive any therapy; these women were grouped into their own category. The women in the control group formed a further category. [3] Multiple imputation performed using treatment arm, recruitment hospital, residence, baseline PHQ-9 score, treatment expectations, education and chronicity, with 20 replications (linear regression model, i.e. not GEE).GEE=generalised estimating equations. [4] Adjusted for recruitment centre, baseline PHQ-9 score, residence, treatment expectations, education and chronicity (see Methods). This parameter is the ratio of the mean PHQ9 score in THPP plus EUC versus EUC alone groups, assuming a Poisson distribution with robust standard errors, or negative binomial distribution. PHQ-9=Patient Health Questionnaire-9. CI=confidence interval.

Adjusted for	Odds ratio (95% CI) for arm A versus B	Prevalence ratio (95% CI) for arm A versus B [1]	P value
Nothing	1·82 (1·07,3·10)	1·22 (1·02,1·46)	0·03
recruitment hospital	1·83 (1·07,3·12)	1·22 (1·02,1·46)	0·03
Residence	1·82 (1·07,3·11)	1·22 (1·02,1·46)	0·03
recruitment hospital, residence	1·83 (1·07,3·12)	1·22 (1·02,1·46)	0·03
recruitment hospital, residence, baseline PHQ-9 score	1·83 (1·07,3·13)	1·22 (1·02,1·46)	0·03
recruitment hospital, residence, baseline PHQ-9 score, treatment expectations	1·77 (1·03,3·04)	1·21 (1·01,1·45)	0·04
recruitment hospital, residence, baseline PHQ-9 score, treatment expectations, education	1·78 (1·03,3·08)	1·21 (1·01,1·45)	0·04
recruitment hospital, residence, baseline PHQ-9 score, treatment expectations, education, chronicity [2]	1·81 (1·04,3·13)	1·21 (1·01,1·45)	0·04

Adjusted for	Odds ratio (95% CI) for arm A versus B	Prevalence ratio (95% CI) for arm A versus B [1]	P value
As above and in addition adjusted for therapist (using GEE) [3]	1·80 (1·13,2·86)	1·21 (1·04,1·38) [4]	0·01

	Odds ratio (95% CI) for arm A versus B	Prevalence ratio (95% CI) for arm A versus B [1]	P value
Multiple imputation for missing outcome data [5]	1·81 (1·03,3·15)	-	0·04

[1] Standard errors estimated using the delta method. [2] Main results reported. P values for the adjustment variables were: recruitment hospital p=0.72, residence p=0.92, baseline PHQ-9 score p=0.84, treatment expectations p=0.76, education p=0.55, chronicity p=0.33. [3] Some women had more than one therapist; the most common therapist for each woman was used for analysis (there was one woman with 5 sessions each with 2 different therapists; the first therapist was used). There were 6 women in the intervention arm who did not receive any therapy; these women were grouped into their own category. The women in the control arm formed a further category. [4] PRs were estimated from models with the following reference categories: GMC, rural residence, moderate PHQ-9 (score 10-14), treatment expectation "very useful", completed primary education, and chronicity ≥12 weeks. PRs were estimated using marginal standardisation with the delta method for the CIs (see Localio AR, Margolis DJ, Berlin JA. Relative risks and confidence intervals were easily computed indirectly from multivariable logistic regression. *Journal of clinical epidemiology* 2007; **60**(9): 874-82) [5] Multiple imputation performed using treatment arm, recruitment hospital, residence, baseline PHQ-9 score, treatment expectations, education and chronicity, with 20 replications (logistic regression model, ie not GEE). GEE=generalised estimating equations. PHQ=Patient Health Questionnaire, CI=Confidence Interval.

	Number		Mean (SD) or n (%)		Mean difference or OR (95% CI) for arm THPP plus EUC vs EUC alone [1]	SMD or PR (95% CI) for THPP plus EUC vs EUC alone [1]	p-value
	THPP plus EUC	EUC alone	THPP plus EUC	EUC alone			
PHQ-9 score at 3 months	123	122	4·3 (4·2)	5·8 (5·7)	-1·68 (-2·97,-0·40)	-0·34 (-0·59,-0·09)	0·01
Remission (PHQ-9<5) at 3 months	123	122	74 (60%)	62 (51%)	1·61 (0·94,2·76)	1·22 (0·97,1·52)	0·08
Response (PHQ<10) at 3 months	123	122	107 (87%)	94 (77%)	2·39 (1·15,4·96)	1·15 (1·03,1·30)	0·02
Response (PHQ<10) at 6 months	122	129	108 (89%)	110 (85%)	1·23 (0·57,2·64)	1·03 (0·93,1·13)	0·60
Recovery (PHQ-9<5 at both 3 and 6 months)	116	120	61 (53%)	49 (41%)	1·82 (1·05,3·15)	1·35 (1·03,1·78)	0·03
WHO-DAS score							
3 months	123	122	11·9 (13·3)	16·2 (15·5)	-4·99 (-8·72,-1·25)	-0·34 (-0·59,-0·10)	0·009
6 months	122	128	9·5 (11·9)	12·1 (13·6)	-2·34 (-5·60,0·92)	-0·18 (-0·43,0·07)	0·16
Number of days unable to work in last month							
3 months	123	122	2·1 (5·2)	2·1 (4·9)	-0·02 (-1·33,1·30)	0·00 (-0·25,0·25)	0·98
6 months	122	128	2·5 (5·6)	2·7 (5·5)	-0·28 (-1·72,1·16)	-0·05 (-0·30,0·20)	0·70
MSPSS score							
3 months	122	122	5·7 (1·0)	5·4 (1·1)	0·34 (0·06,0·63)	0·32 (0·06,0·57)	0·02
6 months	122	128	5·6 (1·1)	5·3 (1·2)	0·32 (0·02,0·62)	0·27 (0·02,0·52)	0.04
Exclusive breastfeeding in last 24 hours							
3 months	121	122	86 (71%)	82 (67%)	1·28 (0·71,2·28)	1·07 (0·91,1·27)	0·41
6 months	121	128	9 (7%)	10 (8%)	0·80 (0·30,2·14)	0·82 (0·35,1·92)	0·65
Infant weight for age z score							
3 months	110	115	-1·21 (0·95)	-1·30 (1·05)	0·11 (-0·16,0·38)	0·11 (-0·15,0·37)	0·42
6 months	106	122	-1·03 (0·88)	-1·05 (1·03)	0·03 (-0·23,0·29)	0·03 (-0·23,0·29)	0·84
Infant height for age z score							
3 months	110	115	-0·38 (1·23)	-0·30 (1·10)	-0·09 (-0·40,0·22)	-0·08 (-0·34,0·18)	0·58
6 months	106	122	-0·43 (1·12)	-0·42 (1·13)	-0·02 (-0·31,0·27)	-0·02 (-0·28,0·24)	0·89

Appendix J: Minimal clinical important difference (MCID).

Relative reduction in PHQ-9 score, by mother's report [1]		
Overall, thinking back since the time since you entered this study how much do you feel your tension (or related health problems) has CHANGED?	N (% of 250)	Relative reduction in PHQ-9 score at 6 months versus baseline, median (range) [2]
Same, or worse	24 (10%)	82% (-91%,100%)
A lot better, or a little better	226 (90%)	85% (-82%,100%)

[1] The table summarises the PHQ-9 score reduction at 6 months versus baseline, by the "gold standard" classification, for the 250 women who had PHQ-9 score at 6 months and responded to the "gold standard" question. We see that 90% of women reported feeling better, the ranges of relative reductions in both groups are large, and the median relative reductions are similar. [2] Calculated as (PHQ-9 score at baseline – PHQ-9 score at 6 months)/PHQ-9 score at baseline *100. PHQ=Patient Health Questionnaire

90% of women reported feeling better. As no cut points for MCID yielded satisfactory sensitivity and specificity, further MCID analyses were not pursued.
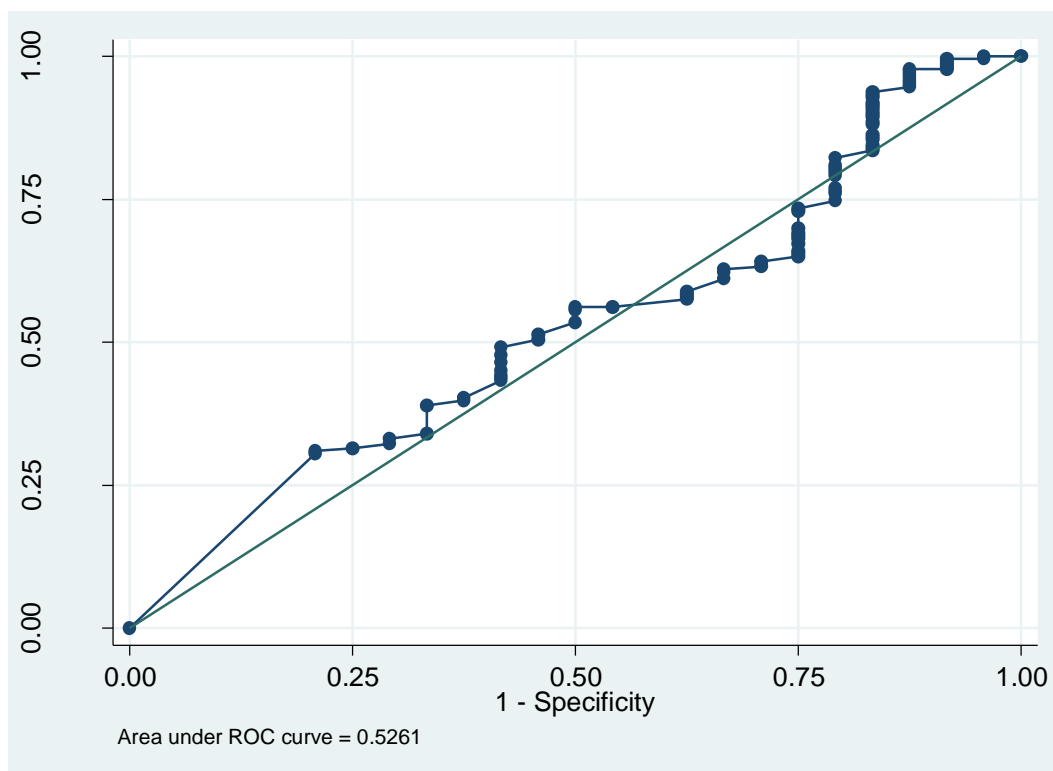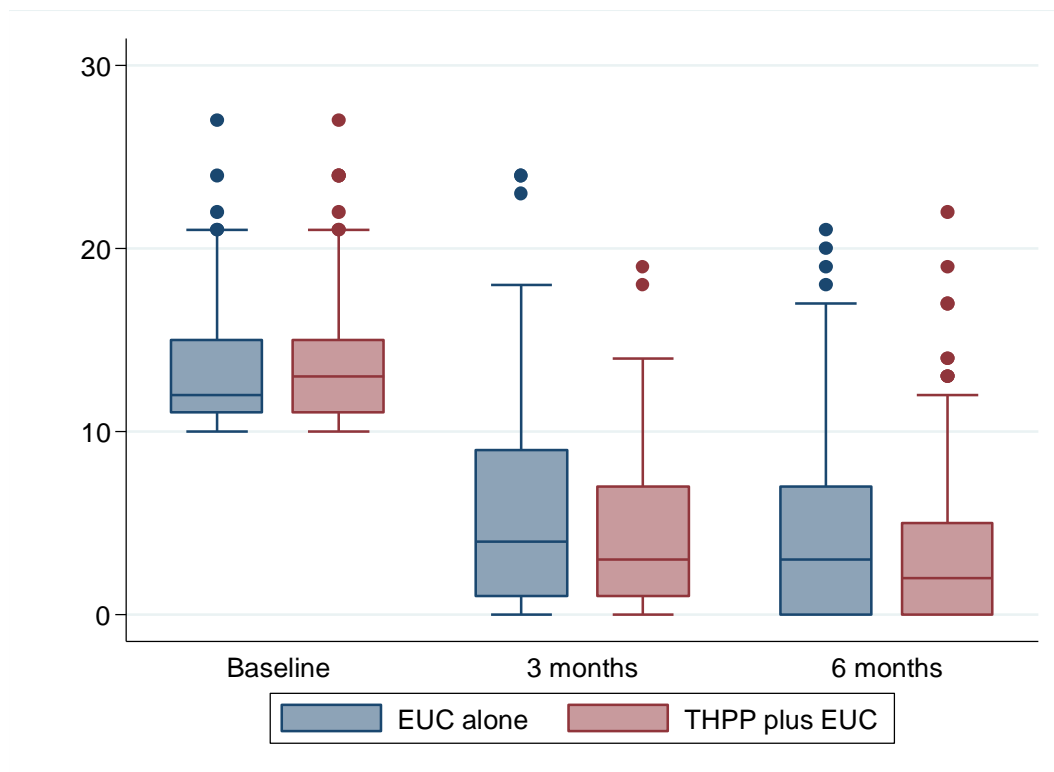
Figure: ROC curve for MCID analysis.

Figure shows the ROC (Receiver Operating Characteristics) curve for the MCID analysis, and the results are summarised in the table below. No cut points for MCID yielded satisfactory sensitivity and specificity; for a specificity of ≥70%, a cut point of ≥93% reduction was required, which yielded a sensitivity of only 33%. We considered other scenarios, also presented in the table below and reached similar conclusions. The MCID analyses were therefore not further pursued.

MCID results		
Approach	Cut off to achieve ≥70% specificity	Associated sensitivity
Original as described (following statistical analysis plan)	93%	33%
Alternative 1. Gold standard classification instead "a lot better" (68% of women)	90%	43%
Alternative 2. As 1, but restricted to women with PHQ-9≥15 at baseline	94%	31%
Alternative 3. As 1, and with absolute rather than relative reduction	12%	37%

The central line shows the median score, the length of the box is the interquartile range, the lines are the most extreme values within 1.5 times the interquartile range, and points outside this range are shown individually. EUC=Enhanced usual care. THPP=Thinking Healthy Programme Peer-Delivered. PHQ=Patient Health Questionnaire

	THPP plus EUC (n=140)	EUC alone (n=140)	P value [1]	Total (n=280)
Participants with at least one SAE	24 (17%)	27 (19%)	0.64	51 (18%)
Participants with at least one:				
Death of participant	0	0	-	0
Loss of child	2 (1%)	6 (4%)	0.28	8 (3%) [2]
Suicide attempt [3]	3 (2%)	2 (1%)	1.00	5 (2%)
Hospitalisation [4]	11 (8%)	7 (5%)	0.33	18 (6%)
Victimisation [5]	7 (5%)	6 (4%)	0.78	13 (5%)
Infant abuse/neglect [6]	0	0	-	0
Stigmatisation [7]	5 (4%)	12 (9%)	0.08	17 (6%)
Reported violence towards others [8]	0	1 (1%)	1.00	1 (<1%)
An event or finding with the potential to become a newsworthy/ international incident	0	0	-	0

Values are number of women experiencing this event at least once (not the total number of events) and column % of those enrolled. Each woman could have more than one type of SAE and therefore could be counted more than once in the table (aside from the first row): 9 women had 2 types of SAEs, and 1 woman had 3 types of SAEs. [1] P-values were calculated using X^2 tests where there were at least 5 observations in each category, and Fisher's exact tests where there were less than 5 observations in any category. [2] Including 2 miscarriages, 1 still birth, 4 children lost before 3 months of age, and 1 child lost between 3 and 6 months of age. [3] Based on responding "yes" at the 3 or 6 month visits to the question "Since you joined the study, have you ever acted on thoughts about hurting yourself?". [4] Includes unplanned mother and unplanned child hospitalisations (9 each). [5] Based on responding "yes" at the 3 or 6 month visits to the question "Have you experienced any form of physical violence since you began participating in our study?". [6] Based on responding "yes" at the 3 or 6 month visits to either of the questions "Since you started taking part in our study, has the baby suffered any beating or any other kind of physical or mental stress?" and "Has there been any indication during the assessment that the infant might have been physically or sexually abused?". [7] Based on responding "a lot" at the 3 or 6 month visits to any of the five questions "Have you or your family members experience any negative comments from others, or been treated in a negative way because you joined the study?", "Did you feel uncomfortable or worried because you thought other people might find out about your problem, make negative comments or treat you differently since you joined the study?", "Did your family members feel uncomfortable or worried because they thought other people might find out about your problem, make negative comments or treat you differently since you joined the study?", "Did you feel uncomfortable or worried because you felt other people were making negative comments or treating you differently since you joined the study?", "Did your family members feel uncomfortable or worried because they felt other people were making negative comments or treating you differently since you joined the study?". Fisher's exact p value=0.13. [8] Based on responding "yes" at the 3 or 6 month visits to the question "From the time you began participating in our study, have you hurt somebody else, for example, by slapping your baby or other children?". THPP= Thinking Healthy Programme Peer-delivered, EUC=Enhanced Usual Care.

Appendix M: Costs per beneficiary at 3 and 6 months, and over duration of trial period in US$, 2015.

Type of cost	Adjusted mean difference between THPP plus EUC and EUC alone at 3 months post-childbirth [1] (US$ Mean, 95% CI) ([#]n=244)	Adjusted mean difference between THPP plus EUC and EUC alone at 6 months post-childbirth [2] (US$ Mean, 95% CI) ([#]n=251)	Adjusted mean difference between THPP plus EUC and EUC alone over total period of trial [3](US$ Mean, 95% CI) ([#]n=236)
THPP intervention	1·36 (1·32 to 1·39)		
Doctor consultations (Primary health care)	0·69 (-2·33 to 3·71)	-0·36 (-1·14 to 0·40)	-0·08 (-3·29 to 3·12)
Doctor consultations (hospital)	-8·62 (-19·54 to 2·29)	-2·21 (-6·56 to 2·14)	-12·18(-24·27 to -0·09)
Hospital admissions	-6·13 (-33·20 to 20·94)	-0·40 (-1·25 to 0·44)	-6·06 (-33·97 to 21·84)
Laboratory tests	-0·25 (-2·12 to 1·62)	-0·41 (-1·32 to 0·49)	-0·93 (-3·03 to 1·78)
Medicines	1·85 (-3·77 to 7·46)	-0·48 (-2·76 to 1·81)	1·57 (-4·64 to 7·78)
* Total health system costs (exc. intervention)	-12·38 (-43·29 to 18·52)	-3·85 (-9·48 to 1·78)	-17·58 (-49·64 to 14·48)
** Total health system costs (incl. intervention)	-11·03 (-41·94 to 19·87)	-2·49 (-8·12 to 3·14)	-14·86 (-46·91 to 17·19)
Time costs	-6.38 (-13.20 to 0.47)	-0.58 (-2.05 to 0.88)	-7.87 (-15.04 to -0.70)
Productivity loss (patient)	-14.56 (-45.90 to 16.77)	-17.87 (-42.21 to 6.46)	-31.04 (-76.01 to 13.93)
Productivity loss (care-giver)	-22.01 (-44.39 to 0.37)	3.25 (-4.19 to 10.69)	-18.64 (-44.37 to 7.08)
Total productivity/time cost	-42·9 (-83·0 to -2·9)	-15·2 (-42·4 to11·9)	-57.55 (-112.89 to -2.21)
*** Total societal costs (health system and productivity costs	-53·98 (-109·64 to 1·68)	-17·69 (-45·94 to 10·54)	-72.41 (-141.84 to -2.98)

one participant declined to respond to the CSRI questions and was dropped from the analysis

*costs Included: in-patient, out-patient, medication, and investigations/tests

**costs included: health service utilisation plus intervention delivery (human resource use associated with the Sakhis)

***costs included: total health systems plus total productivity/time costs

[1] Cost and outcome estimates relate to the 6 month period covering the third trimester and the first 3 months post-childbirth

[2] Cost and outcome estimates relate to the 3 month period since the 3 month post-childbirth assessment

[3] Cost and outcome estimates relate to the 9 month period covering the third trimester and the first 6 months post-childbirth

Appendix N: Health system service utilisation and cost comparison by group and visit.

Table 1: Summary of service use

	3 months		6 months	
Indicator	THPP plus EUC (n=122)	EUC (n=122)	THPP plus EUC (n=122)	EUC (n=129)
	N (%)/mean (SD)	N (%)/Mean (SD)	N (%)/Mean (SD)	N (%)/Mean (SD)
Health system perspective				
In-patient care				
Hospitalised	100%	100%	1 (0.8%)	2 (1.5%)
Number of hospitalisations	151 (1·24; SD=0·66)	148 (1·21; SD=0·45)	1	2
Mean number of nights hospitalised (overall)	7·03 (7·13) [1-38]	7·32 (7·19) [1-51]	1 (-) [-]	2 (1.4) [1-3]
Mean number of nights hospitalised (government hospital)	6·68 (7·24) [1-38]	7·09 (7·33) [1-51]	0	1
Mean number of nights hospitalised (private hospital)	0·34 (1·74) [1-15]	0·22 (1·08) [1-8]	1 [-]	3 [-]
Outpatient care				
Number of contacts to health facility	897 (7·4; SD=7·1)	974 (7·9; SD=6·1)	149 (1·2; SD=1·8)	173 (1·3; SD=1·6)
Time spent with provider (min)	20·6 (12·2) [0-70]	22·3 (18·0) [3-150]	8·9 (13·7) [0-65]	8·5 (11·1) [0-85]
Intervention (Sakhis)				
Session duration (min)	37·5 (6·1) [22·7-53·2]			
Societal perspective				
Productivity/Time				
Time spent on OPD visits (min)	1158.6 (917.5) [40-7560]	1380.8 (1270.3) [50-10250]	253.5 (281.1) [5-1375]	222.2 (275.1) [0-1735
Number of days of lost wages (participant)	2.9 (6.2) [0-30]	3.5 (5.5) [0-25]	2.3 (4.2) [0-16]	3.1 (4.9) [0-20]
Number of days of lost wages (care-giver)	24.1 (24.9) [1-90]	35.5 (43.4) [2-180]	21 (28.7) [2-90]	5.6 (4.8) [1-20]

Table 2: Summary of cost in US$, 2015

Indicator	3 months		6 months	
	THPP plus EUC (n=122)	EUC (n=122)	THPP plus EUC (n=122)	EUC (n=129)
	N (%)/Mean (SD)	N (%)/Mean (SD)	N (%)/Mean (SD)	N (%)/Mean (SD)
Health system perspective				
In-patient care				
Overall Inpatient cost	105·6 (32·2) [14·9-569·1]	111·7 (30·8) [14·9- 763·8]	0·11 (0·73) [0-16·1]	0·51 (0·80) [0-48·4]
Inpatient cost (government)	106·2 (32·5) [14·9-569·1]	112·2 (31·1) [14·9- 763·8]	0	0·13 (0·86) [0-14·9]
Inpatient cost (private)	106·6 (76·6) [16·1-241·9]	69·1 (66·6) [16·1- 129·1]	0·13 (0·69) [0-16·1]	0·38 (0·74) [0-48·4]
Outpatient care				
Travel cost	2·49 (2·00) [0-20·8]	3·36 (1·96) [0-59·5]	0·74 (0·52) [0-14·5]	1·37 (0·51) [0-15·1]
Fee per visit (out of pocket)	1·9 (1·8) [0-121]	1·1 (1·6) [0-15]	1·1 (0·6) [0-21·1]	1·5 (0·6) [0-52·8]
Overall outpatient cost	38·3 (10·3) [0-405·2]	46·3 (9·1) [0·30- 184·9]	5·3 (2·1) [0-52·6]	7·8 (2·1) [0-219·9]
Outpatient cost (government)	28·8 (10·0) [0-248·5]	39·2 (8·1) [0-183·9]	2·1 (1·2) [0-38·5]	1·5 (1·2) [0-29·8]
Outpatient cost (private)	6·2 (5·2) [0-385·2]	3·9 (4·7) [0-51·4]	2·6 (2·6) [0-50·7]	5·4 (2·7) [0-219·9]
Medications and Investigations				
Medication cost	34·7 (6·3) [0-101·3]	33·3 (5·5) [0-117·9]	4·4 (2·9) [0-46·3]	4·8 (2·7) [0-46·3]
Laboratory tests cost	11·9 (1·2) [0-64·2]	12·4 (1·3) [0-56·6]	0·8 (1·1) [0-21·1]	1·2 (0·9) [0-30·2]
Intervention (Sakhis)				
Intervention delivery cost	1·36 (1·32 to 1·39) [0·8-1·9]			
Total health system costs (incl. intervention)	23435 (192·3; SD=32·9) [17·4-714·1]	24833 (203·3; SD=33·2) [60·6-836·3]	1427 (11·9; SD=4·9) [0·82 - 70·85]	1887 (14·4; SD=5·3) [0-220·9]
Difference in health system cost (95% CI)	-11·03 (-41·94 to19·87)		-2·49 (-8·12 to 3·14)	
Societal perspective				
Productivity/Time costs				
Time costs	27·4 (7·5) [0-182·7]	33·7 (6·6) [1·2-247·7]	2·9 (1·4) [0-33·2]	3·5 (1·2) [0-41·9]
Productivity loss (patient)	59·7 (29·0) [0-102·4]	74·3 (23·4) [0-102·4]	47·7 (15·4) [0-102·4]	65·6 (14·8) [0-102·4]
Productivity loss (care-giver)	19·8 (23·9)	41·8 (23·8)	6·6 (7·8)	3·4 (6·8)

Indicator	3 months		6 months	
	THPP plus EUC (n=122)	**EUC (n=122)**	**THPP plus EUC (n=122)**	**EUC (n=129)**
	N (%)/Mean (SD)	N (%)/Mean (SD)	N (%)/Mean (SD)	N (%)/Mean (SD)
	[0-335·8]	[0-671·6]	[0-405·1]	[0-74·6]
Total productivity/time cost	106·9 (39·9) [0-352·5]	149·9 (35·5) [1·2-852·2]	57·3 (13·7) [0-414·4]	72·5 (13·9) [0-140·6]
Difference in productivity/time costs (95% CI)	-42·9 (-83·0 to -2·9)		-15·2 (-42·4 to11·9)	
Total societal costs (health system and productivity costs)	36802·2 (299·3; SD=56·5) [17·4-1005·6]	42806·1 (353·3; SD=31·5) [77·8-1113·7]	8423 (69·3; SD=16·7) [0·8-445·4]	11243 (86·9; SD=16·6) [0-238·3]
Difference in total societal costs	-53·98 (-109·64 to 1·68)		-17·69 (-45·94 to 10·54)	

one participant declined to respond to the CSRI questions and was dropped from the analysis

[xx]: indicates range.

Appendix O: Summary of unit costs used in cost estimations

Category	Unit cost (in US$)	Cost category	Source
Human resource use (Sakhi)	0.04 per minute	Health systems	Primary data on Intervention delivery for this trial
Unemployed or unskilled manual worker	0.007per minute	Productivity	PREMIUM trial* (sourced from daily wage rates recommended in 2015 by the Indian Office of the Labour Commissioner
Skilled manual worker	0.008 per minute	Productivity	PREMIUM trial (same as above)
Professional/clerical	0.009 per minute	Productivity	PREMIUM trial (same as above)
Doctor consultation	0.25 per minute	Productivity	PREMIUM trial (sourced from a previous economic evaluation study in Goa**)
Hospital in-patient bed (private)	16.1 per day	Health systems	PREMIUM trial (same as above)
Hospital in-patient bed (government)	14.9 per day	Health systems	PREMIUM trial (same as above)

* Patel V, Weobong B, Weiss HA, et al. The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: a randomised controlled trial. *Lancet* 2017; **389**(10065): 176-85;

Nadkarni A, Weobong B, Weiss HA et al. Counselling for Alcohol Problems (CAP), a lay counsellor-delivered brief psychological treatment for harmful drinking in men, in primary care in India: a randomised controlled trial. *Lancet*. 2017 Jan 14;389(10065):186-195.

** Buttorff C, Hock RS, Weiss HA, et al. Economic evaluation of a task-shifting intervention for common mental disorders in India. *Bull World Health Organ* 2012; **90**(11): 813-21.)