B. Weobong, H. A. Weiss, D. McDaid, D. R. Singla, S.D. Hollon, A. Nadkarni, A-L. Park, B. Bhat, B. Katti, A. Anand, S. Dimidjian, R. Araya, M. King, L. Vijayakumar, G. T. Wilson, R. Velleman, B. R. Kirkwood, C. G. Fairburn, V. Patel

Sustained effectiveness and cost-effectiveness of the Healthy Activity Program, a brief psychological treatment for depression delivered by lay counsellors in primary care: a randomised controlled trial

Article (Published version)
(Refereed)

Sustained effectiveness and cost-effectiveness of the Healthy Activity Program, a brief psychological treatment for depression delivered by lay counsellors in primary care: twelve-month follow-up of a randomised controlled trial.

Short title: Sustained effects of a lay counsellor delivered brief psychological treatment for depression: twelve-month follow-up of a randomised controlled trial.

Benedict Weobong, Helen A Weiss, David McDaid, Daisy R Singla, Steven D Hollon, Abhijit Nadkarni, A-La Park, Bhargav Bhat, Basavraj Katti, Arpita Anand, Sona Dimidjian, Ricardo Araya, Michael King, Lakshmi Vijayakumar, G. Terence Wilson, Richard Velleman, Betty R Kirkwood, Christopher G Fairburn, Vikram Patel*

Benedict Weobong[1]
Email: Benedict.weobong@lshtm.ac.uk

Helen A Weiss[2]
Email: Helen.weiss@lshtm.ac.uk

David McDaid[3]
Email: D.Mcdaid@lse.ac.uk

Daisy R Singla[4]
Email: daisy.singla@sinaihealthnetwork.ca

Steven D Hollon[5]
Email: steven.d.hollon@vanderbilt.edu

Abhijit Nadkarni[1,6]
Email: Abhijit.nadkarni@lshtm.ac.uk

A-La Park[3]
E-mail: a.park@lse.ac.uk

Bhargav Bhat[6]
Email: bhargav.bhat@sangath.in

Basavraj Katti[6]
Email: basavraj.katti@sangath.in

Arpita Anand[6]
Email: anand.arpita@gmail.com

Sona Dimidjian[7]
Email: Sona.dimidjian@colorado.edu

Ricardo Araya[1,8]
Email: ricardo.araya@kcl.ac.uk

Michael King[9]
Email: Michael.king@ucl.ac.uk

Lakshmi Vijayakumar[10]
Email: lakshmi@vijayakumars.com

G. Terence Wilson[11]
Email: tewilson@rci.rutgers.edu

Richard Velleman[6,12]
Email: r.d.b.velleman@bath.ac.uk

Betty R Kirkwood[1]
Email: betty.kirkwood@lshtm.ac.uk

Christopher G Fairburn[13]
Email: credo@medsci.ox.ac.uk

Vikram Patel*[1,6,14]
Email: vikram_patel@hms.harvard.edu

[1] Centre for Global Mental Health, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK.
[2] MRC Tropical Epidemiology Group, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK.
[3] Personal Social Services Research Unit, London School of Economics and Political Science, London, UK.
[4] Department of Psychiatry, Sinai Health Network, University of Toronto, Toronto, Canada
[5] Department of Psychology, Vanderbilt University, Nashville, TN, USA.
[6] Sangath, Socorro Village, Bardez-Goa, 403501 Goa, India.
[7] Department of Psychology and Neuroscience, University of Colorado, Boulder, CO, USA.
[8] Institute of Psychiatry, Psychology, and Neurosciences, King's College Hospital, London, UK.
[9] Division of Psychiatry, Faculty of Brain Sciences, University College, London, UK.
[10] Sneha, VHS, University of Melbourne, Australia.
[11] Department of Psychology, Rutgers School of Arts and Sciences, USA.
[12] Department of Psychology, University of Bath, Bath, UK.
[13] Department of Psychiatry, University of Oxford, Oxford, UK.
[14] Department of Global Health and Social Medicine, Harvard Medical School, Boston, USA.

*vikram_patel@hms.harvard.edu

ABSTRACT

Background: The Healthy Activity Program (HAP), a brief behavioural intervention delivered by lay counsellors, enhanced remission over 3 months among primary care attendees with depression in peri-urban and rural settings in India. We evaluate the sustainability of the effects after treatment termination and the cost-effectiveness of the HAP over 12 months and the effects of the hypothesized mediator of activation on clinical outcomes.

Methods and Findings: Primary care attenders aged 18-65 screened with moderately severe to severe depression on the Patient Health Questionnaire (PHQ-9) were randomised to either HAP plus Enhanced Usual Care (EUC) (n=247) or EUC alone (n=248), of whom 95% completed assessments at 3 months and 91% at 12 months. Primary outcomes were severity on the Beck Depression Inventory version II (BDI-II) and remission on the PHQ-9.

HAP participants maintained the gains they showed at the end of treatment through the 12-month follow-up (difference in mean BDI-II score between 3 and 12 months=-0.34; 95% CI -.2.37, 1.69; p=0.74), with lower symptom severity scores than EUC alone (adjusted mean difference in BDI-II score=-4.45, 95%CI -7.26, -1.63; p=0.002) and higher rates of remission (adjusted prevalence ratio (aPR) =1.36, 95%CI 1.15, 1.61; p<0.009). They also fared better on most secondary outcomes including recovery (aPR=1.98, 95%CI 1.29, 3.03; p=0.002), any response over time (aPR=1.45, 95%CI 1.27, 1.66), higher likelihood of reporting a minimal clinically important difference (aPR=1.42, 95%CI 1.17, 1.71; p<0.0001); and lower likelihood of reporting of suicidal behavior (aPR=0.71, 95%CI 0.51, 1.01; p=0.06). HAP plus EUC also had a marginal effect on WHO-DAS score at 12 months (aPR=-1.58, 95% CI -3.33, 0.17; p=0.08); other outcomes (days unable to work, intimate partner violence toward females), did not reach statistical significance. Economic analyses indicated that HAP was dominant over EUC alone, with lower costs and better outcomes; uncertainty analysis showed that from this health system perspective there was a 95% chance of HAP being cost-effective, given a willingness to pay threshold of $16,060, equivalent to GDP per capita

in Goa, per QALY gained. Patient-reported behavioural activation levels at 3 months mediated the effects of the HAP intervention on the 12-month depression scores (β=-2.62, 95% CI -3.28, -1.97;p<0.0001). Serious Adverse Events were infrequent and prevalence was similar by arm.

Conclusions: HAP's superiority over EUC at the end of treatment was largely stable over time and mediated by patient activation. HAP provides better outcomes at lower costs adopting a perspective covering publicly funded health care services and productivity impacts on patients and their families.

Main limitations: We were unable to assess possible episodes of remission and relapse which may have occurred between our outcome assessment time points of 3 and 12 months post randomization. We did not account/evaluate the effect of mediators other than behavioural activation

Trial registration: ISRCTN95149997 (http://www.isrctn.com/ISRCTN95149997).

AUTHOR SUMMARY

Background:

- Depression is the leading mental health contributor to the global burden of disease.
- Access to effective treatments is low globally, but especially so in low and middle-income countries (LMICs) like India where a recent national survey reported a treatment gap of 85%.
- The Healthy Activity Program (HAP) is a brief psychological treatment based on the principles of behavioural activation and delivered by non-specialist providers; we have earlier reported the effectiveness of this intervention in reducing depressive symptoms and promoting remission at the end of treatment.

Why Was This Study Done?

- To evaluate the sustained effectiveness and the cost-effectiveness of the HAP over 12 months.
- To assess whether behavioural activation reported by patients at 3 months mediated the effects of the intervention on depression at 12 months.

What Did the Researchers Do and Find?

- We implemented a randomised controlled trial in which 493 adult primary health care attendees with moderately severe or severe depression who were assigned to either the HAP treatment (N=245) or enhanced usual care (EUC) (N=248), and received treatment over two to three months.
- HAP participants maintained the gains they showed at the end of treatment through the 12-month period, with lower symptom severity scores than EUC alone and higher rates of remission; these effects were partly mediated by increased levels of behavioural activation reported at 3 months;
- HAP was highly likely to be cost-effective, and could even save money if productivity costs were taken into account.

What Do These Findings Mean?

- The HAP is associated with sustained effects on depression outcomes over a 12-month period and represents good value for money.
- The HAP is ideally suited for scaling up to reduce the treatment gap for depression.

INTRODUCTION

Depression is a major contributor to the global burden of disease[1], and its treatment is a priority in the global health agenda. Despite the well-documented health and economic consequences of depression[2, 3], investments in mental health are inadequate, resulting in a large treatment gap [3]. Access to treatment remains a challenge particularly in low and middle-income countries (LMICs). The recent National Mental Health Survey in India reported a treatment gap of 85% for major depression[4]. Psychological treatments (PT) are recommended as first line interventions[5], not only because they are as efficacious as pharmacological treatments, but because they also produce sustained effects after treatment termination[6]. However, there are questions about the generalizability of PTs in LMICs, where the lack of trained professionals, variations in explanatory models, and lower literacy may present structural barriers to PT [7, 8]. Some of these barriers could be overcome by the innovative use of task-sharing and there is growing evidence for the acceptability and effectiveness of contextually-sensitive PTs delivered by appropriately-trained and supervised lay health workers in primary care and community settings[9-11]; however, there are very few trials which have reported on the sustained effects, cost-effectiveness or mediation of the effects of these treatments.

The PREMIUM (PRogram for Effective Mental health Interventions in Under-resourced health systeMs) was designed to: 1) implement a methodology for the development of scalable PTs that are culturally appropriate, affordable, and feasible for delivery by non-specialist health workers; and 2) evaluate the effectiveness and cost-effectiveness of the PTs on the two leading mental health causes of the burden of disease, ie. the Counselling for Alcohol Problems program for harmful drinking[12], and the Healthy Activity Program for moderately severe to severe depression (HAP])[13, 14]. The HAP treatment is adapted from behavioural activation (BA), a treatment which has a strong theoretical and empirical evidence base across diverse contexts and patient populations[15]. The stance of BA is particularly attractive as it focusses on the link between activities and mood, whilst

emphasizing increased activation and engagement, problem solving skills, and enhanced social support. A core feature of PREMIUM was the delivery of both treatments by the same lay counsellors in routine primary care settings, as they would be used in actual clinical practice. Usual care in primary care for depression in India is, in effect, no care at all. This was confirmed in the study setting during the pilot study. This is primarily because most cases are not diagnosed and, amongst those who are, most do not receive either antidepressants or PT.

Previously, we reported the favourable results of the impact of 6-8 sessions of HAP on mental health and secondary outcomes at the primary 3-month post-enrolment end-point[16]. The key findings were that HAP produced significantly lower symptom severity (adjusted mean difference in the BDI-II score=-7.57, 95%CI -10.27, -4.86) and higher remission rate (adjusted prevalence ratio=1.61, 95%CI 1.34, 1.93). HAP also showed superior results on the secondary outcomes of disability, days out of work, and intimate partner physical violence in women. The incremental cost of HAP per quality adjusted life year gained was International $9,333 (95% CI 3862, 28169), with an 87% chance of being cost effective from a health systems perspective in the study setting. The question now becomes whether these effects were sustained following the end of treatment in a disorder that is highly prone to relapse and recurrence given the relatively brief duration, minimal dosage and delivery by non-specialized workers, of the HAP ('most brief' PTs, particularly behavioural activation-based treatments in High Income Countries, typically involve at least twice this number of sessions delivered by highly trained professionals). In addition, a meaningful sustained effect should be accompanied by a patient-defined clinically important improvement in symptoms, as well as whether the mediating factor targeted by the PT accounted its effects. In this paper, we address three novel questions: the stability of HAP's effects on depression and other outcomes at 12 months post-enrolment; the mediation of the clinical outcomes by patient activation assessed at 3 months; and the cost-effectiveness of the intervention over 12 months.

METHODS

The methods are described in detail in the protocol. The trial was conducted in accordance with the protocol (S1 Protocol) (ISRCTN95149997) [17], which was approved by the Trial Steering Committee. Approval for the conduct of the trial was obtained from the Institutional Review Boards of the London School of Hygiene and Tropical Medicine, Sangath (the implementing institution in India), and the Indian Council of Medical Research. Written (or witnessed, if the participant was illiterate) informed consent was mandatory for enrolment. This study is reported as per CONSORT guidelines (S1 Checklist).

Study design and participants: This was a parallel-arm individually randomized controlled trial (RCT) in ten primary health centres in Goa, a state on the west coast of India. Participants were adult primary health care (PHC) patients aged 18-65 years with a probable diagnosis of moderately severe to severe depression ascertained with the Patient Health Questionnaire (PHQ-9) score>14, a cut-point previously validated in the study setting, and who gave informed consent. Pregnant women and patients presenting with severe medical conditions requiring urgent medical attention, and those with hearing/speech difficulties were excluded. Participants were interviewed to collect data on socio-demographic factors and potential moderators of treatment outcome: gender, illness severity, duration of the illness, and expectations for treatment[18]. Sequential numbered opaque envelopes were used to randomize consenting participants in a 1:1 allocation scheme[19]. Enrolment was conducted between 28th October 2013 and 29th July 2015 and the final 12-month assessment was completed on 30th August 2016.

Sample size estimation: Our sample size estimations for the 3-month primary outcomes assumed an intra-cluster correlation (ICC) between clinics of 0.04, with one counsellor per PHC at any one time, loss to follow-up of 15% over 3 months, and a 1:1 allocation ratio. Based on these we aimed to recruit 500 participants (425 in our analysis sample) to detect

the hypothesized effects: i) a standardised mean difference (effect size) of 0.42 for the primary continuous outcome of depression severity with 90% power; and ii) proportion recovered of 65% in the HAP plus EUC compared with 44% in EUC with 92% power. The high follow-up rate (attrition rate of 9%), at 12 months means that we have 90% power to detect these effect sizes at 12 months.

Interventions: *Enhanced Usual Care (EUC)* comprised routine consultation with the PHC physician, enhanced by providing the screening results to both PHC physician and patient, and providing copies of a contextualized version of the mhGAP guidelines to the PHC physician including when and where to refer for psychiatric care[20]. EUC was available to all trial participants.

*Healthy Activity Program (*HAP): The HAP is a contextually adapted brief psychological treatment based on behavioural activation[13], which focused on increasing patient activation levels in pleasurable or mastery activities, and comprised the following strategies: psychoeducation, behavioral assessment, activity monitoring, activity structuring and scheduling, activation of social networks, and problem solving. HAP was delivered in an individual format, and involved six-to-eight sessions, each lasting 30-40 minutes, with the initial sessions being at weekly intervals. The beginning phase focused on orienting to treatment, a multi-session middle phase on teaching core intervention strategies, and a late phase on reviewing gains and termination. The middle phase could be extended with up to 2 additional sessions for patients who did not show sufficient improvement, allowing a maximum of 8 sessions across all phases. Patients who did not respond by the end of treatment were referred for specialist care. Details about the intervention are reported elsewhere[13] and can be accessed online (http://hap.nextgenu.org). A description of counsellor selection, training and supervision is published elsewhere [21] [22]. Counsellors were members of the local community, above 18 years of age, completed at least high school education, and did not have prior professional mental health training. Counsellors

underwent a three-week participatory workshop covering both PTs, followed up by an internship phase of 6 months, in which trainee counsellors delivered the treatment to eligible patients in primary health-care clinics. Eleven counsellors who met competency standards participated in the trial. They received weekly peer-led supervision in groups of four to six and individual supervision twice monthly.

The same counsellor delivered the Counselling for Alcohol Problems treatment to adult males who met criteria for harmful drinking. Counsellors maintained separate clinical registers for both group of patients and reviewed individual patient records before each session. In order to ensure their treatment-specific counselling skills were maintained throughout the trial, weekly peer-led group supervision sessions were structured in ways that involved holding separate sessions for each of the two treatments. This arrangement allowed the expert supervisors for each of the two treatments to provide more focused feedback to the counsellors.

Treatment fidelity was assessed at two levels: the quality with which the HAP was delivered; and the quantity of the dose of HAP administered. The quality of HAP was assessed based on a random selection of 10% of audio-recorded sessions, rated on a therapy quality scale[22], by peers and experts. The quantity of HAP delivered was assessed based on treatment completion records maintained by the counsellors.

Outcomes: The two primary outcomes for the 12 month analyses were: 1) depression severity assessed by the Beck Depression Inventory version II (BDI-II) (dropping the item related to sex for cultural reasons); and 2) partial remission from depression (defined as PHQ-9 score <10). Our cut-off is in alignment with the depression treatment literature which defines remission as either the complete absence of symptoms, which is reflected by a PHQ-9 score< 5 or a partial absence of symptoms defined as PHQ-9 score<10 [23, 24]. A range of secondary outcomes included recovery from depression (PHQ9 score <5 at both 3

312 and 12 months); relapse (partial or full); disability; suicidal behaviour; and inter-personal
313 violence.
314
315 We estimated the Minimal Clinically Important Difference (MCID) as a patient-centred metric
316 that captures both the magnitude of improvement and the value the patient places on that
317 improvement[25]. We used the anchor-based approach for estimating MCID that ties change
318 in outcome on the PHQ-9 to the patient's subjective sense of improvement[26]; patients'
319 rating of perceived improvement on a 'global rating of change' scale[27] was used to
320 calculate the corresponding difference in score (see S1 Table for definition of all secondary
321 outcomes). In addition, we assessed patient-reported activation levels, using a 5-item Likert
322 Scale (0-5) based on the Behavioural Activation for Depression Scale—Short Form[28]
323 (BADS-SF), at 3 months to test for mediation. This variable was pre-specified as a potential
324 mediator of the HAP on depression outcomes because patient activation levels are the
325 primary focus of treatments for depression based on the theory of behavioral activation. All
326 measures were carefully selected based on their psychometric properties and contextual
327 appropriateness. The BDI is a widely-used measure for evaluating depression in trials, and
328 has been used in surveys in India[29]; the PHQ-9 has been validated in primary care and
329 Konkani (widely spoken local language in trial area) version validated in Goa[30]; the WHO
330 disability assessment schedule version 2 (WHO-DAS II) is validated for international use and
331 used in previous trials in Goa[31, 32]; the Client Service Receipt Inventory (CSRI) has been
332 previously used in trials in the study setting[33, 34]; the two items on intimate partner
333 violence (IPV) were selected based on interviews used in earlier studies in Goa[35], and the
334 BADS-SF was translated into Konkani using standardized procedures followed by
335 piloting[13].
336
337 ***Statistical methods:*** Analyses were on an intention-to-treat basis using multiple
338 imputations (20 iterations) for missing outcome data via a data augmentation algorithm in
339 Stata 14.0. All models adjusted for PHC as a fixed effect to allow for within-PHC clustering

and baseline PHQ-9 scores. For continuous outcomes, intervention effects were estimated using linear regression and reported as adjusted mean differences (AMDs) and effect sizes (ESs), with 95% confidence intervals (CIs). For binary outcomes, intervention effects are reported as adjusted prevalence ratios (aPRs) estimated from logistic regression using the marginal standardisation technique for the prevalence ratios and the delta method for the CIs[36]. Sensitivity analyses included adjustment for counsellor as a random effect and complete case analyses. Repeated measures analyses were conducted to estimate the time-by-treatment interaction effect. In addition, we examined changes in mean outcome scores over time, by treatment condition. The MCID was estimated using Receiver Operator Characteristics (ROC) analysis in order to establish the minimum relative change in PHQ-9 score that best differentiates those individuals who felt better from those who did not. We applied the cut-point for minimum specificity of 70% suggested by Button and colleagues[26]. Following cut-point determination, a binary outcome variable was created and intervention effects reported as adjusted prevalence ratios (aPRs) estimated from logistic regression. Results are described in terms of strength of evidence rather than statistical significance; hence we did not adjust p-values for multiple comparisons[37]. Our approach to the mediation analysis involved the Monte Carlo Method for Assessing Mediation [38, 39] which has been shown to be more rigorous than the Sobel test and as accurate as bootstrapping[40]. In the current study, we computed a 95% CI with 20,000 repetitions. All regression models controlled for individual patient's baseline PHQ-9 scores as well as any variables that were found to be significantly related to either the proposed mediator or 12-month BDI-II scores. Variance inflation factor (VIF) was conducted for each independent variable that was entered into each regression model to assess multicollinearity between independent variables, with a conservative estimate (VIF≥5).

Economic evaluations were conducted from both the health care system (costs to the health system only) and the societal perspectives (health system costs plus impacts on productivity of patients and their families). Information on the use of health services, including contacts

with PHC, hospital doctor contacts and inpatient stays, medication use and diagnostic tests was collected from service users using a tailored version of the CSRI at 3 and 12 months. Unit costs for doctor contacts and inpatient stays were inflated to 2015 prices using unit costs that had previously been used for an economic evaluation in Goa[41]. Detailed information on medications and laboratory tests used, as well as costs to the public purse were recorded. Mean costs were then extrapolated to cover the full 12 months. Detailed information was also recorded on the time taken to deliver each HAP session, whether delivered at a PHC, over the telephone or at a patient's home. Travel time and transportation costs were also recorded for home visits, including 'no-show' home visits. Per minute unit costs for counsellors, taking account of their training, supervision and other overheads were then attached to time to estimate the total costs of intervention delivery.

Productivity costs consisted of patient time out of usual activities because of their health, as well as time costs for patients (and accompanying family members) related to the use of health services. The number of days completely out of normal role over the previous 30 days were based on patient responses to the WHO-DAS II at 3 months and 12 months. WHO-DAS II data on days of activity cutback over this period were also included, with the assumption that each day of cutback would have half the value of a complete day out of role, an approach that has been adopted in High Income settings[42]. Patients reported how much time was spent attending health services using the CSRI; patients were also asked to report if they were accompanied by someone. In this case it was also assumed that one family member incurred the same level of productivity losses. We assumed that the mean of patient and family time costs at 3 months and 12 months would also apply to the rest of the year. Costs due to cutback and complete days out of role were adjusted to avoid double counting time that patients spent attending health services. All patient and family time was valued using the human capital approach making use of different daily wage rates recommended in 2015 by the Indian Labour Commission. The rate used was dependent on whether the patient was classified as an unskilled, skilled or a clerical/professional worker.

We assumed the value of days out of role for those classified unemployed were the same as those for unskilled workers.

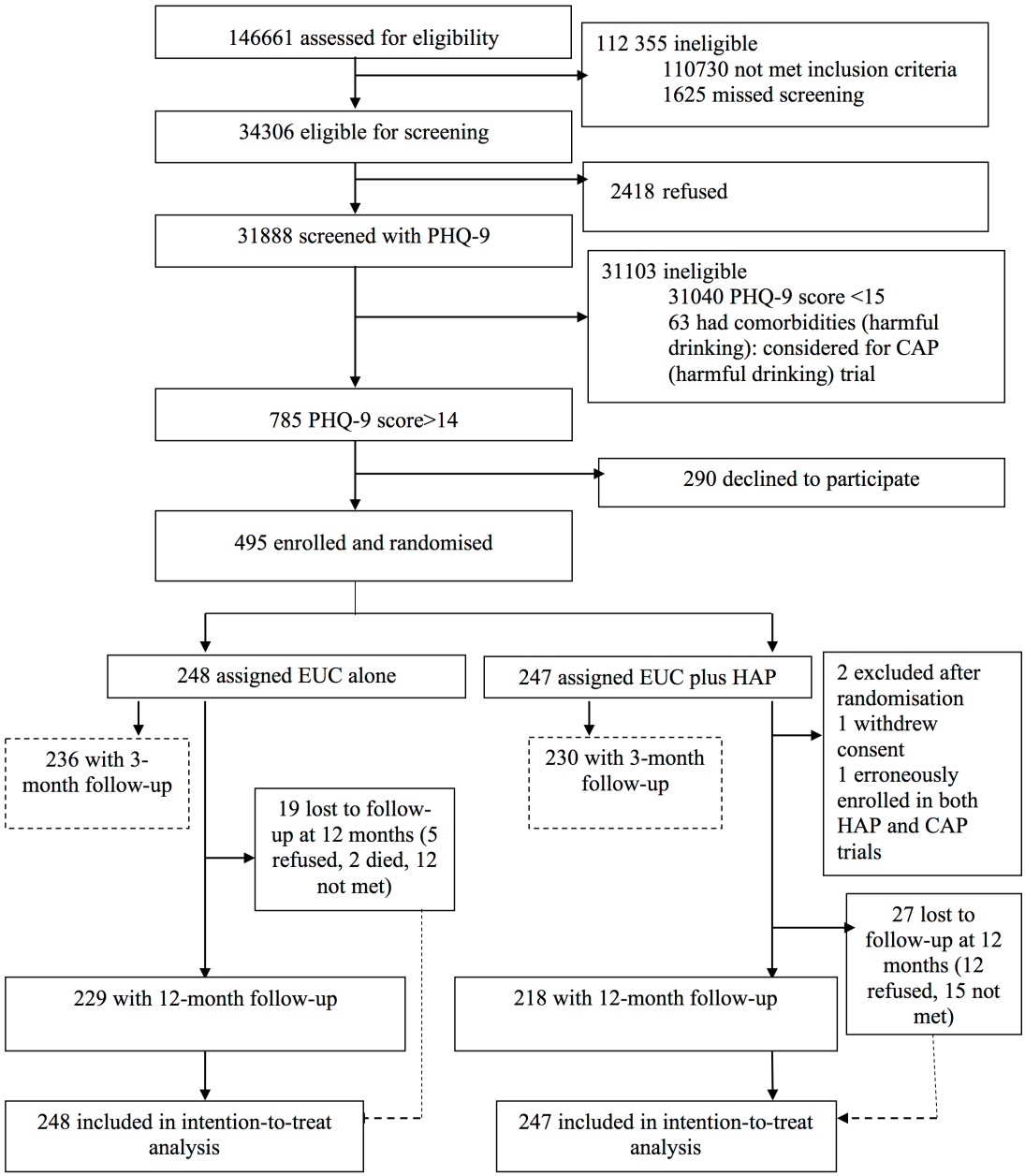
Quality Adjusted Life Year (QALY) scores were derived through transformation of WHO-DAS 12 item scores as in earlier Indian trials[41]. Incremental Cost-Effectiveness Ratios (ICERs) were bootstrapped, randomly resampling pairs of outcomes and costs for intervention and comparator groups to derive 95% Cis with a distribution of mean incremental costs and effects shown on cost effectiveness planes to test the robustness of cost results. Cost-effectiveness acceptability curves (CEAC) were also generated showing the likelihood that HAP would be cost-effective at different levels of willingness-to-pay. All statistical analyses were conducted using Excel 2016, SPSS 21 for the cost-effectiveness analyses; SAS, R-Studio for the mediation analyses; and STATA 13/14 for all other analyses. All costs are presented in 2015 International Dollars (http://eppi.ioe.ac.uk/costconversion/).

RESULTS

Trial conduct: A detailed description of the conduct of the trial is provided in the primary trial paper[16]. Between October 28, 2013, and July 29, 2015, 34,306 (23%) of the 146661 PHC attenders assessed met inclusion/exclusion criteria. Of these 31,888 adult PHC attenders were screened for depression using the PHQ-9 of whom 785 (2.5%) were eligible (PHQ-9 score>14) for inclusion in the trial, and 495 (63%) consented to participate and were enrolled. A total of 248 participants were randomized to EUC and 247 to HAP plus EUC. Of the latter, two were subsequently excluded (one withdrew consent and the other was erroneously enrolled in both trials) leaving a total of 245 participants treated with HAP plus EUC (Fig 1). The modal reason for non-participation was lack of time, and participants had similar baseline characteristics to non-participants. Baseline characteristics were similar by arm. 466 participants (95%) were assessed at the 3-month post-treatment endpoint and 447 participants (91%) at 12-month follow-up; rates were similar between arms. A total of 438

(89%) participants had observations for both follow up time-points. In all, only 18 (3.6%) participants did not have any follow-up data. Those lost to follow-up at 12-months were younger (S2 Table), and this was similar at the 3-month post-treatment endpoint. The intra-class correlation of BDI-II within PHCs was 0.02.

Figure 1: The Healthy Activity Program trial flow chart



Flow chart legend: (CAP=Counselling for Alcohol Problems. EUC=enhanced usual care. HAP=Healthy Activity Program. PHQ-9=Patient Health Questionnaire 9)

Impact on clinical outcomes: There was an intervention effect on both primary outcomes at the 12-month follow-up. The mean endpoint BDI-II score was 19.73 (SD 15.53) among participants in the HAP plus EUC arm and 24.09 (SD 14.67) among participants in the EUC arm (AMD=-4.45; 95%CI -7.26, -1.63; ES=0.23, 95%CI 0.18, 0.28; p=0.002; Table 1). This main effect at 12 months was influenced by the passage of time (p-value for time-by-treatment interaction 0.04), such that participants in the EUC arm continued to improve through the 12-month follow-up (difference in mean BDI-II score between 3 and 12 months=3.2; 95% CI 1.34, 5.06; p=0.001; S3 Table) while HAP plus EUC essentially retained the greater gains that it had made at the earlier assessment (difference in mean BDI-II score between 3 and 12 months=-0.34; 95% CI -.2.37, 1.69; p=0.74; S3 Table). Participants in the HAP plus EUC arm also had a higher probability of remission than those in the EUC arm (63.1% vs 48%; aPR=1.36, 95%CI 1.15, 1.61; p<0.001). As was the case for mean scores on the BDI-II, remission rates stayed relatively constant from 3 to 12 months among participants in the HAP plus EUC arm, whereas those in the EUC arm showed a slight increase by 12 months (Fig 2). Sensitivity analysis showed similar results (S4 Table). There was no evidence of moderation by gender, severity, chronicity, or patient expectancies (S5 Table).

Table 1: Effects of the HAP plus EUC compared with EUC alone on primary and secondary clinical outcomes at 12 months

Outcome	EUC arm (n=248)	HAP+EUC arm (n=245)	[1]Adjusted mean difference (AMD), effect size (ER), prevalence ratio (PR), prevalence difference (PD) (95% CI)	p-value
Primary outcomes at 12 months				
*Mean BDI-II score (SD)	24.09 (14.67)	19.73 (15.53)	AMD:-4.45 (-7.26, -1.63) ES: 0.23 (0.18, 0.28)	p=0.002
**Remission: PHQ-9<10 no. (%)	117 (46.98%)	155 (63.14%)	PR: 1.36 (1.15, 1.61) PD: 16.66% (7.85%, 25.47%)	p<0.001 p<0.001
Secondary outcomes at 12 months				
***Recovery: PHQ-9<5 at 3 & 12m no. (%)	33 (13.27%)	64 (26.10%)	PR: 1.98 (1.29, 3.03) PD: 12.96% (5.31%, 20.61%)	p=0.002 p=0.001
***Full relapse: PHQ-9 score 15-27 no. (%)	12 (4.92%)	21 (8.78%)	PR: 1.79 (0.87, 3.69)	p=0.14
***Partial relapse: PHQ-9 score 10-14 no. (%)	7 (2.70%)	21 (8.60%)	PR: 3.19 (1.27, 7.88)	p=0.01
***Mean PHQ-9 score (SD)	10.46 (7.54)	8.16 (6.96)	AMD: -2.36 (-3.70, -1.02) ES: 0.37 (0.32, 0.42)	p<0.001
Any response over 12 months no. (%)	266 (53.97%)	383 (77.65%)	PR: 1.45 (1.27, 1.66)	p<0.0001
#Suicidal behaviour	66 (26.55%)	47 (19.10%)	PR: 0.71 (0.51, 1.01)	p=0.06

Outcome	EUC arm (n=248)	HAP+EUC arm (n=245)	[1]Adjusted mean difference (AMD), effect size (ER), prevalence ratio (PR), prevalence difference (PD) (95% CI)	p-value
(Suicide thoughts) – no. (%)				
$MCID (% reduction in baseline PHQ-9 score)	102 (41.25%)	142 (58.10%)	PR: 1.42 (1.17, 1.71) PD: 17.08% (7.89%, 26.26%)	p<0.0001 p<0.0001

[1] Adjusted for PHC as a fixed effect and PHQ-9 baseline score

* Sensitivity analysis (point estimate: AMD): Random effects =-4.41(-7.21, -1.61); complete case = -4.57 (-7.34, -1.81); excluding unmasked (3.7%)=-4.40 (-7.29, -1.51)

** Sensitivity analysis (point estimate: PR): complete case=1.36 (1.14, 1.61)

***Not previously specified in trials protocol but specified in published analysis plan

#Suicidal thoughts over the past two weeks were assessed through the relevant PHQ-9 item while suicide attempts were assessed over the 3-month period leading up to the 12-month outcome follow up assessment. Attempts not included as numbers very small (only 2 patients (1 in each arm) reported suicide attempt over the period). ## Among married participants.

$Minimal Clinically Important Difference: estimated based on relative difference in baseline and outcome score, and how this compares with overall subjective global rating of 'feeling better' at the end of the trial. The optimal cut-off in relative change in score with maximum specificity (>70%) is 55

Figure 2: Remission rates over time between HAP plus EUC and EUC

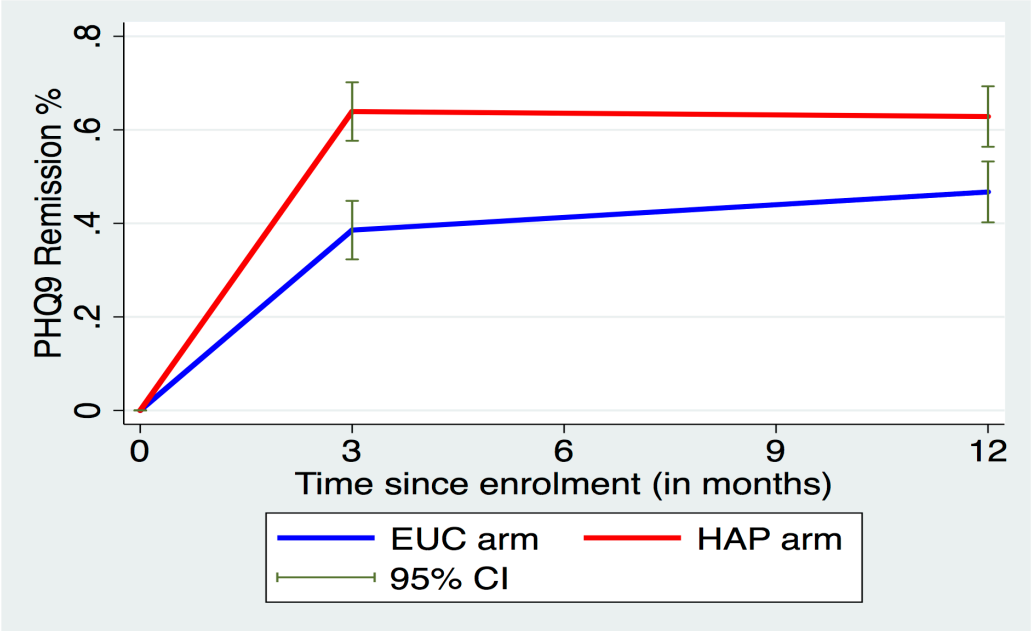


Figure legend: EUC arm: Enhanced Usual Care arm
HAP arm: Healthy Activity Programme arm

Participants in the HAP plus EUC arm had a higher probability of remission and recovery compared to those in the EUC arm (Table 1). While participants in the HAP plus EUC arm who had remitted at 3 months had a higher probability of partial relapse at 12 months compared to those in the EUC arm, the proportion with full relapse was similar between arms (Table 1). Participants in the HAP plus EUC arm also had a higher probability of *any* response over the 12 months (Table 1, Fig 3). More participants remitted in HAP plus EUC in the short-term compared to EUC alone, but as expected were more likely to relapse following treatment termination than patients who remitted in EUC alone (Fig 3). Participants in the HAP plus EUC arm had marginally lower prevalence of suicidal behavior (mainly

suicide thoughts as there were only two attempts) at 12 months. Our analysis on what constitutes a MCID, revealed a relative score change of 55% from baseline. Based on this score change, HAP plus EUC was superior to EUC at 12 months (aPR=1.42, 95% CI 1.17, 1.71; $p<0.0001$, Table 1).

Figure 3: Clinical trajectories in cases with 3 and 12-month outcome data (n=438)

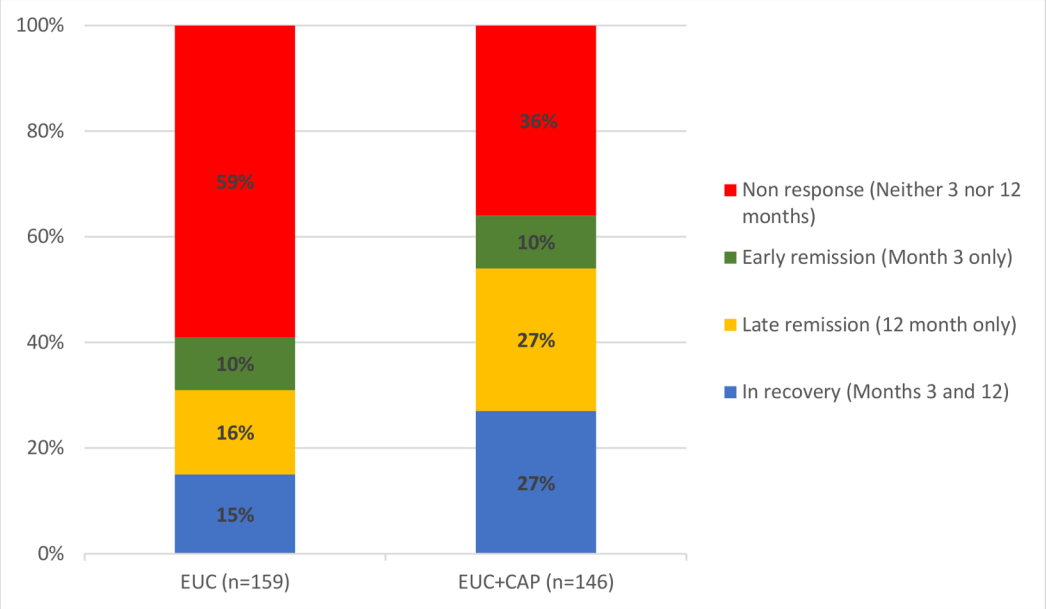


Figure legend: EUC arm: Enhanced Usual Care arm
HAP arm: Healthy Activity Programme arm

Impact on other outcomes and mediation analyses: HAP plus EUC also had a marginal effect on WHO-DAS score at 12 months (aPR=-1.58, 95% CI -3.33, 0.17; p=0.08); other outcomes (days unable to work, intimate partner violence toward females), did not reach statistical significance (Table 2). The prevalence of SAEs (HAP plus EUC=23; EUC=23) and proportion of participants prescribed antidepressant medications (HAP plus EUC=7; EUC=11) did not differ between the treatments (S6 Table). Our assessment of mediation demonstrated that patient-reported behavioural activation levels at 3-months partially mediated the superiority of HAP plus EUC relative to EUC in terms of reduced depression severity at 12-months (Beta coefficient=-2.62, 95% CI -3.28, -1.97; p<0.0001; Fig 4, also S7 Table). Patient-reported behavioural activation could account for 58% of the total effect of HAP plus EUC. None of the models evidenced multi-collinearity between the independent variables (VIF<5).

Of the 245 participants in the HAP group (receiving a total of 1181 sessions), 169 (69%) had a planned discharge, of whom seven (4%) were referred for specialist care. The median number of sessions was six (IQR five to seven). Patients with an unplanned discharge were likely to stop attending early (median one session [IQR none to two])

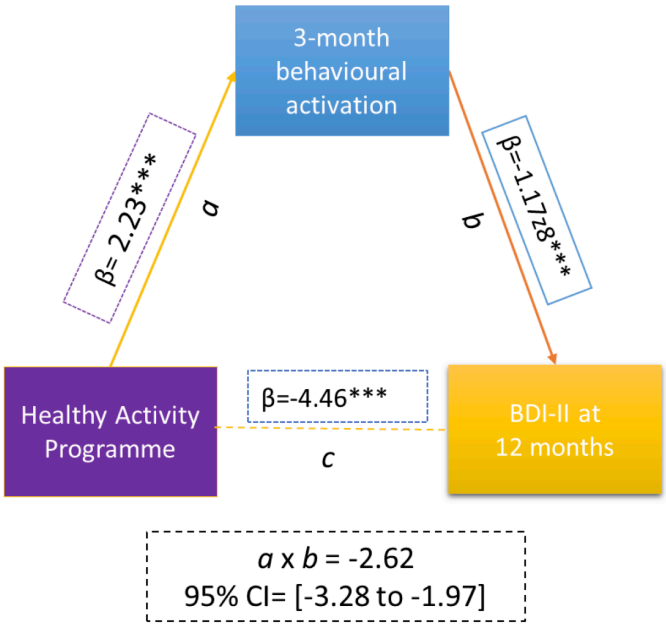
Table 2: Effect of HAP plus EUC compared with EUC alone on disability and intimate partner violence at 12 months

Outcome	EUC arm (n=248)	HAP+EUC arm (n=245)	[1]Adjusted mean difference (AMD), effect size (ER), prevalence ratio (PR), (95% CI)	p-value
Mean disability score (SD)	10.89 (9.22)	9.38 (9.61)	AMD: -1.58 (-3.33, 0.17) ES: 0.03 (-0.03, 0.8)	p=0.08
Mean days unable to work (SD)	6.05 (8.81)	4.81 (8.24)	AMD: -1.29 (-2.89, 0.31) ES: 0.09 (0.04, 0.15)	p=0.12
Intimate partner physical violence## – females no. (%)	20/118 (16.57%)	11/109 (9.86%)	PR: 0.60 (0.29, 1.22)	p=0.16
Intimate partner psychological/emotional violence## – females no. (%)	40/118 (33.86%)	28/109 (26.10%)	PR: 0.75 (0.50, 1.13)	p=0.17
Intimate partner psychological/emotional violence## – males no. (%)	12/40 (28.75%)	7/34 (19.23%)	PR: 0.82 (0.36, 1.84)	p=0.62

[1] Adjusted for PHC as a fixed effect and PHQ-9 baseline score

535

Figure 4. The mediating effect of behavioural activation at 3 months on the effectiveness of the HAP on depression severity at 12 months.



Note. Beta estimates are unstandardized. Multiple linear regression models controlled for baseline PHQ-9 scores, primary health centre, and age.
*p<0.05. **p<0.01. ***p<0.001

536
537
538 Figure legend:
539 Beta coefficient
540 *a*': a-path (HAP-mediator)
541 *b:* b-path (mediator-outcome)
542 c: direct effect (HAP-outcome)
543 *axb:* indirect effect
544
545
546

Costs and cost effectiveness: While health system costs had been significantly higher at 3 months follow up due to the cost of providing HAP, by 12 months these costs were offset by reductions in the use of health services through month 12 and there was no statistically significant difference in health system costs between the two arms (S8 Table). From a wider societal perspective, which combines impacts on the health system with impacts on productivity costs, the HAP plus EUC group had significantly lower costs at 12 months (mean difference -$154.93, 95% CI -$305.51, -$4.35; p=0.044), this was due to lower costs of days out of work and work cutback (mean difference -$146.28 (-$218.08, -$74.47 p=0.000). While there is still a gain in mean QALYs per person at 12 months, this difference was not quite statistically significant (mean difference 0.011, 95% CI 0.006, -0.002 p=0.092). Table 3 provides an assessment of cost effectiveness showing ICERs. It indicates that the incremental cost per QALY gained is -$1,721; thus, the HAP is associated with both lower costs and better outcomes than EUC alone. To test the robustness of the ICER results, two cost effectiveness analysis planes were generated using 1000 randomly resampled pairs of costs and QALY outcomes from both health system and societal perspectives to generate further incremental cost per QALY gained values (Fig 5). This can help policymakers by showing the likelihood that any intervention will be cost effective or even cost saving. Figure 5 (A) indicates that HAP plus EUC has a 58% chance of being cost saving from a health system perspective, i.e. 58% of the 1000 pairs of costs and QALYs are in the south-east quadrant, which indicates that the intervention (in this case HAP plus EUC) has both lower costs and better QALY outcomes than EUC, while a further 39% of the 1000 pairs of cost and QALYs fall in the north-east quadrant, where HAP plus EUC is more effective but more expensive than EUC. Nearly all of the observations in this quadrant were still below the cost effectiveness threshold used in the analysis (shown by the red line) of GDP per capita per additional QALY gained, a threshold which has been applied in economic evaluations in LMIC[43]. This threshold in the state of Goa expressed in international dollars in 2015 was $16 060[44]. Overall this means that the case for investment is very strong with a 95% likelihood that investment in the intervention will be cost effective, including a 58% chance

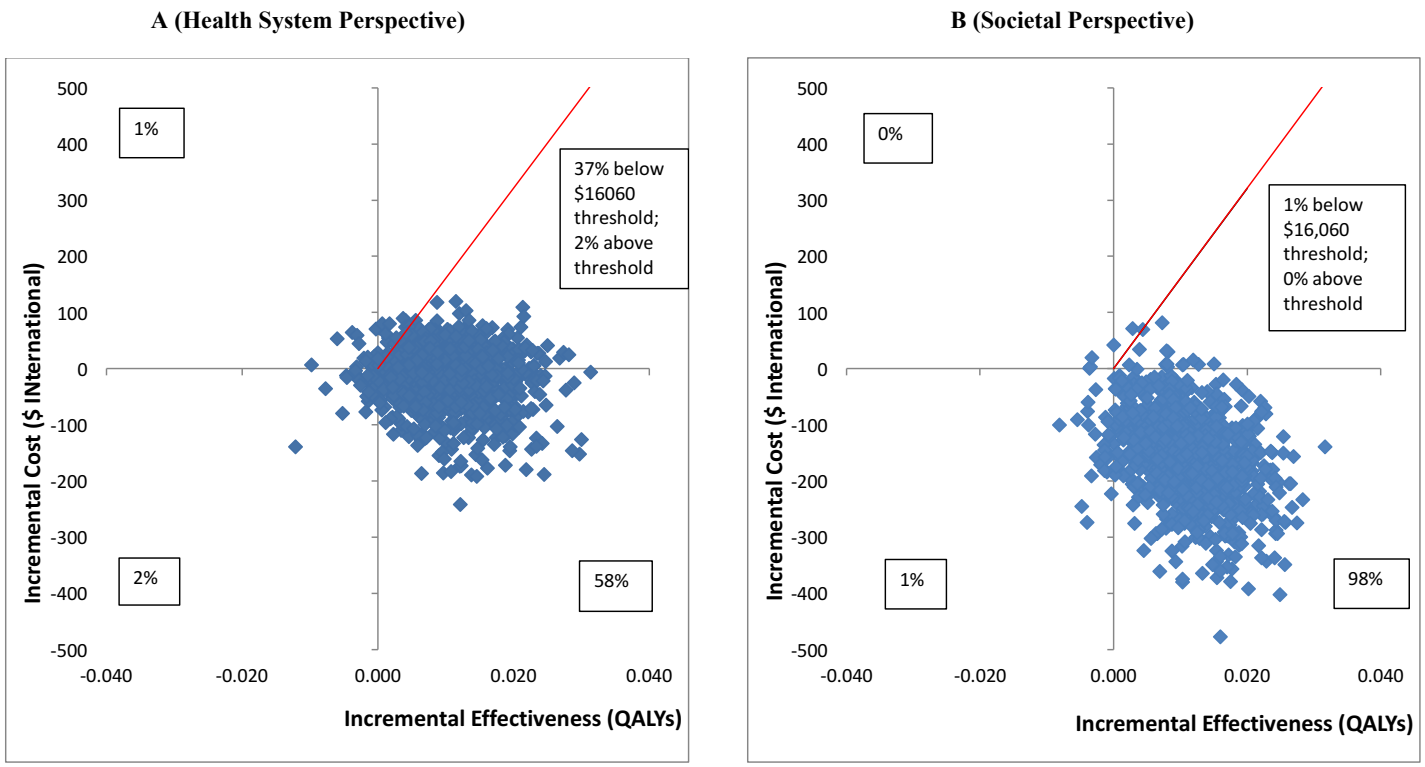
575 that it will be cost saving. Similarly, in Figure 5 (B) when costs also include a conservative
576 estimate of productivity losses to patients and families, 98% of the pairs of costs and QALYs
577 fall in the south-east quadrant, where HAP plus EUC is cost saving with lower costs and
578 better outcomes compared to EUC. As Table 3 shows, if the same approach is used to look
579 at costs per additional remission achieved compared to EUC from a health system
580 perspective, HAP plus EUC would be considered a highly worthwhile investment (S2 Fig)
581 with a 90% chance of being cost effective, including a 59% chance of being cost saving.
582

583 **Table 3: Cost-effectiveness analyses from health system and societal**
584 **perspectives (2015 International Dollars)**

	Health system perspective	Likelihood ICER cost saving (CS) and cost effective (CE)	**Societal perspective**	Likelihood cost saving (CS) and cost effective (CE)
Cost per QALY gained at 12 months (95% CI)*	-1721 (-23,966, 18,158)	CS: 58% CE: 95%	-14,438 (-81,359, 13,966)	CS: 98% CE: 99%
Cost per remission at 12 months (95% CI)**	-149 (-1,304, 988)	CS: 59% CE: 90%	-1,250 (-3,869, -186)	CS: 99% CE: 100%

585
586 *Assumes willingness to pay threshold equivalent to GDP per capita in Goa ($16,060)
587 **Assumes willingness to pay threshold equivalent to one month's wages for unskilled manual worker
588 in Goa ($415)

Figure 5: Cost effectiveness planes: HAP plus EUC compared to EUC



589

DISCUSSION

We report on the sustained effects, the cost-effectiveness, and the role of behavioural activation in mediating the effectiveness of the Healthy Activity Program, a brief psychological treatment delivered by lay counsellors to primary care attenders with moderately severe to severe depression in a randomized controlled trial in India. We have two main findings.

First, the effects of the HAP on acute depression observed shortly after the end of treatment (3 months) were largely sustained through the 12-month follow-up. This is striking because depression tends to return after treatment termination among recently remitted patients, one of the reasons why physicians are encouraged to keep patients on active medications for at least four months following initial remission[24]. What makes that less surprising is that HAP is adapted from behavioral activation and that approach was found to reduce risk for subsequent relapse by more than half, relative to prior medications in the one study in which they have been compared[45]. Patients who remitted on HAP in the short-term were more likely to relapse following treatment termination than patients who remitted in EUC, but that is to be expected since more patients remitted on HAP than in EUC and it is plausible that those additional remitters were tougher patients at higher risk (Fig 3). That being said, HAP's effects were relatively stable over time (i.e. depression severity scores did not change) and absolute relapse rates were lower than those observed for behavioral activation in the largest comparable trials[45]. In a disorder that is prone to relapse, that augers well for the possibility that HAP might have an enduring effect.

Our second major finding was that HAP essentially pays for itself and more. It cost $65.66 per patient to provide HAP but those extra treatment costs were completely offset by reductions in other health care expenses across the course of a year so that health care costs between the two trial arms were no longer significantly different at 12 months (they

had been significantly higher in the 3-month analysis[16]). Moreover, there was a very high 95% probability of HAP plus EUC being cost effective from a health system perspective, including a 58% probability that it would be cost saving. What our data suggest therefore is that the initial additional costs of providing HAP will be at least budget neutral from a health system perspective, while improving clinical outcomes.

When we factor in societal costs in terms of productivity, the economic benefits of HAP become even more evident. Poor mental health has been associated with significantly lower rates of participation in employment in Low, Middle and High-Income Countries, including in India, where severe mental illness has been associated with a 40% reduction in individual earnings[46]. Poor mental health also reduces the opportunity to contribute in other ways to the economy, such as household activities, because of time of usual activity; it also increases the use of informal care and support from families. Our analysis indicates that we also make major gains in terms of productivity that have real implications for the individuals involved and for the larger society in which they are embedded. The United Kingdom has committed over £700 million pounds to train therapists to deliver empirically-supported treatments like behavioral activation on the premise that doing so would be good for the economy[47]. Our data suggest that this assumption might well hold for this Indian setting despite the substantial structural differences which mean that the interventions and their contexts are not directly comparable.

Additionally, we observed that patients who received HAP reported feeling better subjectively at 12 months post-enrolment than patients who received EUC alone. HAP patients not only were better in terms of reported symptoms but they had the subjective sense that they were better in ways that actually mattered to them. This adds a patient-centered outcome to our main effectiveness results. At the same time, our mediation analysis suggested that patient-reported levels of behavioral activation at 3 months

mediated the effects of HAP in reducing depression severity at 12-months. This suggests that behavioural activation may underlie HAP's sustained effects and, thus, adds to existing evidence suggesting that patient-reported activation levels mediate response to behavioral activation therapy as specified by theory[48, 49].

Our effects were modest and about a third of patients treated with HAP remained at least moderately symptomatic. That being said, HAP was a very brief treatment by western standards (only 6-8 sessions) and was delivered by lay counsellors; most efficacy trials provide two-to-three times that many sessions delivered by highly trained professionals[50, 51]. Treatment differences did narrow over time from the 3-month post-treatment assessment to the 12-month follow-up but that was largely a function of continued improvement in the EUC condition (likely due to spontaneous remission) and not any loss of efficacy for HAP over time (within condition changes were not significant). Even the elevated relapse rate for HAP relative to EUC was limited to partial relapse (requiring a change of as little as a point to rise to 10 or above on the PHQ-9); there were no differences with respect to full relapse (scores of 15 or above). Notwithstanding these notable benefits, it is clear that HAP is not sufficient as a stand-alone treatment for depression for a sizeable minority of patients in primary care. Whether its dosage or duration needs to be extended or nonresponders switched or augmented with another treatment (like medications) remains to be determined.

We acknowledge limitations of this study design. First, from a methodological perspective, we only had two assessment time points at 3 months and 12 months, thus precluding detection of possible episodes of remission and relapse between these two time points[52]. Second, we continue to observe a pattern of discordance between our two primary outcome measures similar to what we found in our 3-month outcome assessments; patients at 12 months were in the low end of the moderate range of severity on the BDI-II, but the same patients were indicated as having mild residual symptoms on the PHQ-9. This suggests

potential cross-cultural challenges with the use of the BDI-II, which we are currently investigating in a separate report. Third, and according to the sequential ignorability assumption[53], there is a chance that there may be other confounders that we did not assess that may explain the relation between the proposed mediator (in this case, patient activation) and depression outcomes. While our proposed mediator was selected apriori and based on the conceptual theory of behavioural activation, future studies considering additional mediators through, for example, comprehensive structural equation models are required to verify our findings and address the sequential ignorability assumption[54]. Lastly, we did not apply diagnostic criteria in recruiting patients at baseline or in our definition of outcome, but we note that the PHQ-9 is widely used to define case-level morbidity in trials and, importantly, we used locally validated cutoffs in this study[30].

Clinical implications and conclusions

In conclusion, our findings are consistent with the small but growing body of evidence suggesting an enduring effect for behavioural activation or more cognitive behavioural approaches [45, 51, 55]. HAP is unique in that, despite its brevity and delivery by a lay counsellor, it is able to sustain short-term gains in a LMIC primary care setting. In addition, HAP is only one of two[56] brief behavioural activation theory-based PT delivered by lay counsellors in primary care settings yet evaluated. The low levels of antidepressant medication (ADM) noted in our study, even after the diagnosis was conveyed to the primary care physician confirms that the effect of HAP could hot have been confounded by ADM use, and further supports the applicability of the HAP treatment in this treatment naïve population. The ecological validity of the trial was enhanced by the fact that the lay counsellors had no prior professional mental health training (as would be the case in most real-world settings) and that they were concurrently delivering a completely different PT for harmful drinking (as would be the case in actual practice) (Nadkarni et al, companion paper[57]). The importance of establishing sustained effects of treatments cannot be

overemphasized given depression tends to relapse or recur. We have demonstrated that brief psychological treatments like HAP and the Counselling for Alcohol Problems (CAP) program delivered by non-specialist mental health workers in routine primary care can have sustained clinical effects and are good value for the money. Such treatments are ideal for scaling up and future research should focus on: 1) employing SMART designs to assess how different interventions can be applied in sequence to achieve higher rates of remission and recovery[58]; and 2) examining the potential roles of multiple mediators within randomised trial designs so that the effectiveness of treatments can be enhanced through a focus on these mediators.

ACKNOWLEDGMENTS

We acknowledge the generous partnership and support from the Directorate of Health Services of the Government of Goa.

REFERENCES

1. Patel V, Chisholm D, Parikh R, Charlson FJ, Degenhardt L, Dua T, et al. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. Lancet. 2016;387(10028):1672-85.
2. Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010. PLoS One. 2015;10(2):e0116820.
3. Chisholm D, Sweeny K, Sheehan P, Rasmussen B, Smit F, Cuijpers P, et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. Lancet Psychiatry. 2016;3(5):415-24.
4. Gururaj G, Varghese M, Benegal V, Rao G, Pathak K, Singh L, et al. National Mental Health Survey of India, 2015-16: Summary. Bengaluru: National Institute of Mental Health and Neuro Sciences, NIMHANS Publication No. 128; 2016.
5. WHO. mhGAP Mental Health Gap Action Programme: Scaling up Care for Mental, Neurological and Substance Use Disorders Geneva: WHO; 2008 [Available from: http://www.who.int/mental_health/mhgap/evidence/en/.
6. Cuijpers P, Hollon SD, van Straten A, Bockting C, Berking M, Andersson G. Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy? A meta-analysis. BMJ Open. 2013;3(4).
7. Maziak W, Eissenberg T, Klesges RC, Keil U, Ward KD. Adapting smoking cessation interventions for developing countries: a model for the Middle East. Int J Tuberc Lung Dis. 2004;8(4):403-13.
8. Patel V. The need for treatment evidence for common mental disorders in developing countries. Psychol Med. 2000;30(4):743-6.
9. van Ginneken N TP LS, Rao GN, Meera SM, Pian J, Chandrashekar S, Patel V. Non-specialist health worker interventions for the care of mental, neurological and substance-abuse disorders in low- and middle-income countries. Cochrane Database of Systematic Reviews. 2013.
10. Singla DR, Kohrt, B.A., Murray, L.K., Anand, A., Chorpita, B.C., Patel, V. Psychological treatments for the world: Lessons from low- and middle-income countries. Annual Review of Clinical Psychology. 2017;13(15):1-5.
11. Chibanda D, Weiss HA, Verhey R, Simms V, Munjoma R, Rusakaniko S, et al. Effect of a Primary Care-Based Psychological Intervention on Symptoms of Common Mental Disorders in Zimbabwe: A Randomized Clinical Trial. JAMA. 2016;316(24):2618-26.
12. Nadkarni A, Velleman R, Dabholkar H, Shinde S, Bhat B, McCambridge J, et al. The systematic development and pilot randomized evaluation of counselling for alcohol problems, a lay counselor-delivered psychological treatment for harmful drinking in primary care in India: the PREMIUM study. Alcohol Clin Exp Res. 2015;39(3):522-31.
13. Chowdhary N, Anand A, Dimidjian S, Shinde S, Weobong B, Balaji M, et al. The Healthy Activity Program lay counsellor delivered treatment for severe depression in India: systematic development and randomised evaluation. Br J Psychiatry. 2016;208(4):381-8.
14. Vellakkal S, Patel V. Designing Psychological Treatments for Scalability: The PREMIUM Approach. PLoS One. 2015;10(7):e0134189.
15. Dimidjian S, Barrera Jr M, Martell C, Muñoz RF, Lewinsohn PM. The origins and current status of behavioral activation treatments for depression. Annual Review of Clinical Psychology. 2011;7:1-38.

16. Patel V, Weobong B, Weiss HA, Anand A, Bhat B, Katti B, 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.
17. Patel V, Weobong B, Nadkarni A, Weiss H, Anand A, Naik S, et al. The effectiveness and cost-effectiveness of lay counsellor-delivered psychological treatments for harmful and dependent drinking and moderate to severe depression in primary care in India: PREMIUM study protocol for randomized controlled trials. Trials. 2014;15(1):101.
18. Fournier JC, DeRubeis RJ, Shelton RC, Hollon SD, Amsterdam JD, Gallop R. Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. J Consult Clin Psychol. 2009;77(4):775-87.
19. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet. 2002;359(9306):614-8.
20. World Health Organisation. mhGAP intervention guide for mental, neurological and substance use disorders in non-specialized health settings: mental health Gap Action Programme (mhGAP). Geneva: WHO; 2010.
21. Patel V, Weobong B, Nadkarni A, Weiss HA, Anand A, Naik S, et al. The effectiveness and cost-effectiveness of lay counsellor-delivered psychological treatments for harmful and dependent drinking and moderate to severe depression in primary care in India: PREMIUM study protocol for randomized controlled trials. Trials. 2014;15(1):101.
22. Singla DR, Weobong B, Nadkarni A, Chowdhary N, Shinde S, Anand A, et al. Improving the scalability of psychological treatments in developing countries: an evaluation of peer-led therapy quality assessment in Goa, India. Behav Res Ther. 2014;60:53-9.
23. Cuijpers P, Karyotaki E, Weitz E, Andersson G, Hollon SD, van Straten A. The effects of psychotherapies for major depression in adults on remission, recovery and improvement: a meta-analysis. J Affect Disord. 2014;159:118-26.
24. Rush AJ, Kraemer HC, Sackeim HA, Fava M, Trivedi MH, Frank E, et al. Report by the ACNP Task Force on response and remission in major depressive disorder. Neuropsychopharmacology. 2006;31(9):1841-53.
25. McGlothlin AE, Lewis RJ. Minimal clinically important difference: defining what really matters to patients. JAMA. 2014;312(13):1342-3.
26. Button KS, Kounali D, Thomas L, Wiles NJ, Peters TJ, Welton NJ, et al. Minimal clinically important difference on the Beck Depression Inventory--II according to the patient's perspective. Psychol Med. 2015;45(15):3269-79.
27. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4):407-15.
28. Manos RC, Kanter JW, Luo W. The behavioral activation for depression scale-short form: development and validation. Behav Ther. 2011;42(4):726-39.
29. Kumar GS, Jain A, Hegde S. Prevalence of depression and its associated factors using Beck Depression Inventory among students of a medical college in Karnataka. Indian J Psychiatry. 2012;54(3):223-6.
30. Patel V, Araya R, Chowdhary N, King M, Kirkwood B, Nayak S, et al. Detecting common mental disorders in primary care in India: a comparison of five screening questionnaires. Psychol Med. 2008;38(2):221-8.
31. Patel V, Weiss HA, Chowdhary N, Naik S, Pednekar S, Chatterjee S, et al. Effectiveness of an intervention led by lay health counsellors for depressive and anxiety

disorders in primary care in Goa, India (MANAS): a cluster randomised controlled trial. Lancet. 2010;376(9758):2086-95.
32. Patel V, Chisholm D, Kirkwood BR, Mabey D. Prioritizing health problems in women in developing countries: comparing the financial burden of reproductive tract infections, anaemia and depressive disorders in a community survey in India. Trop Med Int Health. 2007;12(1):130-9.
33. Patel V, Chisholm D, Rabe-Hesketh S, Dias-Saxena F, Andrew G, Mann A. Efficacy and cost-effectiveness of drug and psychological treatments for common mental disorders in general health care in Goa, India: a randomised, controlled trial. Lancet. 2003;361(9351):33-9.
34. Chisholm D, Sekar K, Kumar KK, Saeed K, James S, Mubbashar M, et al. Integration of mental health care into primary care. Demonstration cost-outcome study in India and Pakistan. Br J Psychiatry. 2000;176:581-8.
35. Maselko J, Patel V. Why women attempt suicide: the role of mental illness and social disadvantage in a community cohort study in India. J Epidemiol Community Health. 2008;62(9):817-22.
36. Localio AR, Margolis DJ, Berlin JA. Relative risks and confidence intervals were easily computed indirectly from multivariable logistic regression. J Clin Epidemiol. 2007;60(9):874-82.
37. Sterne JA, Davey Smith G. Sifting the evidence-what's wrong with significance tests? BMJ. 2001;322(7280):226-31.
38. Mackinnon DP, Lockwood CM, Williams J. Confidence Limits for the Indirect Effect: Distribution of the Product and Resampling Methods. Multivariate Behav Res. 2004;39(1):99.
39. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. Annu Rev Psychol. 2007;58:593-614.
40. Preacher KJ SJ. Advantages of Monte Carlo confidence intervals for indirect effects. . Communication Methods and Measures 2012;6(2):77-98.
41. Buttorff C, Hock RS, Weiss HA, Naik S, Araya R, Kirkwood BR, et al. Economic evaluation of a task-shifting intervention for common mental disorders in India. Bull World Health Organ. 2012;90(11):813-21.
42. Kessler RC, Barber C, Birnbaum HG, Frank RG, Greenberg PE, Rose RM, et al. Depression in the workplace: effects on short-term disability. Health Aff (Millwood). 1999;18(5):163-71.
43. Hutubessy R, Chisholm D, Edejer TT. Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. Cost Eff Resour Alloc. 2003;1(1):8.
44. Implementation MoSaP. State Domestic Product and other aggregates, 2011–2012 series 2016 [Available from: http://mospi.nic.in/data
45. Dobson KS, Hollon SD, Dimidjian S, Schmaling KB, Kohlenberg RJ, Gallop RJ, et al. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. J Consult Clin Psychol. 2008;76(3):468-77.
46. Levinson D, Lakoma MD, Petukhova M, Schoenbaum M, Zaslavsky AM, Angermeyer M, et al. Associations of serious mental illness with earnings: results from the WHO World Mental Health surveys. Br J Psychiatry. 2010;197(2):114-21.
47. Layard R, & Clark, D. M. Thrive: The power of evidence-based psychological therapies.

. London, UK: Penguin Books; 2014.
48. Dimidjian S, Goodman SH, Sherwood NE, Simon GE, Ludman E, Gallop R, et al. A pragmatic randomized clinical trial of behavioral activation for depressed pregnant women. J Consult Clin Psychol. 2017;85(1):26-36.
49. Ryba MM, Lejuez CW, Hopko DR. Behavioral activation for depressed breast cancer patients: the impact of therapeutic compliance and quantity of activities completed on symptom reduction. J Consult Clin Psychol. 2014;82(2):325-35.
50. DeRubeis RJ, Hollon SD, Amsterdam JD, Shelton RC, Young PR, Salomon RM, et al. Cognitive therapy vs medications in the treatment of moderate to severe depression. Arch Gen Psychiatry. 2005;62(4):409-16.
51. Dimidjian S, Hollon SD, Dobson KS, Schmaling KB, Kohlenberg RJ, Addis ME, et al. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. J Consult Clin Psychol. 2006;74(4):658-70.
52. Möller HJ, Riedel M, Seemüller F. Relapse or recurrence in depression: why has the cutoff been set at 6 months? MEDICOGRAPHIA. 2011;33(2):125-31.
53. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychol Methods. 2010;15(4):309-34.
54. MacKinnon DP, Pirlott AG. Statistical approaches for enhancing causal interpretation of the M to Y relation in mediation analysis. Pers Soc Psychol Rev. 2015;19(1):30-43.
55. Wiles N, Lewis G, Peters T, Kuyken W, Williams C. Cognitive behavioural therapy for treatment-resistant depression - Authors' reply. Lancet. 2013;381(9880):1814-5.
56. Rahman A, Hamdani SU, Awan NR, Bryant RA, Dawson KS, Khan MF, et al. Effect of a Multicomponent Behavioral Intervention in Adults Impaired by Psychological Distress in a Conflict-Affected Area of Pakistan: A Randomized Clinical Trial. JAMA. 2016;316(24):2609-17.
57. Nadkarni A, Weiss, A.H, Weobong, B, MacCambridge, J, Bhat, B, Katti, B, et al. Sustained effectiveness and cost-effectiveness of the Counselling for Alcohol Problems Program, a brief psychological treatment for harmful drinking delivered by lay counsellors in primary care: twelve-month follow-up of a randomised controlled trial PLoS Med-provisional acceptance. 2017.
58. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A "SMART" design for building individualized treatment sequences. Annu Rev Clin Psychol. 2012;8:21-48.
59. Bedirhan Üstün T, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, et al. Developing the World Health Organization Disability Assessment Schedule 2.0 Bull World Health Organ 2010;88:815–23.
60. Chisholm D, Knapp MR, Knudsen HC, Amaddeo F, Gaite L, van Wijngaarden B. Client Socio-Demographic and Service Receipt Inventory--European Version: development of an instrument for international research. EPSILON Study 5. European Psychiatric Services: Inputs Linked to Outcome Domains and Needs. Br J Psychiatry Suppl. 2000(39):s28-33.

Supplementary tables and figures

S1 Table: Secondary outcomes at 12 months.

Secondary outcome	Measure of outcome
Recovery from depression	PHQ-9 score <5 at both 3 and 12 months
Full relapse	PHQ-9 score >14 at 12 months amongst patients with at least partial remission at 3 months.
Partial relapse	PHQ-9 score 10 to 14 at 12 months amongst patients with at least partial remission at 3 months.
Disability	Mean disability score on the WHO disability assessment schedule version 2 (WHO-DAS II[59])
Total days unable to work	Mean total days unable to work in the previous month on the WHO-DAS II.
Suicidal behaviour	Proportion reporting suicide thoughts in the last two weeks on the PHQ-9; proportion reporting any suicide attempts in the last 3 months
Intimate partner violence	Proportion reporting experience of intimate partner violence (physical/psychological/emotional) over the past 3 months.
Minimal Clinically Important Difference (MCID)	Change in PHQ-9 outcome score from baseline compared with the corresponding score on patient's subjective sense of improvement.
Any response over 12 months	PHQ-9 score <10 at either of the 3- and 12-month outcome assessment points (added post-hoc but before analysing the data).
Resource impacts for the health system	Estimates of cost-effectiveness/cost-saving using detailed electronic records on HAP delivery, as well as other use of primary and secondary care services collected from patients using the Client Service Receipt Inventory[60].

S2 Table: Comparison of participants who were followed up and those lost to follow up at 3 and 12 months

	Lost before 3-month evaluation N=27 (5.5%)	**Completed 3-month outcome evaluation N=466 (94.5%)**	**Lost before 12-month evaluation N=46 (9.3%)**	**Completed 12-month outcome evaluation N=447 (90.7%)**	**p-value (12-month follow up)**
*Age (years) (mean [SD])	36.2 (11.6)	42.9 (12.0)	35 (11.8)	43 (11.8)	p<0.001
Gender (Female) (n [%])	23 (85%)	356 (76%)	35 (76%)	344 (77%)	p=0.86
Marital status (n [%])					
Married	16 (59.3%)	321 (68.9%)	33 (71.7%)	304 (68.0%)	
Single	8 (29.6%)	49 (10.5%)	11 (23.9%)	46 (10.3%)	p=0.004
Separated/Divorced	1 (3.7%)	3 (0.6%)	0 (0.0%)	4 (0.9%)	
Widowed	2 (7.4%)	93 (20.0%)	2 (4.4%)	93 (20.8%)	
Education status (n [%])					
None	6 (22%)	124 (27%)	7 (15%)	123 (28%)	p=0.004
Primary	13 (48%)	236 (50%)	23 (50%)	226 (51%)	
Secondary	5 (19%)	73 (16%)	6 (13%)	72 (16%)	
Higher Secondary	2 (7%)	22 (5%)	6 (13%)	18 (4%)	

	Lost before 3-month evaluation N=27 (5.5%)	**Completed 3-month outcome evaluation N=466 (94.5%)**	**Lost before 12-month evaluation N=46 (9.3%)**	**Completed 12-month outcome evaluation N=447 (90.7%)**	**p-value (12-month follow up)**
Graduate/above	1 (4%)	11 (2%)	4 (9%)	8 (2%)	
Occupation (n [%])					
Unemployed	14 (52%)	278 (60%)	22 (48%)	270 (60%)	p=0.02
Unskilled manual labour	12 (41%)	162 (35%)	18 (39%)	156 (35%)	
Skilled manual labour	0 (0%)	7 (2%)	0 (0%)	7 (2%)	
Clerical & professional	1 (3%)	19 (4%)	6 (13%)	14 (3%)	
Patient's expectation of counselling (n [%])					
Not useful	0 (0%)	1 (0.2%)	0 (0%)	1 (0.2%)	p=0.13
A little/somewhat useful	8 (31%)	218 (47%)	14 (30%)	212 (47%)	
Moderately useful	8 (28%)	108 (23%)	13 (28%)	103 (23%)	
Very useful	11 (38%)	139 (30%)	19 (41%)	131 (29%)	
Chronicity of symptoms-wks (median [IQR])	4 (4-15)	12 (4-48)	9 (4-24)	12 (4-48)	p=0.22
Median PHQ score (median [IQR])	17 (15-18)	17 (16-20)	18 (15-19)	17 (16-20)	p=0.55
Mean PHQ-score (SD)	17.3 (1.9)	18.0 (2.8)	17.6 (2.4)	17.9 (2.7)	p=0.43
PHQ category (n [%])					
Score 15-19 (Mod. severe)	24 (90%)	348 (75%)	37 (80%)	335 (75%)	p=0.48
Score 20-27 (severe)	3 (10%)	118 (25%)	9 (20%)	112 (25%)	
Trial Arm					
EUC	12 (44%)	236(49%)	19 (41%)	229 (51%)	p=0.22
HAP+EUC	15 (56%)	230 (51%)	27 (59%)	218 (49%)	

905

S3 Table: Results of t-test and descriptive statistics for change in mean primary outcome score between 3 and 12 month endpoints by trial arm (complete case N=438)

Trial arm	Endpoint		95% CI for Mean Difference	t	df	p-value
	3 months Mean (SD)	12 months Mean (SD)				
BDI-II						
EUC	27.66 (13.27)	24.46 (14.66)	3.2 (1.34, 5.06)	3.39	224	p=0.001
HAP+EUC	19.64 (15.45)	19.97 (15.59)	-0.34 (-.2.37, 1.69)	-0.33	212	p=0.74

S4 Table: Effect of the HAP treatment plus EUC on scores for depression symptoms, disability, suicide behavior, and intimate partner violence over 9 months, based on complete case and random effects

SENSITIVITY ANALYSIS	EUC arm (n=229)	HAP+EUC arm (n=218)	[1]Adjusted mean difference (AMD), effect size (ER), prevalence ratio (PR), prevalence difference (PD) (95% CI)	p-value
PRIMARY OUTCOMES				
Complete case				
Mean BDI-II score (SD)	24.37 (14.65)	19.83 (15.56)	AMD: -4.57 (-7.34, -1.81) ES: 0.30 (0.11-0.48)	p=0.001
Remission: PHQ-9<10- no. (%)	107 (46.72%)	137 (62.84%)	PR: 1.36 (1.14, 1.61)	p=0.0004
Random effects				
Complete case adjusting for counsellor/PHC as random effect (BDI-II score)	24.37 (14.65)	19.83 (15.56)	AMD: -4.57(-7.34, -1.81) ES: 0.30 (0.11, 0.48)	p=0.001
Multiple imputation adjusting for counsellor/PHC as	24.09 (14.67)	19.73 (15.53)	AMD: -4.41(-7.21, -1.61) ES: 0.23 (0.17, 0.28)	p=0.002

SENSITIVITY ANALYSIS	**EUC arm (n=229)**	**HAP+EUC arm (n=218)**	**[1]Adjusted mean difference (AMD), effect size (ER), prevalence ratio (PR), prevalence difference (PD) (95% CI)**	**p-value**
random effect (BDI-II score)				
SECONDARY OUTCOMES				
Complete case				
Recovery: PHQ-9<5 at 3 and 12 months- no. (%)	48 (20.96%)	66 (30.28%)	PR: 1.44 (1.05, 1.97) PD: 9.27% (1.43%, 17.11%)	p=0.022 p=0.021
***Full relapse: PHQ-9 score>14- no. (%)	11 (4.80%)	19 (8.72%)	PR: 1.81 (0.88, 3.69)	p=0.11
***Partial relapse: PHQ-9 score>9<15- no. (%)	2 (2.62%)	18 (8.26)	PR: 3.15 (1.27, 7.79)	p=0.013
***Mean PHQ-9 score (SD)	10.57 (7.58)	8.19 (6.91)	AMD: -2.41 (-3.72, -1.09) ES: 0.33 (0.14, 0.51)	p<0.0001
Any response over 12 months no. (%)	129/240 (53.75)	182/235 (77.45)	PR: 1.44 (1.26, 1.26)	p<0.0001
Mean disability score (SD)	10.89 (9.22)	9.38 (9.61)	AMD: -1.58 (-3.33, 0.17) ES: 0.03 (-0.03, 0.8)	p=0.08
Mean days unable to work (SD)	6.05 (8.81)	4.81 (8.24)	AMD: -1.29 (-2.89, 0.31) ES: 0.09 (0.04, 0.15)	p=0.12
Suicidal behaviour (Suicide thoughts) – no. (%)#	61/229 (26.64)	41/218 (18.81)	PR: 0.70 (0.49, 0.99)	p=0.046
Intimate partner physical violence##– females no. (%)	19/116 (16.38)	10/103 (9.71)	PR: 0.59 (0.29, 1.21)	p=0.149
Intimate partner psychological/emotional violence## –	39/116 (33.62)	27/103 (26.21)	PR: 0.75 (0.49, 1.13)	p=0.173

SENSITIVITY ANALYSIS	**EUC arm (n=229)**	**HAP+EUC arm (n=218)**	**[1]Adjusted mean difference (AMD), effect size (ER), prevalence ratio (PR), prevalence difference (PD) (95% CI)**	**p-value**
females no. (%)				
Intimate partner psychological/emotional violence## – males no. (%)	11/39 (28.21)	6/32 (18.75)	PR: 0.36 (0.37, 1.94)	p=0.711
$MCID (% reduction in baseline PHQ-9 score)	93 (40.61)	125 (57.60)	PR: 1.21 (1.05, 1.39)	p=0.009
Random effects				
Complete case adjusting for counsellor as random effect (PHQ-9 score)	10.57 (7.58)	8.19 (6.91)	AMD: -2.41 (-3.72, -1.09) ES: 0.33 (0.14, 0.51)	p<0.0001
Multiple imputation adjusting for counsellor as random effect (PHQ-9 score)	10.46 (7.54)	8.16 (6.96)	AMD: -2.34 (-3.67, -1.00) ES: 0.37 (0.31, 0.42)	p<0.0001
Complete case adjusting for counsellor as random effect (Mean disability score)	11.05 (9.22)	9.43 (9.62)	AMD: -1.64 (-3.34, 0.05) ES: 0.17 (-0.01, 0.36)	p=0.057
Multiple imputation adjusting for counsellor as random effect (Mean disability score)	10.89 (9.22)	9.38 (9.61)	AMD: -1.55 (-3.29, 0.19) ES: 0.03 (-0.03, 0.08)	p=0.082
Complete case adjusting for counsellor as random effect (Mean days	6.14 (8.83)	4.81 (8.21)	AMD: -1.31 (-2.86, 0.23) ES: 0.16 (-0.03, 0.34)	p=0.096

SENSITIVITY ANALYSIS	EUC arm (n=229)	HAP+EUC arm (n=218)	[1]Adjusted mean difference (AMD), effect size (ER), prevalence ratio (PR), prevalence difference (PD) (95% CI)	p-value
unable to work)				
Multiple imputation adjusting for counsellor as random effect (Mean days unable to work)	6.05 (8.81)	4.81 (8.24)	AMD: -1.26 (-2.86, 0.33) ES: 0.09 (0.04, 0.15)	p=0.121

Note:

[1] Adjusted for PHC as a fixed effect and PHQ-9 baseline score

***Not previously specified in trials protocol but specified in published analysis plan. #Suicidal thoughts over the past two weeks were assessed through the relevant PHQ-9 item while suicide attempts were assessed over the 3-month period leading up to the 12 month outcome follow up assessment. Attempts not included as numbers very small (only 2 patients (1 in each arm) reported suicide attempt over the period). ## Among married participants. $Minimal Clinically Important Difference: estimated based on relative difference in baseline and outcome score, and how this compares with overall subjective global rating of 'feeling better' at the end of the trial. The optimal cut-off in relative change in score with maximum specificity (>70%) is 55%.

S5 Table: Interaction effect of baseline depression severity, gender, chronicity of depression, and expectations of treatment, on the effect of HAP plus EUC on scores for depression symptoms (BDI-II outcome)

[1] Adjusted for PHC as a fixed effect and PHQ-9 baseline score

Analysis	**EUC arm (n=248)**	**HAP+EUC arm (n=245)**	**[1]Adjusted mean difference (95% CI)**	**p-value**
Baseline severity			p-effect modification = 0.227	
Moderate	22.77 [14.28]	19.69 [15.72]	-3.54 (-6.85, -0.22)	p=0.037
Severe	28.59 [15.17]	19.42 [15.04]	-8.41 (-14.29, -2.53)	p=0.006
Gender			p-effect modification= 0.857	
Males (mean score[SD])	24.28 [14.84]	20.07 [13.60]	-3.64 (-9.44, 2.17)	p=0.216
Females (mean score[SD])	24.24 [14.69]	19.62 [16.15]	-4.90 (-8.26, -1.55)	p=0.004
Chronicity			p-effect modification=0.181	
<12 weeks	21.76 [14.78]	18.68 [15.28]	-2.89 (-6.87, 1.10)	p=0.155
>=12 weeks	27.59 [14.04]	20.59 [15.74]	-6.86 (-10.86, -2.86)	p=0.001
Expectation			p-effect modification=0.629	
Not or somewhat useful	23.70 [13.96]	19.69 [15.28]	-4.15 (-8.11, -0.19);	p=0.040
Moderate or very useful	24.38 [15.58]	20.00 [16.04]	-5.32 (-9.36, -1.29);	p=0.010

S6 Table: SAEs and medication use by arm in the last 3 months

SAE/psychotropic medication	**EUC number of SAEs (No. of participants)**	**HAP+EUC Number of SAEs (No. of participants)**	**p-value**
SAEs			
Total SAEs	29 (34)	17 (18)	p=0.12
Death	2 (2)	0 (0)	p=0.49
Suicide attempt	1 (1)	1 (1)	p=1.00
Unplanned hospitalisation	26 (31)	18 (17)	p=0.26
Psychotropic medication	**No. of participants**	**No. of participants**	
Total psychotropic medication use	11	7	p=0.47

S7 Table. Mediation results examining patient-reported activation levels at 3-months on 12-month depression outcomes.

Assessment point	BA score (imputed data)		Model	Regression Result		Bootstrap 95% CI
	EUC (n=248)	HAP+EUC (n=245)		B*	SE	
3 months mean (SD)	9.81 (4.31)	12.01 (4.71)	*c*' (HAP+EUC → BDI-II at 12-months)	-4.46***	0.79	(-6.01, -2.91)
12 months	10.02 (4.64)	11.00 (4.49)	*a*' (HAP+EUC → activation at 3-months)	2.23***	0.23	(1.77, 2.68)
			b' (activation at 3-months → BDI-II at 12 months)	-1.17***	0.09	-1.35, -1.00)
			a x b	-2.62***	0.33	(-3.28, -1.97)

Note: *Beta estimates are unstandardized. Multiple linear regression models controlled for baseline PHQ-9 scores, participant age, and PHC). *p<0.05. **p<0.01. ***p<0.001
c' total effect; a x b: indirect effect

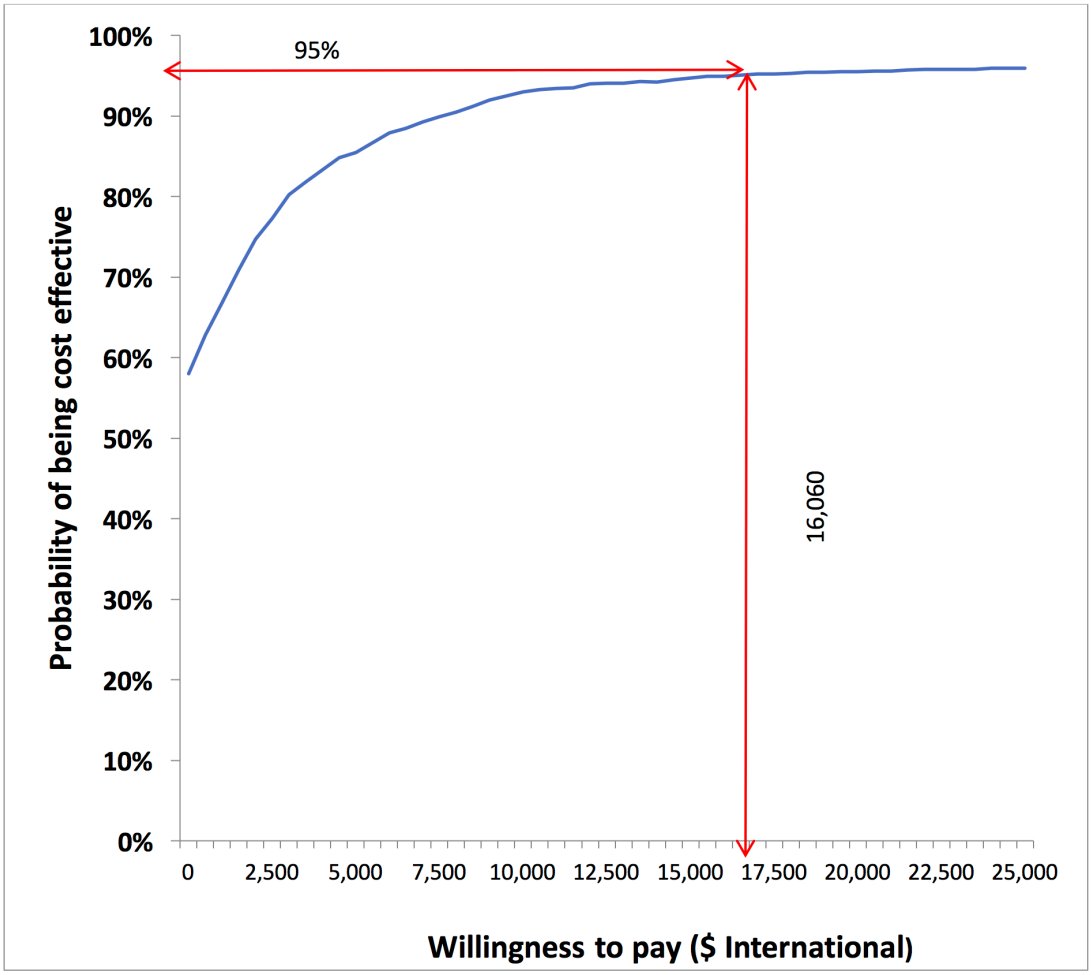
S8 Table: Mean costs (2015 International Dollars) and QALYs gained per person over 12 months

Type of Cost	HAP+EUC arm (n=245)	EUC arm (n=248)	Mean Difference (95% CI)	p-value
HAP intervention costs				
HAP Intervention (SE)	65.66 (3.48)	0 (0)	65.66 (58.80, 72.52)	0.000
Health Service Utilisation				
PHC Doctor Consultations (SE)	51.64 (3.86)	58.77 (6.23)	-7.13 (-21.54, 7.28)	0.331
Hospital Doctor Consultations (SE)	84.06 (14.51)	116.96 (47.55)	-32.90 (-130.75, 64.95)	0.509
Hospital Admissions (SE)	19.92 (5.32)	39.08 (9.12)	-19.16 (-39.92, 1.60)	0.070
Laboratory Tests (SE)	23.91 (2.99)	39.08 (6.55)	-15.16 (-29.32, -1.01)	0.036
Medicines (SE)	24.62 (2.91)	34.39 (5.07)	-9.77 (-21.27, 1.73)	0.096
Total Health Service Utilisation Costs (SE)	204.15 (19.56)	288.27 (50.85)	-84.12 (-191.32, 23.07)	0.124
Total Health System Costs				
Total Health System Costs (SE)	269.81 (19.53)	288.27 (50.85)	-18.47- (-125.64, 88.71)	0.735
Productivity Costs				
Time costs to service users and families (SE)	164.70 (12.89)	154.89 (12.77)	9.81 (-25.83, 45.46)	0.589
Productivity losses (SE)	344.95 (24.85)	491.22 (26.80)	-146.28 (-218.08, -74.47)	0.000

Total Societal Costs				
Societal perspective (SE)	779.46 (40.84)	934.39 (64.81)	-154.93 (-305.51, -4.35)	0.044
QALYs				
QALYs gained (SE)	0.848 (0.005)	0.837 (0.004)	0.011 (0.006, -0.002)	0.092

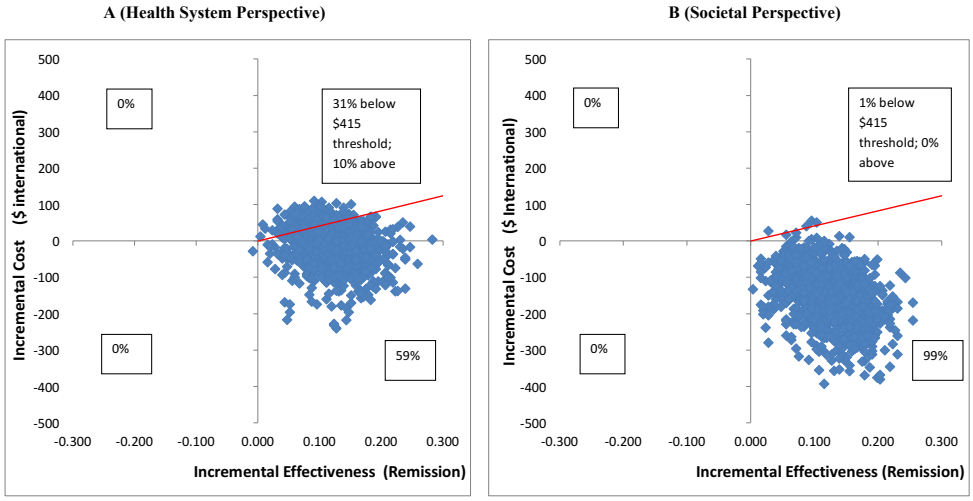
948
949
950
951

S1 Figure: Cost-effectiveness acceptability curve: willingness to pay per QALY gained from HAP from a health system perspective.



952

S2 Figure: Cost effectiveness planes: HAP plus EUC compared to EUC per remission achieved



953
954
955