

# Sussex Research

## Behavioural activation versus treatment as usual for depressed older adults in primary care: a pragmatic cluster-randomised controlled trial

NP Janssen, P Lucassen, MJH Huibers, D Ekers, T Broekman, JE Bosmans, Harm van Marwijk, J Spijker, R Oude Voshaar, GJ Hendriks

### Publication date

29-06-2023

### Licence

This work is made available under the [CC BY 4.0](#) licence and should only be used in accordance with that licence. For more information on the specific terms, consult the repository record for this item.

### Document Version

Published version

### Citation for this work (American Psychological Association 7th edition)

Janssen, N., Lucassen, P., Huibers, M., Ekers, D., Broekman, T., Bosmans, J., van Marwijk, H., Spijker, J., Oude Voshaar, R., & Hendriks, G. (2023). *Behavioural activation versus treatment as usual for depressed older adults in primary care: a pragmatic cluster-randomised controlled trial* (Version 1). University of Sussex. <https://hdl.handle.net/10779/uos.24014973.v1>

### Published in

Psychotherapy and Psychosomatics

### Link to external publisher version

<https://doi.org/10.1159/000531201>

### Copyright and reuse:

This work was downloaded from Sussex Research Open (SRO). This document is made available in line with publisher policy and may differ from the published version. Please cite the published version where possible. Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners unless otherwise stated. For more information on this work, SRO or to report an issue, you can contact the repository administrators at [sro@sussex.ac.uk](mailto:sro@sussex.ac.uk). Discover more of the University's research at <https://sussex.figshare.com/>

# Behavioural Activation versus Treatment as Usual for Depressed Older Adults in Primary Care: A Pragmatic Cluster-Randomised Controlled Trial

Noortje P. Janssen<sup>a,b,c</sup> Peter Lucassen<sup>b</sup> Marcus J.H. Huibers<sup>d</sup> David Ekers<sup>e</sup>  
Theo Broekman<sup>f</sup> Judith E. Bosmans<sup>g</sup> Harm van Marwijk<sup>h</sup> Jan Spijker<sup>a,c</sup>  
Richard C. Oude Voshaar<sup>i</sup> Gert-Jan Hendriks<sup>a,c,j</sup>

<sup>a</sup>Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands; <sup>b</sup>Department of Primary and Community Care, Research Institute of Health Sciences, Radboud University Medical Centre Nijmegen, Nijmegen, The Netherlands; <sup>c</sup>Institute for Integrated Mental Health Care Pro Persona, Nijmegen, The Netherlands; <sup>d</sup>NPI Centre for Personality Disorders/Arkin Amsterdam, Amsterdam, The Netherlands; <sup>e</sup>Mental Health and Addictions Research Group, Tees Esk and Wear Valleys NHS FT/University of York, York, UK; <sup>f</sup>Bureau Bêta, Nijmegen, The Netherlands; <sup>g</sup>Department of Health Sciences, VU University, Amsterdam, The Netherlands; <sup>h</sup>Department of Primary Care and Public Health, Brighton and Sussex Medical School, Brighton, UK; <sup>i</sup>Department of Psychiatry, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands; <sup>j</sup>Department of Psychiatry, Radboud University Medical Centre, Nijmegen, The Netherlands

## Keywords

Behavioural activation · Depression · Older adults · Primary care

## Abstract

**Introduction:** Effective non-pharmacological treatment options for depression in older adults are lacking. **Objective:** The effectiveness of behavioural activation (BA) by mental health nurses (MHNs) for depressed older adults in primary care compared with treatment as usual (TAU) was evaluated. **Methods:** In this multicentre cluster-randomised controlled trial, 59 primary care centres (PCCs) were randomised to BA and TAU. Consenting older ( $\geq 65$  years) adults ( $n = 161$ ) with clinically relevant symptoms of depression (PHQ-9  $\geq 10$ ) participated. Interventions were an 8-week individual MHN-led BA programme and unrestricted TAU in which general

practitioners followed national guidelines. The primary outcome was self-reported depression (QIDS-SR16) at 9 weeks and 3, 6, 9, and 12-month follow-up. **Results:** Data of 96 participants from 21 PCCs in BA and 65 participants from 16 PCCs in TAU, recruited between July 4, 2016, and September 21, 2020, were included in the intention-to-treat analyses. At post-treatment, BA participants reported significantly lower severity of depressive symptoms than TAU participants (QIDS-SR16 difference =  $-2.77$ , 95% CI =  $-4.19$  to  $-1.35$ ),  $p < 0.001$ ; between-group effect size = 0.90; 95% CI = 0.42–1.38). This difference persisted up to the 3-month follow-up (QIDS-SR16 difference =  $-1.53$ , 95% CI =  $-2.81$  to  $-0.26$ ,  $p = 0.02$ ; between-group effect size = 0.50; 95% CI = 0.07–0.92) but not up to the 12-month follow-up [QIDS-SR16 difference =  $-0.89$  ( $-2.49$  to 0.71)],  $p = 0.28$ ; between-group effect size = 0.29 (95% CI =  $-0.82$  to 0.24). **Conclusions:** BA led to a greater symptom reduction of depressive symptoms

in older adults, compared to TAU in primary care, at post-treatment and 3-month follow-up, but not at 6- to 12-month follow-up.

© 2023 The Author(s).  
Published by S. Karger AG, Basel

## Introduction

Depressed older adults ( $\geq 65$  years) often do not receive adequate psychological care [1]. Research shows that older adults have less access to psychological treatment than younger adults, even though its effectiveness does not differ between age-groups [1, 2]. Many older adults report not knowing where to get help or feel like they should be able to solve their problems themselves [3]. Additionally, older adults have a 2–3 times lower chance of being referred for a guideline-focused psychological treatment [4]. Even though the majority of older adults with depression would rather receive psychological than pharmacological treatment, most patients are treated with antidepressants [5, 6]. Antidepressants often take several weeks to show effect, and their use is associated with risks and adverse events, such as falls and possible adverse interactions with two-thirds of the medications prescribed to older adults [7].

Older people tend to visit their general practitioner (GP) frequently, making the primary care setting a practical option for increasing accessibility of mental health care [8]. Despite the common constraints of limited time, competing priorities, and limited availability of mental health experts in primary care [9], behavioural activation (BA) may still be a viable option due to its simplicity and feasibility [10]. Within BA, the therapist and patient work together to create a personal environment of positive reinforcement by increasing functional and pleasurable behaviour and decreasing avoidant and depressed behaviour [11]. BA is as effective as cognitive behavioural therapy and more effective than control conditions in reducing depression in adults [12, 13]. BA uses behavioural techniques, leaving out the more complicated cognitive techniques, and thus can be delivered by less specialised health practitioners such as mental health nurses (MHNs) [10]. It is feasible, effective, and cost-effective in managing depression in primary care [14].

In their meta-analysis, Orgeta and colleagues conclude that BA might also be effective for treating older adults with depression but that sample sizes were too small to draw firm conclusions [15]. Several studies also showed that collaborative care, which included telephone-

delivered or face-to-face-delivered BA, was more effective in treating and preventing major depression, dysthymic disorder, and clinically relevant depressive symptoms in older adults than usual primary care [16–19]. Nonetheless, adequately powered real-world RCTs investigating the effectiveness of BA as a standalone therapy for depressed older adults are lacking. Therefore, we aimed to evaluate the effectiveness of BA delivered by MHNs, compared to treatment as usual (TAU) for older adults with moderate to severe depressive symptoms in primary care. A cluster-randomised design was chosen because of possible contamination of knowledge by participating MHNs. There was no other way to ensure that the therapy was delivered by separate mental health professionals across treatment arms [20]. Our primary hypothesis was that BA leads to a greater reduction of depressive symptoms at the patient level compared to TAU after treatment and at subsequent follow-ups. Results regarding cost-effectiveness, moderators, and mediators of BA will be described elsewhere.

## Materials and Methods

### *Trial Design*

We performed a multicentre, cluster-randomised trial in primary care with two treatment arms and a 12-month follow-up period. Clusters were PCCs which were recruited by contacting two umbrella organisations (networks of several PCCs sharing the same pool of MHNs) as well as standalone PCCs in the Radboudumc network. Two umbrella organisations (Zorroo [19 PCCs] and NEO [24 PCCs]) joined as well as 16 standalone practices. PCCs were eligible when they had at least one MHN employed (which applies to almost all PCCs in the Netherlands). Details about the design have been described elsewhere [21]. This study was funded by ZonMw (843001606) and registered at the Dutch Trial Register (NL5436)<sup>1</sup> retrospectively, 2 months after recruitment started, due to an administrative error. At that point in time, ten participants were recruited, but no changes in the trial design have been made after the start of recruitment.

### *Participants*

Participants were older adults ( $\geq 65$  years) who presented with depressive symptoms at their general practice either spontaneously or after reading an information leaflet about the ongoing trial. Inclusion criteria were aged 65 years or older and current clinically relevant depressive symptoms as measured with the Patient Health Questionnaire 9 (PHQ-9  $\geq 10$ ) [22], the shortest validated screening instrument for older adults in primary care [23]. Exclusion criteria were current severe mental illness (except severe depression), high risk of suicide, drug and/or alcohol abuse in need of specialised treatment, as assessed with the Mini International Neuropsychiatric Interview (MINI5.0.0) [24], psychotherapy in

<sup>1</sup>Formerly known as Netherlands Trial Register NTR6013.

the previous 12 weeks or current treatment by a mental health specialist, as confirmed by the GP, and moderate to severe cognitive impairment, as measured with the Montreal Cognitive Assessment (MoCA < 18) [25].

In case the MINI 5.0.0 results suggested a severe mental condition, an old-age psychiatrist as well as the patients' GP were consulted to determine whether participating in the study and thereby withholding another specialised treatment would be damaging to the patient. Only patients in need of specialised care were excluded. Physical illness, comorbid psychological disorders that were not in need of specialised treatment, disability, or mild cognitive impairment were not exclusion criteria. Additionally, illiteracy or a nonperfect understanding of Dutch language were not exclusion criteria as independent, trained, and blinded research assistants offered assistance with filling in questionnaires at participants' homes. These assistants were employed as quality officers at the research department of a large mental health institute (Pro Persona) and were neither involved in the study design nor the analyses. Patients with antidepressants were eligible, provided that a stable dose had been maintained for at least 12 weeks before participating in the study.

#### *Randomisation and Masking*

PCCs were randomly allocated to BA or TAU, taking into account the number and sex of the employed MHNs, the size of the patient population, and affiliation with an umbrella organisation. An independent statistician was provided an anonymised list of the abovementioned characteristics per PCC, without any traceable data and created two equal groups of PCCs based on these characteristics. Computer generated random numbers were used to assign these groups to conditions (BA or TAU). The allocation sequence was concealed until clusters were enrolled and assigned to one of the interventions. Participants and MHNs were not blinded to treatment allocation due to the nature of the intervention. Research assistants involved in collection of outcome assessment were blinded to allocation.

#### *Procedures*

Older adults that visited their GP or MHN with depressive symptoms, complaints of loneliness, or medically unexplained symptoms were informed about the study with an information letter. After their consent, they were asked whether they agreed to be contacted by our research assistant. Some older adults visited their GP after having read posters, newsletters, or information booklets about the study in the waiting room. These eligible participants were possibly referred by their GP to the research assistant that contacted them to plan a home visit to check for eligibility and inform them about the allocated treatment option in their PCC (BA or TAU).

During this baseline home visit, an experienced and trained research assistant conducted the MINI5.0.0 and cognitive assessment. Participants received the baseline questionnaire (online or on paper), which they were asked to complete within the following week. In case of illiteracy, language barriers, or difficulty answering questions, an extra meeting with a different, independent research assistant was planned to assist with the questionnaires. Approximately 1 week after the baseline meeting, treatment started. Nine weeks after the baseline meeting, the post-treatment home visit and questionnaires were planned. In the follow-up period

(3–12 months after the post-treatment meeting), participants received a follow-up questionnaire every 3 months, either on paper or online. Questionnaires included questions about depressive symptoms, possible mediators, and a cost-effectiveness measure. A detailed description of all used measures can be found elsewhere [21]. Participants received assistance with filling in questionnaires, if necessary. GPs informed the research team immediately about any serious adverse events (SAEs) such as hospitalisation or death during the study and 12-month follow-up. Research assistants asked whether SAEs had taken place at follow-up appointments. SAEs were reported to the sponsor as well as the medical ethical committee. Non-serious adverse events were not collected in agreement with the medical ethical committee due to an expected high frequency in the older population and limited likelihood these were associated with the intervention.

#### *Interventions*

In the intervention arm consisting of 30 PCCs, BA was delivered in one 45-min face-to-face session followed by seven 30-min face-to-face sessions. All key elements of the original BA protocol of Martell et al. were included, such as functional analysis, activity registration, activity scheduling, and relapse prevention [11]. BA was delivered by primary care MHNs at the PCC or at participants' homes. All MHNs had at least a bachelor degree in nursing and were educated about different treatment techniques such as motivational interviewing and problem solving but not of protocolised therapies for mental health problems such as cognitive behavioural therapy or BA for depression [26]. All MHNs ( $n = 29$ ) in the BA arm received a 2-day training by licensed specialists in the delivery of the BA protocol. They received biweekly online supervision in small groups. MHNs filled in a session checklist for every patient to check for therapy adherence to the BA protocol after every session. Before the start of treatment, participants were made aware that BA consisted of a limited number of 8 weekly sessions and that the goal was to finish treatment within these 8 sessions. During the course of the intervention, no additional mental health interventions were delivered to the BA group. Any prescribed antidepressant medication remained on a stable dose.

In the TAU arm, consisting of 29 PCCs, treatment options were consistent with the guidelines of the Dutch College of GPs [27]. Health professionals in the TAU arm were free to determine the frequency and duration of care. Usual MHN support consists of 30-min eclectic counselling sessions.

#### *Outcomes*

##### *Depressive Symptoms*

The Quick Inventory of Depressive Symptomatology self-report 16 items (QIDS-SR16) was the primary outcome measure of severity of depressive symptoms at patient level at baseline, post-treatment (9 weeks), and all follow-ups (3, 6, 9, and 12 months after post-treatment) [28]. The QIDS-SR16 is a self-report instrument assessing depressive symptoms during the last 2 weeks with a score range of 0–27 with higher scores indicating more depressive symptoms. It is widely used, can be administered in 5–7 min, and has good psychometric properties [28].

The PHQ-9 was used to assess the eligibility of potential participants and was also used as a secondary outcome to assess changes in depressive symptoms over the course of treatment (weeks 2, 4, and 7), post-treatment (week 9), and all follow-ups.

The PHQ-9 is a 9-item self-report instrument that assesses depressive symptoms during the last 2 weeks. The psychometric properties of the PHQ-9 are good. The PHQ-9 has a score range from 0 to 27. Patients scoring PHQ  $\geq$  10 can be classified as having clinically relevant depressive symptoms [29].

#### Depressive Disorder and Psychiatric Comorbidity

A trained research assistant administered the complete M.I.N.I.5.0.0 to distinguish patients with a major depressive disorder from patients with only clinically relevant symptoms and assess psychiatric comorbidity at baseline. At post-treatment and 6- and 12-month follow-up, only the section about depression was administered to determine whether participants had a major depressive disorder. The M.I.N.I.5.0.0 is a short diagnostic interview that assesses current psychological disorders based on DSM-IV diagnostic criteria.

#### Statistical Analyses

Prior to the study, we calculated a sample size of  $n = 200$  ( $\alpha = 0.05$ , power  $[1 - \beta] = 0.80$ , 2-tailed test) with an expected effect size of 0.50 based on available data from studies with younger participants [21]. Since the COVID-19 pandemic posed the risk that we had to discontinue the trial prematurely, we performed an interim analysis on the 161 participants to determine the need to restart recruitment after the first lockdown. We took into account the actual correlation between pre- and post-treatment (0.468), the design effect based on actual ICC (0.017) and cluster size (4.35), actual dropout during the treatment phase (17.5%), and cluster variation (0.275), which led to a necessary number of 137 participants ( $\alpha = 0.05$ , power  $[1 - \beta] = 0.80$ , 2-tailed test) to be able to detect an effect size of 0.50. The new calculation was conducted by an independent statistician.

Baseline characteristics of the participants and the PCCs were summarised using the means for continuous data and percentages for categorical data. We used the R software version 4.1.1 [30] within RStudio version 2021.09.0 + 351a, specifically the *lme4* and *lmerTest* packages [31], to model the differences in the course of depressive symptoms between BA and TAU with linear mixed-effects models. Prior to entering treatment effects in the models, we fitted two models to examine whether the PCC level added additional predictive power to the model (see online suppl. material; for all online suppl. material, see <https://doi.org/10.1159/000531201>). To model the non-linear effect of treatment on depression over time, while accounting for the intra-subject correlation between the different time points, we used flexible functions known as regression splines, including pre-treatment assessment of depression as a covariate and the random part of the model with one internal knot placed at the 9-week mark of end of treatment. We modelled the effect of time using linear and quadratic splines and selected the models with the best fit based on the Akaike information criterion values and Bayesian information criteria [32]. Fixed effects were decomposed by ANOVA and evaluated by the F-statistic. Because a quadratic spline with one knot is difficult to interpret, we computed estimated marginal means (EMMs) using the R *emmeans* package [33] and examined whether the differences between the groups were significant at post-treatment and 3, 6, 9, and 12 months after post-treatment by performing *t* tests on these EMMs and computing standardised between and within effect sizes (Cohen's *d*). For the effect size calculation, we used the estimation of the SD of the intercept in our model but without the

baseline covariate of the outcome because the SD of the intercept is nearly zero when this covariate is added.

The intention-to-treat analyses included all participants, regardless of the number of sessions received. The use of mixed models with time, treatment, interaction between time and treatment, and baseline covariate as parameters obviates separate multiple imputation methods for missing data [34]. We replicated our final model with the secondary outcome measure (PHQ-9) which also included measurements at weeks 2, 4, and 7 of the treatment phase.

We determined the proportion of patients that showed reliable and clinically significant improvement and deterioration on the outcome measures at the 12-month follow-up, based on the model of Jacobson and Truax [35, 36]. Relapse and recurrence were assessed with survival analysis (Cox proportional hazards regression), using treatment group as a covariate as well as any covariates that significantly influenced outcomes in the main analysis. Relapse was defined as the return of clinically relevant depressive symptoms (PHQ-9  $\geq$  10) once remission (PHQ  $\leq$  5) had occurred – but before recovery (minimum of 4 months with PHQ  $\leq$  5) had been attained [37]. Recurrence referred to a return of clinically relevant depressive symptoms (PHQ-9  $\geq$  10) after recovery had taken place [37]. We utilised the PHQ-9 (relapse or recurrence of clinically relevant depressive symptoms) as it was used as an inclusion criterion, while MINI.5.0.0 (presence or absence of a depressive disorder) was not.

For the per protocol subgroup analysis, we used the final model with a subset of the data. We defined BA delivered “per protocol” as completion of at least five out of eight sessions before the end of treatment measurement (9 weeks) while adhering to the BA protocol in these sessions, as registered by MHNs in a provided checklist. When all treatment steps were checked in the checklist, the session was counted as “per protocol” session. Not fulfilling these per protocol requirements was regarded as preliminary discontinuation of treatment. Data of all TAU participants were used in the per protocol model since the requirements of TAU treatment were not defined. Furthermore, we analysed whether baseline characteristics were related to per protocol treatment completion within the BA group, as tested with *t* tests and Pearson's  $\chi^2$  tests.

To investigate whether “dropout” influenced treatment outcome, dropout was inserted as a binary variable in the final model. This variable indicated whether a participant did or did not finish the follow-up measurement of 12 months after post-treatment. We also analysed whether baseline characteristics of participants were related to later dropout, as tested with *t* tests and Pearson's  $\chi^2$  tests.

Lastly, we performed some ancillary analyses. We performed a “COVID” analysis to examine the effect of the COVID-19 period on our results. In this analysis, a binary variable stating whether the measurement point was before or during the pandemic was added as a factor in the final model. Furthermore, we investigated whether age, education, cognitive impairment, sex, and help with filling in questionnaires were related to treatment outcome when added as a factor in the final model.

## Results

We recruited 27 PCCs before the start of the study, and because of concerns that the incidence rate would be

lower than expected, we recruited the remaining 32 PCCs in two subsequent recruitment waves. Between July 4, 2016, and September 21, 2020, we assessed 395 older adults for eligibility. Of the 212 potential participants in the BA condition, 96 (45.2%) met the eligibility criteria and gave informed consent. In the TAU arm, 183 potential participants were screened, of whom 65 (35.5%) met eligibility criteria and gave informed consent. The twenty-two centres that did not recruit any participants were considered “lost to follow-up,” even though they did not formally drop out of the study. At any given time during the study period, we had missing data for 43/96 (44.8%) participants in the BA arm and 20/65 (30.8%) participants in the TAU arm (shown in Fig. 1). Eight participants received BA treatment at home rather than at the GP’s office. Three SAEs were reported which resulted in death. SAEs were not related to treatment or to depressive symptoms.

Baseline characteristics are shown in Table 1. Participants had a mean age of 75.2 years ( $SD = 7.0$ ) and were predominantly female (60.2%). One in three participants had a comorbid anxiety disorder. The average duration of their depressive symptoms was 9.7 years ( $SD = 16.6$ ), and more than half of the participants fulfilled the criteria for major depressive disorder. Out of all participants, 69% had pre-existing symptoms that were known and formerly treated by their GP, while 26% reported experiencing an incident depression. Details on the treatment history of the remaining 5% were missing. On average, participants used 4.8 medications prescribed by a medical professional for physical conditions and 0.6 medications for mental health conditions. Thirty-five (21.7%) participants needed help filling in the questionnaires. There were no significant baseline differences between groups. A detailed description of baseline characteristics can be found in the online supplementary material.

#### *BA Adherence and Actual TAU*

In the BA group, 62 participants (65%) received at least 5 protocolised sessions within 8 weeks were considered per protocol treatment completers. On average, BA participants received 5.8 sessions ( $SD = 3.3$ ). The treatment adherence of conducted sessions recorded by MHN checklists varied between 97% in session 1 to 84% in session 8. TAU consisted of sessions with an MHN (50.6%), antidepressants (14.3%), watchful waiting (7.8%), sessions with a GP (6.5%), psychotherapy (3.9%), mindfulness (2.6%), and physiotherapy (1.4%). For 10 participants (12.9%), received treatment modalities were not registered. Some participants received more than one treatment option.

#### *Main Results*

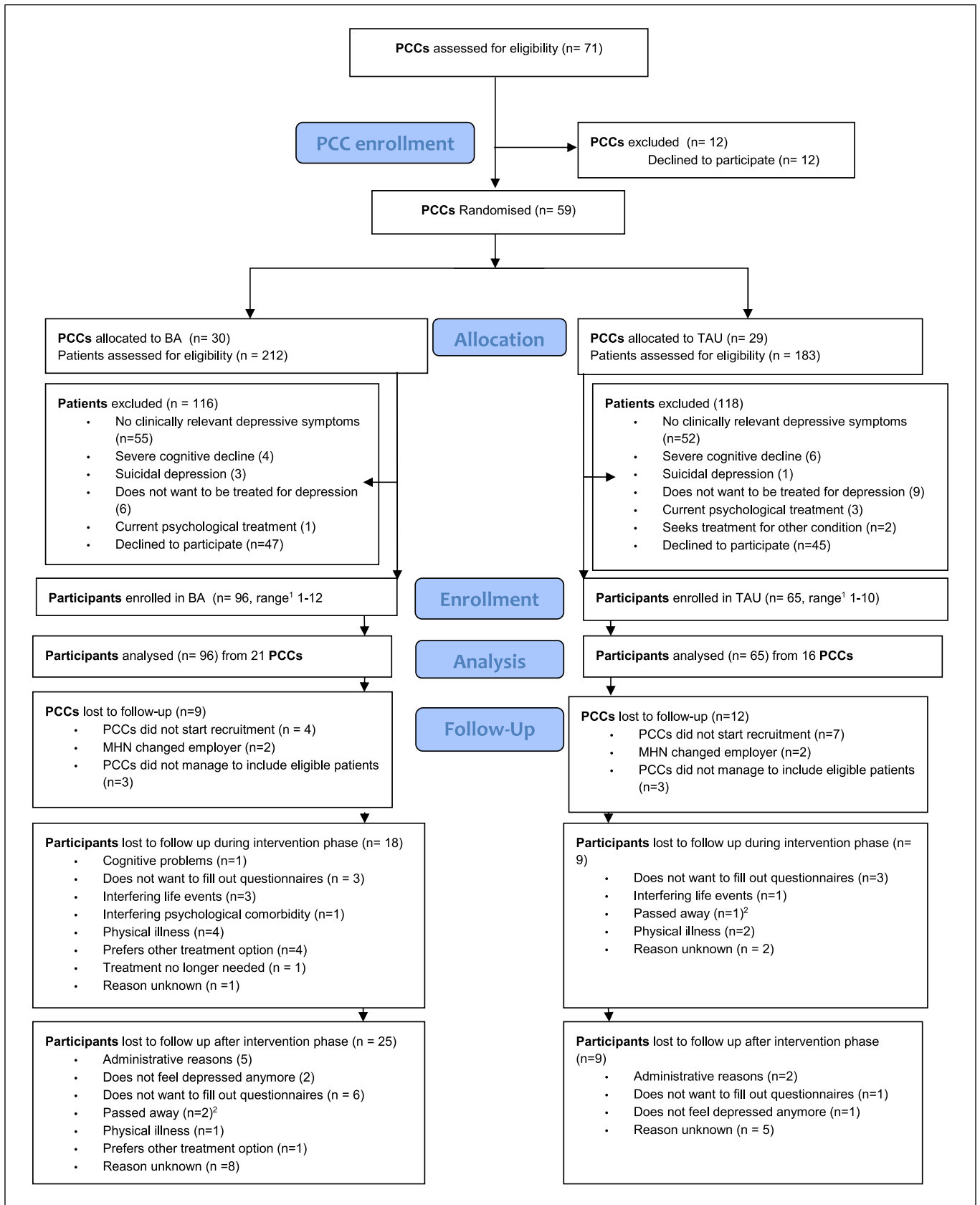
One participant dropped out before the first measurement. All available data from all other participants ( $n = 160$ ) were used in the intention-to-treat model. The best fitting model used quadratic splines without adding PCC as a level. The comparison between different models can be found in online supplementary Table S1.

The interaction between treatment arm and depression over time was significant [ $F(3, 127) = 4.60, p < 0.001$ ]. Figure 2 shows the predicted values of QIDS-SR16 based on our final model. At the end of treatment (9 weeks), EMM results show significantly fewer depressive symptoms in BA participants than in TAU participants (QIDS-SR16 difference =  $-2.77, p < 0.001$ ), and a significant difference persisted until the 3-month follow-up after post-treatment (QIDS-SR16 difference =  $-1.53, p = 0.02$ ) but not in the follow-ups at 6 to 12 months. The between-groups effect size at the end of treatment was  $d = 0.90$  (95% CI, 0.42–1.38), a large effect [38]. At the 12-month follow-up, the between-groups effect size was  $d = 0.29$  (95% CI,  $-0.81$  to 0.24), a small to medium effect [39]. All EMMs and effect sizes can be found in Table 2.

#### *Secondary Outcomes*

The PHQ-9 model showed a similar pattern as the QIDS-SR16 model. The PHQ-9 was administered before and during treatment and at follow-up, showing that BA participants reported significantly fewer depressive symptoms than TAU participants, starting 2 weeks after baseline (difference = 1.18,  $p < 0.001$ ) and reaching a difference of 3.21 ( $p < 0.001$ ) at the end of treatment, which was reduced to a difference of 1.09 ( $p = 0.25$ ) at the 12 month follow-up.

Based on the model data, using the standard deviation (SD) of baseline depression in our sample ( $SD = 3.96$ ) at post-treatment, 50% of the BA participants and 17% of the TAU participants showed reliable clinical improvement, and none of the BA participants and 3.1% of the TAU participants showed reliable deterioration. At the 12-month follow-up, the percentage of reliable improvement increased to 52% of the BA participants and 29% of TAU participants, and 1.1% of BA participants and 1.6% of TAU participants deteriorated. A Cox proportional hazards regression analysis, with treatment group, and help with questionnaires as covariates showed that the estimated risk of relapse (hazard ratio, 1.67 [95% CI, 0.63–4.38];  $p = 0.30$ ) and recurrence (hazard ratio, 0.88 [95% CI, 0.05–12.98];  $p = 0.88$ ) in the BA condition did not differ from the risk of relapse in the TAU condition.



1

(For legend see next page.)

### Per Protocol Analysis and Dropout

The per protocol analyses ( $n = 127$ ) yielded similar results as the intent-to-treat analyses. Participants who completed BA treatment per protocol ( $N = 62$ ) were more likely to live alone than participants who discontinued treatment preliminarily ( $N = 34$ ); ( $\chi^2 [1, N = 91] = 5.60, p = 0.018$ ).

Adding dropout to the final model did not significantly influence treatment results, and there was no significant interaction between dropout and treatment group. Participants who dropped out of the study ( $n = 54$ ) did not differ significantly at baseline from completers ( $n = 107$ ) on all baseline measures except on cognitive impairment (MoCA score ( $t(161) = 10.3, p = 0.002$ ) and age ( $t(161) = 4.42, p = 0.037$ ). Participants who dropped out of the study had more cognitive impairment (MoCA,  $M = 24.72$ ;  $SD = 2.81$ ) and higher age ( $M = 77$ ;  $SD = 7$ ) at baseline than participants who did not dropout (MoCA,  $M = 26.10$ ;  $SD = 2.36$ ; age,  $M = 74$ ;  $SD = 6$ ). Details of these analyses can be found in online supplementary Table S2.

### Ancillary Analyses

Adding a binary variable stating whether a measurement point was during a COVID-19 lockdown as a factor did not alter the results of the main analysis. Adding baseline variables age, education, cognitive impairment, and sex as a covariate did not alter the results of the main analysis. There was a significant interaction effect of time with received help during the study [ $F(3, 127) = 3.09, p = 0.03$ ]. Participants who received help during the study performed significantly worse on the 3-, 6-, and 9-month follow-up than participants filling in the questionnaires alone, with differences on QIDS-SR16 ranging from 1.66 at 3 months to 2.59 at 9 months. There was no interaction of received help with treatment arm [ $F(3, 127) = 0.006, p = 0.99$ ]. Detailed results of these ancillary analyses can be found in the online supplementary material.

## Discussion

This cluster-RCT showed that symptom severity was significantly lower in BA participants compared to TAU participants at post-treatment, starting in the second week of the study and up to 3 months after the end of treatment but not at 6–12 month follow-up. We observed

between-group effect sizes favouring BA of  $d = 0.90$  directly after treatment which were reduced to  $d = 0.29$  at the 12-month follow-up. Some bias might have occurred as a result of the cluster-randomised design. Consent and dropout rates were in line with expectations during the treatment phase, but BA (45%) had more dropouts than TAU (27%) after treatment. The BA arm had a higher consent rate (45%) than TAU (35%), and in TAU (13 PCCs), more PCCs were lost to follow-up than in BA (9 PCCs). The majority of PCCs (17 PCCs) that were lost to follow-up did not manage to or did not intent to recruit participants but did not formally withdraw. Our ancillary analyses showed that adding dropout as a factor in the main analyses did not alter the results and that age and cognitive impairment influenced dropout of the study but not treatment completion or treatment results. The only factor explaining preliminary discontinuation of treatment was “living alone.” Participants living with a partner were more likely to discontinue treatment than participants living alone. This factor has not been found in earlier studies.

Orgeta et al. [15] showed that BA was a promising treatment for depressed older adults with moderate to large effect sizes (SMD =  $-0.72$ , 95% CI,  $-1.04$  to  $-0.41$  at 4–12 weeks) compared to TAU but that higher quality research was needed to confirm this. To our knowledge, our study was the first adequately powered cluster-RCT assessing the effectiveness of a standalone BA treatment for depression in older adults, which shows similar results as the before-mentioned meta-analysis. Our results are comparable to those reported in a recent meta-analysis on psychotherapy for depression [2] that found large effect sizes for older adults ( $\geq 55$  to 75 years;  $g = 0.66$ ) as well as for the oldest old ( $\geq 75$  years;  $g = 0.97$ ) and to the results of the CASPER-plus trial that studied the effectiveness of collaborative care, with a focus on a brief BA intervention by telephone, as compared to TAU for older adults with depression [16]. This study found that collaborative care was more effective in the short term than TAU, with an effect size of 0.32, but the effects did not remain significant at the 12-month follow-up and beyond [16]. Collaborative care, however, includes appointing a case manager who discusses the patients’ health with other health professionals, keeps track of and adds medication if needed, and calls patients to provide information about other possibly helpful services. Implementing a collaborative care system may therefore be more complex and

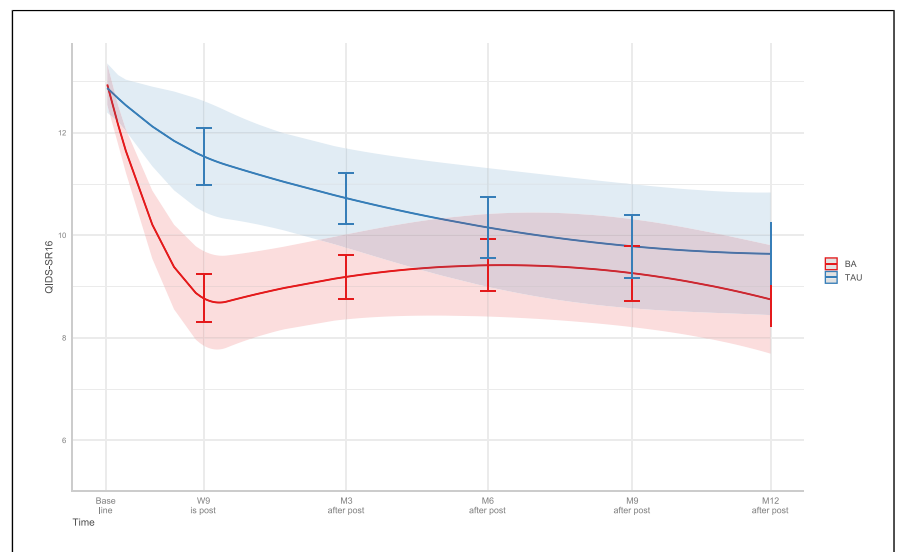
**Fig. 1.** Consort flow diagram. <sup>1</sup>“Range” refers to the number of enrolled participants per PCC. <sup>2</sup>All serious adverse events during treatment were assessed and presented to the Dutch Central Committee on research involving human subjects (CCMO), and all were categorised as unrelated to treatment.



**Table 1.** Baseline information

| Patient characteristics                                 | BA (n = 96) | TAU (n = 65) |
|---------------------------------------------------------|-------------|--------------|
| Age, years (SD)                                         | 75.8 (6.8)  | 74.3 (6.0)   |
| Sex, n (%) <sup>1</sup>                                 |             |              |
| Female                                                  | 61 (63.5)   | 36 (55.4)    |
| Marital status, n (%)                                   |             |              |
| Living alone                                            | 48 (53)     | 37 (59)      |
| Level of education, n (%)                               |             |              |
| Low                                                     | 39 (42.4)   | 27 (42.2)    |
| Intermediate                                            | 28(30.4)    | 28 (43.8)    |
| High                                                    | 25 (27.2)   | 9 (14.1)     |
| Duration of current depression, years <sup>2</sup> (SD) | 8.6 (15.8)  | 11.4 (17.7)  |
| Baseline psychotropic medication use, n (%)             |             |              |
| Antidepressants                                         | 32 (33)     | 19 (28)      |
| Tranquilizers                                           | 22 (23)     | 20 (31)      |
| Antipsychotics                                          | 4 (4)       | 0 (0)        |
| Baseline physical medications <sup>3</sup>              | 4.6 (3.2)   | 5.0 (3.1)    |
| Current depressive disorder, n (%)                      | 50 (52.1)   | 36 (55.4)    |
| Baseline QIDS-SR16 score (SD)                           | 13.4 (4.1)  | 12.7 (3.7)   |
| MoCA <sup>4, 5</sup> score, n (%)                       |             |              |
| Mild cognitive impairment (18–23)                       | 21 (22.1)   | 8 (12.7)     |
| No cognitive impairment (24–30)                         | 74 (77.9)   | 39 (87.3)    |
| At least one comorbid anxiety disorder, n (%)           | 33 (34.4)   | 19 (29.2)    |
| Loneliness SES <sup>6</sup> (SD)                        | 4.34 (1.29) | 4.39 (1.33)  |
| Received help with questionnaires, n (%)                | 18 (18.8)   | 17 (26.2)    |

<sup>1</sup>All percentages are based on available data. <sup>2</sup>Depression is 'clinically relevant depressive symptoms (PHQ-9<sub>≥</sub> 10)'. <sup>3</sup>All medications for physical diseases prescribed by a medical professional. <sup>4</sup>Montreal Cognitive Assessment. <sup>5</sup>Missing for three participants recruited during the COVID-lockdown, when home visits for testing were not allowed. During this time, the GP only referrec patients without cognitive decline. <sup>6</sup>SES = the 6-item De Jong Gierveld Social-Emotional Loneliness Scale ranging from 0 (not lonely) to 6 (very lonely).



**Fig. 2.** The predicted values of QIDS-SR16 with 95% confidence interval and error bars for the effect of treatment on depression over time based on a quadratic spline model with baseline depression as a covariate and a knot placed on day 78 (median day of end of treatment measurement).

**Table 2.** Tests of QIDS-SR EMMs

| Assessment                  | Treatment group (estimated marginal mean [SE]) |                 |            |                 | Estimated treatment difference |                          |                |
|-----------------------------|------------------------------------------------|-----------------|------------|-----------------|--------------------------------|--------------------------|----------------|
|                             | BA                                             | ES <sup>1</sup> | TAU        | ES <sup>1</sup> | difference (95% CI)            | ES <sup>2</sup> (95% CI) | <i>p</i> value |
| Baseline                    | 13.1 (0.2)                                     | –               | 12.9 (0.3) | –               | 0.16 (–0.48 to 0.80)           | –                        | 0.63           |
| Post-treatment <sup>3</sup> | 8.8 (0.5)                                      | 1.39            | 11.5 (0.6) | 0.44            | –2.77 (–4.19 to –1.35)         | 0.90 (0.42–1.38)         | <0.001         |
| Follow-up 3M <sup>4</sup>   | 9.2 (0.4)                                      | 1.26            | 10.7 (0.5) | 0.71            | –1.53 (–2.81 to –0.26)         | 0.50 (0.07–0.92)         | 0.02           |
| Follow-up 6M                | 9.4 (0.5)                                      | 1.18            | 10.2 (0.6) | 0.89            | –0.74 (–2.27 to 0.80)          | 0.24 (–0.26–0.74)        | 0.35           |
| Follow-up 9M                | 9.3 (0.5)                                      | 1.23            | 9.8 (0.6)  | 1.01            | –0.52 (–2.13 to 1.08)          | 0.17 (–0.36–0.70)        | 0.52           |
| Follow-up 12M               | 8.7 (0.5)                                      | 1.40            | 9.6 (0.6)  | 1.06            | –0.89 (–2.49 to 0.71)          | 0.29 (–0.24–0.82)        | 0.28           |

$y \sim bs$  (dy, knots = 78, degree = 2) \* Treatment arm +  $y_0$  + [bs (dy, knots = 78, degree = 2) + 0 |sid]. <sup>1</sup>ES = within group effect size per measurement point. <sup>2</sup>ES = between-group effect size. <sup>3</sup>Post-treatment is at 9 weeks after baseline. <sup>4</sup>Follow-up months are the months after post-treatment.

possibly costlier than implementing BA. Moreover, it should be noted that at 12 months after treatment ended, the effects ( $d = 0.29$ ) were similar to the effects of antidepressants as compared to placebo in other studies [40].

A main strength of this study was its ecological validity. Procedures closely matched the daily practice of PCCs, and our BA protocol was adapted to a primary care setting. Furthermore, compared to other studies, the participants in our sample were diverse in terms of age, symptom duration, and level of education. Whereas studies tend to be criticised for their stringent inclusion criteria, our study included older adults with physical and psychological comorbidity and mild cognitive impairment. The home visits by research assistants kept the one in five participants in the study who would have otherwise declined to participate, either because of illiteracy or difficulties with questionnaires. These participants had a higher severity of depressive symptoms on follow-up measures than participants who filled in questionnaires independently, regardless of treatment arm.

The study has some limitations that require careful consideration in light of the methodological recommendations for psychology trials by Guidi et al. [41]. First, while the pragmatic design of the study improved its ecological validity by closely resembling daily practice, the resulting heterogeneity may have limited the ability to draw definitive conclusions about the efficacy of BA, as opposed to its real-world effectiveness. For example, by using a PHQ-9 cut-off as an inclusion criterion, we missed an opportunity to differentiate between patients in different phases of depressive disorder [41]. It remains

unclear whether participants who experienced first incident depression benefit similarly from a BA treatment to those with persistent or recurrent treatment-resistant depressive symptoms. Future studies could adopt a staging approach that incorporates a patient's history into the randomisation, gaining insights into BA's effectiveness in differing stages of depressive disorder. Furthermore, the varying frequency of attended sessions and the heterogeneous nature of the treatment options in TAU may have impacted outcome differences post-treatment. Possibly, BA participants received more attention of a health care provider, and potential treatment ingredients of TAU remained unclear [42]. One way to eliminate such a "frequency effect" is to standardise the number of sessions in both treatment arms, for example, with an "attention placebo group" [41]. This may result in a better understanding of BA's efficacy in this population but would be less pragmatic and limit understanding of the effectiveness in usual health care settings. Future pragmatic studies could provide a more detailed description of the frequency of depression-related care in the TAU group, allowing the frequency effect to be explored in analyses.

Second, potential bias may have occurred in this study. Outcome measures were self-report questionnaires filled out by participants who were aware of the treatment group they were assigned to. The use of a single self-rating scale fails to consider the wide range of variables that can impact the clinical representation of a disorder, while adverse events are usually ignored [41]. Future studies should consider combination of observer- and self-rated tools, considering a wider range of factors beyond clinical symptoms. However, the QIDS-SR is valid when

compared to clinician-rated instruments such as the QIDS-C and MADRS but at a lower cost and burden on the patient [43, 44] which is also an important consideration in study design. Furthermore, assessors assisting with questionnaires were independent and blinded to treatment condition, and no significant baseline differences between groups were detected. Both treatment groups in this study were presented as active and likely to be effective conditions, limiting impact of an expectancy effect on results.

Third, dropout rates were high at both the patient as the PCC level. The enthusiasm expressed by participating umbrella organisations did not translate to active recruitment by all individual PCCs. Furthermore, MHNs in BA practices were offered additional training for the intervention, while MHNs in the TAU arm were not. Therefore, health care providers in the BA arm might have been more engaged in the recruitment process. Possibly, this has led to the inclusion of less motivated patients in the BA arm, which would explain the higher dropout rate in that group, though these non-significant rates may also be due to chance. Our ancillary analyses showed that adding dropout as a factor in the main analyses did not alter the results.

Fourth, recruitment ended prematurely due to COVID-19 lockdowns, but an interim sample size analysis conducted by an independent statistician confirmed that the study had sufficient power to answer research questions even with a slightly smaller dataset, as approved by the medical ethical committee. Ancillary analyses showed that the timing of the intervention either before or after COVID did not significantly alter treatment results. It remains unclear whether COVID-19 has affected participants in areas that questionnaires did not cover. A final limitation is the lack of formal registration of ethnicity as >95% was expected to be Caucasian. Therefore, results cannot be generalised to other ethnic groups.

These limitations notwithstanding, this study shows that MHNs can deliver BA, a low-intensity treatment for depressed older adults, and the oldest old in primary care that leads to a faster symptom reduction than TAU, resulting in significant differences between treatment groups at post-treatment but not in the long run (6–12 months after treatment ended). Two weeks into treatment, participants in the BA arm showed a significantly lower severity of symptoms than in TAU, lasting until 3 months post-treatment. A fast symptom reduction has individual and societal benefits since depressive symptoms are associated with a decreased quality of life, disability, greater healthcare utilisation, and increased mortality [45, 46]. BA is easy to train and easy to implement, which

provides an opportunity to reduce the use of psychopharmaceuticals that seem to lack effectiveness in older adults. However, in the long term, BA does not outperform TAU, and the treatment elements that make up the effectiveness of BA in older adults remain unknown. Future studies could use staging methods to determine whether the treatment is effective in different stages of depression and investigate whether booster sessions could sustain the symptom reduction longer.

## Acknowledgments

We would like to thank all participants and their spouses, all participating primary practices, including GPs and MHNs that referred their patients to us and assistants that informed their patients about the study, as well as the patients and GPs that have participated in writing the study protocol. Furthermore, we would like to thank the research assistants for all the kilometres they drove to help participants with their questionnaires and the hours they spent calling participants and PCCs and entering data.

## Statement of Ethics

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All procedures involving human subjects were approved by the Local Medical Ethical Committee, CMO Arnhem-Nijmegen, approval number 2016-2346, on May 18, 2016. All participants gave written informed consent before taking part.

## Conflict of Interest Statement

The study was funded by the Ministry of Health Funding-program for Health Care Efficiency Research ZonMw (843001606). All authors have completed the ICMJE uniform disclosure form at <http://www.icmje.org/disclosure-of-interest/>. GJH occasionally receives honoraria for lectures at symposia. On a yearly basis, it is never more than EUR 1,000. MJHH received a payment for training MHNs for the purpose of this study. DE is chief investigator for UK NIHR-funded multimorbidity in older adults' programme of research. A programme in a similar clinical area looking at BA in people 65 and over with 2 or more health problems and depression. HvM was supported by the National Institute of Health Research (NIHR) Applied Health Collaboration Kent, Surrey, Sussex. The views expressed are those of the author(s) and not necessarily the NHS, the NIHR, or the Department of Health and Social Care. All other authors declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years; and no other relationships or activities that could appear to have influenced the submitted work.

## Funding Sources

The Ministry of Health Funding-program for Health Care Efficiency Research ZonMw, the funding source of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

## Author Contributions

N.P.J., M.J.H.H., P.L., R.C.O.V., J.E.B., H.V.M., J.S., and G.-J.H. conceptualised the study. G.-J.H., M.J.H.H., and R.C.O.V. acquired funding. N.P.J. and G.-J.H. performed the investigation. M.J.H.H. and D.E. provided training for MHNs. N.P.J. and T.B. curated the data. N.P.J., G.-J.H., and T.B. acquired, analysed, and/or interpreted the data. G.-J.H., P.L., and M.J.H.H. supervised the study.

N.P.J. drafted the manuscript. P.L., R.C.O.V., D.E., T.B., J.E.B., H.V.M., J.S., and G.-J.H. critically revised the manuscript. All authors approved the final version of the manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

## Data Availability Statement

The data will be made publicly available at DOI <https://doi.org/10.34973/1dk9-tj89>, after completion of the project, but no later than December 2025, due to ethical considerations, which include ensuring that the release of the data does not interfere with the ongoing analyses of the research group. Until that time, researchers can obtain the data through consultation with the corresponding author.

## References

- 1 Chaplin R, Farquharson L, Clapp M, Crawford M. Comparison of access, outcomes and experiences of older adults and working age adults in psychological therapy. *Int J Geriatr Psychiatry*. 2015 Apr;30(2):178–84.
- 2 Cuijpers P, Karyotaki E, Eckshtain D, Ng MY, Corteselli KA, Noma H, et al. Psychotherapy for depression across different age groups: a systematic review and meta-analysis. *JAMA Psychiatry*. 2020 Mar;77(7):694–702.
- 3 Wuthrich VM, Frei J. Barriers to treatment for older adults seeking psychological therapy. *Int Psychogeriatr*. 2015 Mar;27(7):1227–36.
- 4 Pettit S, Qureshi A, Lee W, Stirzaker A, Gibson A, Henley W, et al. Variation in referral and access to new psychological therapy services by age: an empirical quantitative study. *Br J Gen Pract*. 2017 Dec;67(660):453–9.
- 5 Verhaak PF, van Dijk CE, Nuijen J, Verheij RA, Schellevis FG. Mental health care as delivered by Dutch general practitioners between 2004 and 2008. *Scand J Prim Health Care*. 2012 Jul;30(3):156–62.
- 6 Hetlevik Ø, Garre-Fivelsdal G, Bjorvatn B, Hjørleifsson S, Ruths S. Patient-reported depression treatment and future treatment preferences: an observational study in general practice. *Fam Pract*. 2019 Dec;36(6):771–7.
- 7 Holvast F, Massoudi B, Oude Voshaar RC, Verhaak PFM. Non-pharmacological treatment for depressed older patients in primary care: a systematic review and meta-analysis. *PLoS One*. 2017;12(9):e0184666.
- 8 Hu T, Dattani ND, Cox KA, Au B, Xu L, Melady D, et al. Effect of comorbidities and medications on frequency of primary care visits among older patients. *Can Fam Physician*. 2017 Jan;63(1):45–50.
- 9 Lau R, Stevenson F, Ong BN, Dziedzic K, Treweek S, Eldridge S, et al. Achieving change in primary care—causes of the evidence to practice gap: systematic reviews of reviews. *Implement Sci*. 2016;11(1):40–39.
- 10 Ekers D, Richards D, McMillan D, Bland JM, Gilbody S. Behavioural activation delivered by the non-specialist: phase II randomised controlled trial. *Br J Psychiatry*. 2011 Jan; 198(1):66–72.
- 11 Martell CR, Dimidjian S, Herman-Dunn R. *Behavioral activation for depression: a clinician's guide*. Guilford Press; 2013.
- 12 Stein AT, Carl E, Cuijpers P, Karyotaki E, Smits JA. Looking beyond depression: a meta-analysis of the effect of behavioral activation on depression, anxiety, and activation. *Psychol Med*. 2020 Jul;51(9):1491–504.
- 13 Cuijpers P, Karyotaki E, Harrer M, Stikkelbroek Y. Individual behavioral activation in the treatment of depression: a meta analysis. *Psychother Res*. 2023 Apr:1–12.
- 14 Ekers D, Godfrey C, Gilbody S, Parrott S, Richards DA, Hammond D, et al. Cost utility of behavioural activation delivered by the non-specialist. *Br J Psychiatry*. 2011 Dec;199(6):510–1.
- 15 Orgeta V, Brede J, Livingston G. Behavioral activation for depression for older people with or without cognitive impairment: a systematic review and metaanalysis of randomised controlled trials. *BJPsych*. 2017 Oct;13(7):P499–500.
- 16 Gilbody S, Lewis H, Adamson J, Atherton K, Bailey D, Birtwistle J, et al. Effect of collaborative care vs. usual care on depressive symptoms in older adults with subthreshold depression: the CASPER randomized clinical trial. *JAMA*. 2017 Feb;317(7):728–37.
- 17 Unützer J, Katon W, Callahan CM, Williams JW Jr, Hunkeler E, Harpole L, et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA*. 2002 Dec;288(22):2836–45.
- 18 Almeida OP, Patel H, Kelly R, Ford A, Flicker L, Robinson S, et al. Preventing depression among older people living in rural areas: a randomised controlled trial of behavioural activation in collaborative care. *Int J Geriatr Psychiatry*. 2021 Oct;36(4):530–9.
- 19 Lewis H, Adamson J, Atherton K, Bailey D, Birtwistle J, Bosanquet K, et al. Collaborative care and active surveillance for Screen-Positive ElDeRs with subthreshold depression (CASPER): a multicentred randomised controlled trial of clinical effectiveness and cost-effectiveness. *Health Technol Assess*. 2017; 21(8):1–196.
- 20 Jacobsen P, Wood L. Risk of contamination when planning psychological therapy trials can be assessed using a simple framework. *J Clin Epidemiol*. 2020 Aug;124:8–15.
- 21 Janssen N, Huibers MJH, Lucassen P, Voshaar RO, van Marwijk H, Bosmans J, et al. Behavioural activation by mental health nurses for late-life depression in primary care: a randomized controlled trial. *BMC Psychiatry*. 2017;17(1):230.
- 22 Cameron IM, Cardy A, Crawford JR, du Toit SW, Hay S, Lawton K, et al. Measuring depression severity in general practice: discriminatory performance of the PHQ-9, HADS-D, and BDI-II. *Br J Gen Pract*. 2011 May;61(588):e419–26.
- 23 Costa MV, Diniz MF, Nascimento KK, Pereira KS, Dias NS, Malloy-Diniz LF, et al. Accuracy of three depression screening scales to diagnose major depressive episodes in older adults without neurocognitive disorders. *Braz J Psychiatry*. 2016 Apr;38(2):154–6.
- 24 Van Vliet I, De Beurs E. *Het Mini Internationaal Neuropsychiatrisch Interview (MINI): een kort gestructureerd diagnostisch psychiatrisch interview voor DSM-IV-en ICD-10-stoornissen*. Tijdschr Psychiatr; 2007.
- 25 Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005 Apr;53(4):695–9.
- 26 Vereniging P. *Functie-en competentieprofiel POH-GGZ*. Utrecht: LHV; 2014.

- 27 Van Gelderen MG, Hassink-Franke LJA, Van Heest FB, Seeleman J, Terluin B, Woutersen-Koch H, et al. [NHG-standaard depressie](#). 2019.
- 28 Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry*. 2003 Sep; 54(5):573–83.
- 29 Rodda J, Walker Z, Carter J. Depression in older adults. *BMJ*. 2011 Sep;343:d5219.
- 30 R Core Team R. [R language definition](#). R foundation for statistical computing; 2020.
- 31 Kuznetsova A, Brockhoff PB, Christensen RHB. [Package “lmerTest”](#). R package version. 2015;2:734.
- 32 Tamura Y, Sato T, Ooe M, Ishiguro M. A procedure for tidal analysis with a Bayesian information criterion. *Geophys J Int*. 2007; 104(3):507–16.
- 33 Russell L. Emmeans: estimated marginal means, aka least-squares means. [R package version](#). 2022;1(7.2).
- 34 Twisk JW, Rijnhart JJ, Hoekstra T, Schuster NA, Ter Wee MM, Heymans MW. Intention-to-treat analysis when only a baseline value is available. *Contemp Clin Trials Commun*. 2020 Nov;20:100684.
- 35 Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol*. 1991;59(1):12–9.
- 36 Cuijpers P, Karyotaki E, Ciharova M, Miguel C, Noma H, Furukawa TA. The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: a meta-analysis. *Acta Psychiatr Scand*. 2021 Jun;144(3):288–99.
- 37 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 Sep;31(9):1841–53.
- 38 Cohen J. [The effect size](#). *Statistical power analysis for the behavioral sciences*. 1988.
- 39 Kim H-Y. Statistical notes for clinical researchers: effect size. *Restor Dent Endod*. 2015 Oct;40(4):328–31.
- 40 Khan A, Brown WA. Antidepressants versus placebo in major depression: an overview. [World Psychiatry](#). 2015 Oct;14(3):294–300.
- 41 Guidi J, Brakemeier E-L, Bockting CL, Cosci F, Cuijpers P, Jarrett RB, et al. Methodological recommendations for trials of psychological interventions. [Psychother Psychosom](#). 2018 Jul;87(5):276–84.
- 42 Bruijniks SJ, Lemmens LH, Hollon SD, Peeters FP, Cuijpers P, Arntz A, et al. The effects of once-versus twice-weekly sessions on psychotherapy outcomes in depressed patients. *Br J Psychiatry*. 2020 Apr; 216(4):222–30.
- 43 Doraiswamy PM, Bernstein IH, Rush AJ, Kyutoku Y, Carmody TJ, MacLeod L, et al. Diagnostic utility of the quick inventory of depressive symptomatology (QIDS-C16 and QIDS-SR16) in the elderly. [Acta Psychiatr Scand](#). 2010 Sep;122(3):226–34.
- 44 Bernstein IH, Rush AJ, Stegman D, Macleod L, Witte B, Trivedi MH. A comparison of the QIDS-C16, QIDS-SR16, and the MADRS in an adult outpatient clinical sample. *CNS Spectr*. 2010 Jul;15(7):458–68.
- 45 Meeks TW, Vahia IV, Lavretsky H, Kulkarni G, Jeste DV. A tune in “a minor” can “b major”: a review of epidemiology, illness course, and public health implications of subthreshold depression in older adults. *J Affect Disord*. 2011 Mar;129(1–3):126–42.
- 46 van den Berg KS, Wiersema C, Hegeman JM, van den Brink RH, Rhebergen D, Marijnissen RM, et al. Clinical characteristics of late-life depression predicting mortality. [Aging Ment Health](#). 2021 Mar;25(3):476–83.