

Self-help mobile messaging intervention for depression among older adults in resource-limited settings: a randomized controlled trial

Received: 10 October 2023

Accepted: 12 February 2024

Published online: 14 March 2024

 Check for updates

A list of authors and their affiliations appears at the end of the paper

Scalable solutions to treat depression in older adults in low-resourced settings are urgently needed. The PRODIGITAL-D pragmatic, single-blind, two-arm, individually randomized controlled trial assessed the effectiveness of a mobile messaging psychosocial intervention in improving depressive symptomatology among older adults in socioeconomically deprived areas of Guarulhos, Brazil. Older adults (aged 60+ years) registered with 24 primary care clinics and identified with depressive symptomatology (9-item Patient Health Questionnaire (PHQ-9) scores ≥ 10) received the 6-week Viva Vida intervention based on psychoeducation and behavioral activation ($n = 298$) or a single message ($n = 305$). No health professional support was offered. The primary outcome was improvement from depressive symptomatology (PHQ-9 < 10) at 3 months. Of the 603 participants enrolled (mean age = 65.1 years; 451 (74.8%) women), 527 (87.4%) completed the follow-up assessment. In the intervention arm, 109 of 257 (42.4%) participants had an improved depressive symptomatology, compared with 87 of 270 (32.2%) participants in the control arm (adjusted odds ratio = 1.57; 95% confidence interval = 1.07–2.29; $P = 0.019$). No severe adverse events related to trial participation were observed. These results demonstrate the usefulness of a digital messaging psychosocial intervention in the short-term improvement from depressive symptomatology that can potentially be integrated into primary care programs for treating older adults with depression. Brazilian Registry of Clinical Trials registration: ReBEC ([RBR-4c94dtn](https://www.clinicaltrials.gov/ct2/show/study?term=RBR-4c94dtn)).

Reducing the burden of depression in older adults is a global health priority, key to ensuring healthy aging and promoting well-being, especially in low- and middle-income countries (LMICs) where 69% of the world's older population live¹. In Brazil, the number of adults aged 65 years or older has increased by 57% since 2010, now representing 11% of the total population². Of the older adults interviewed in the Brazilian National Health Survey of 2019, 12% reported that they had experienced

depression in the past³. Estimates from low-resource settings in Brazil found that 30% of the older population had experienced depressive symptoms suggestive of clinical depression; only one-third of them had received a previous diagnosis⁴.

Programs led by trained nonmental health specialists (task-sharing) and with a team-based approach (collaborative care) have been effective in treating depression in all ages^{5,6}, including

✉ e-mail: scazufca@gmail.com; ricardo.araya@kcl.ac.uk

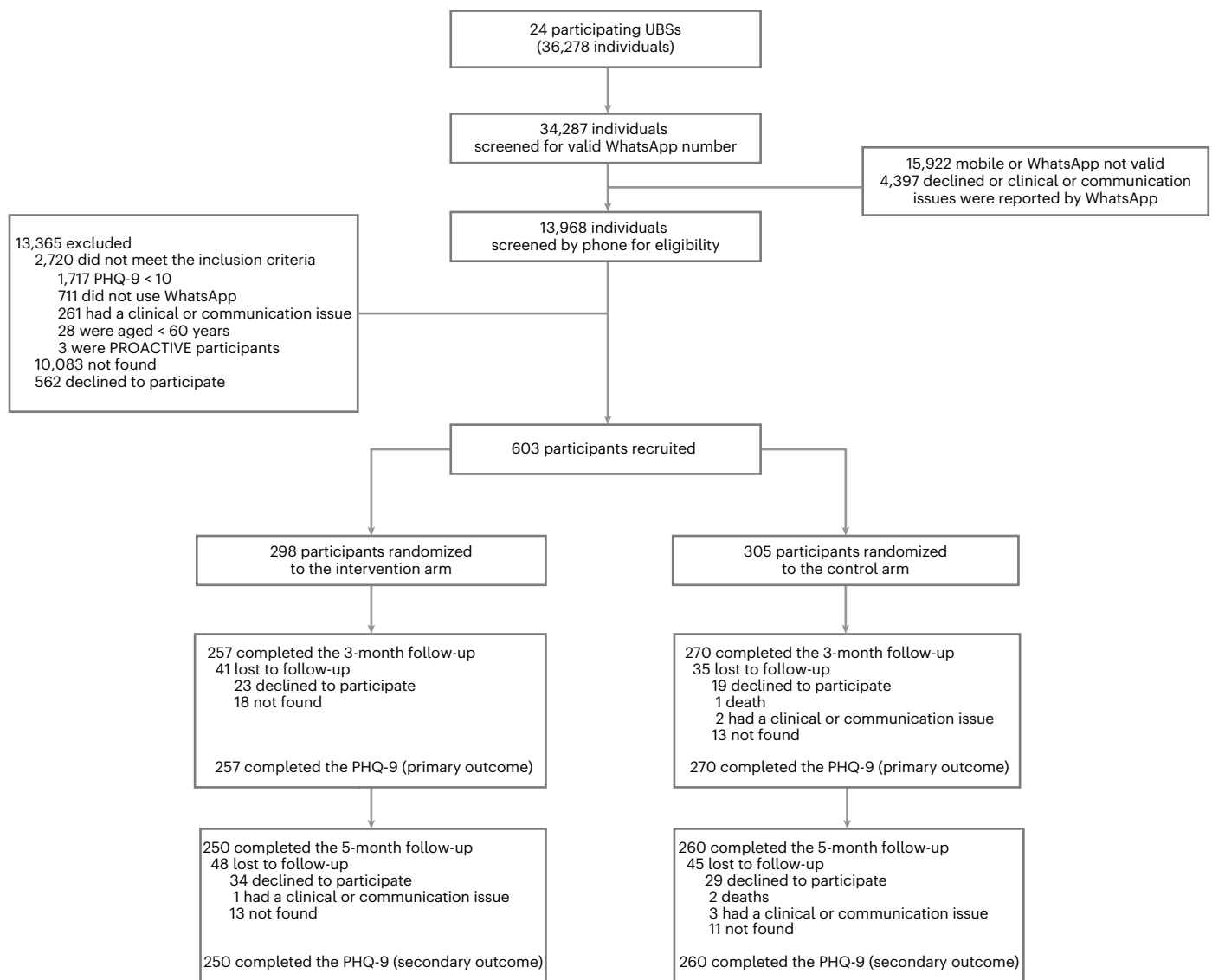


Fig. 1 | Consolidated Standards of Reporting Trials (CONSORT) diagram.

old age⁷⁻⁹. However, these programs often require a prominent role for health professionals, which is not feasible in health systems with few professionals who are already overstretched with other priorities^{10,11}. For example, during the coronavirus disease 2019 pandemic, non-coronavirus disease health services were reduced to a minimum in Brazil¹² because of restrictions on in-person consultations and the health system's unpreparedness to provide remote care.

A useful strategy to improve access to treatment for depression without increasing health professionals' workloads is the use of self-help digital mental health interventions¹³, which are low-cost alternatives requiring minimal or no health professional support^{14,15}. However, there is little evidence on the effectiveness of such interventions for older adults with depressive symptoms, especially in LMICs¹⁶.

Therefore, we developed a self-help mobile messaging intervention for older adults with substantial depressive symptoms¹⁷. Viva Vida is a 6-week digital psychosocial intervention delivered by WhatsApp audio and visual messages, without health professional support. It was adapted from a task-shared, collaborative care psychosocial program led by community health workers in Brazil, PROACTIVE¹⁸, which improved recovery from depression at the 8-month and 12-month follow-ups⁷. In this study, we evaluated the effectiveness of Viva Vida in

a pragmatic, single-blind, two-arm, individually randomized controlled trial among older adults registered in primary care in socioeconomically deprived areas of Guarulhos, Brazil (PRODIGITAL-D)¹⁷.

Results

Participant disposition

Participants were enrolled between 8 September 2021 and 8 April 2022. A list with 36,278 eligible older adults registered with the 24 primary care clinics, also called Unidades Básicas de Saúde (UBSs), was provided by the Guarulhos Health Secretariat. Among the 13,968 (38.5%) individuals with a valid WhatsApp number and screened for eligibility by phone: 2,720 (19.5%) did not meet the eligibility criteria (1,717 did not have a depressive symptomatology); 562 (4.0%) declined participation; and 10,083 (72.2%) could not be contacted (they did not answer either of the two calls on different days made by the research assistants) (Fig. 1). A total of 603 participants were recruited, 298 (49.4%) randomized to the intervention arm and 305 (50.6%) to the control arm, with the small discrepancy in numbers attributable to the (blocked) stratified allocation. With variations driven largely by differences in catchment populations, between five and 45 participants were recruited in each UBS, although most of them recruited 20–30 participants.

Table 1 | Baseline characteristics in the intervention and control arms

Characteristics	n of total sample size (%)	
	Intervention (n=298)	Control (n=305)
Female	225/298 (75.5)	226/305 (74.1)
Age group		
60–69 years	251/298 (84.2)	251/305 (82.3)
70+ years	47/298 (15.8)	54/305 (17.7)
Education		
None	45/297 (15.2)	51/302 (16.9)
1–4 years	83/297 (27.9)	84/302 (27.8)
5–8 years	94/297 (31.6)	85/302 (28.1)
>8 years	75/297 (25.3)	82/302 (27.2)
Personal income ^a		
Up to 1 minimum wage	197/298 (66.1)	191/305 (62.6)
>1–2 minimum wage	60/298 (20.1)	56/305 (18.4)
>2 minimum wage	41/298 (13.8)	58/305 (19.0)
Hypertension (self-reported)	207/298 (69.5)	222/305 (72.8)
Diabetes (self-reported)	123/298 (41.3)	121/305 (39.7)
Receiving pharmacological treatment for depression (self-reported)	38/298 (12.8)	46/305 (15.1)
PHQ-9 severity		
10–14	120/298 (40.3)	119/305 (39.0)
15–19	99/298 (33.2)	100/305 (32.8)
20+	79/298 (26.5)	86/305 (28.2)

^aIn 2021, the minimum wage in Brazil was R\$1,100 (approximately US\$213).

The descriptive statistics of sociodemographic and other relevant characteristics demonstrated no major differences between trial arms at baseline (Table 1); Extended Data Tables 1 and 2 show that baseline characteristics remained balanced at both follow-up assessments. The 70–79 and 80+ age groups were merged because of the small number of participants aged 80+ years ($n = 4$, all in the control arm). Most participants were aged 60–69 years (83.3%) and women represented 74.8%, of whom most had fewer than 8 years of education (73.8%) and were earning up to a minimum wage (64.3%). The proportion of participants with self-reported hypertension and diabetes were 71.1% and 40.5%, respectively.

A total of 527 (87.4%) participants recruited were followed up at 3 months; 41 (13.8%) intervention participants were lost to follow-up, 35 (11.5%) of which were controls. At the 5-month follow-up, 510 (84.6%) of 603 participants were assessed; 48 (16.1%) participants in the intervention arm and 45 (14.8%) participants in the control arm were lost to follow-up. Because the extent of the missing data exceeded the threshold of 10% specified in the statistical analysis plan (SAP), the primary analyses were based on imputed data as presented below alongside (for the primary outcome) the parallel complete case analyses. Extended Data Tables 3 and 4 demonstrate the similarity of baseline demographics between participants with complete data and those with missing data, at the 3-month and 5-month follow-up assessments, respectively.

Primary outcome

Estimates suggest a beneficial effect of the intervention for the primary outcome of improvement of depressive symptomatology (a 9-item Patient Health Questionnaire (PHQ-9) score < 10) at 3 months. While 109 (42.4%) participants in the intervention arm reached our threshold of

improvement, only 87 (32.2%) did so among controls (Table 2); the crude absolute difference was 10.2 (95% confidence interval (CI) = 2.0–18.4) percentage points. The adjusted odds ratio (OR) for this improvement, after imputing missing values, was 1.57 (95% CI = 1.07–2.29; $P = 0.019$). Findings from the complete case analyses were virtually identical (adjusted OR = 1.58, 95% CI = 1.09–2.29; $P = 0.016$); adjusting for the time between randomization and the first follow-up using complete data only also had no appreciable effect on the results (adjusted OR = 1.54, 95% CI = 1.05–2.24; $P = 0.025$).

Secondary outcomes

At 5 months, virtually identical proportions had improved in the two groups (Table 2); 50.8% ($n = 127$) and 50.0% ($n = 130$) for the intervention and control arms, respectively (adjusted OR = 1.01; 95% CI = 0.70–1.46; $P = 0.938$). Table 2 shows that the secondary outcome of PHQ-9 scores reducing by at least 50% from baseline values also demonstrated a benefit of the intervention at 3 months (1.58; 95% CI = 1.08–2.29; $P = 0.016$), but not at 5 months (1.04; 95% CI = 0.73–1.48; $P = 0.841$). The same general conclusions can be drawn from the (post hoc) analyses of the effects presented in terms of (crude) absolute risk differences (Table 2). Complete case analyses again showed extremely similar results (Extended Data Table 5). There was no evidence to support an effect of the intervention for any other secondary outcome at the 3-month and 5-month follow-up assessments using the imputed data: PHQ-9 scores (–1.14, 95% CI = –2.33–0.05; $P = 0.060$ and –0.21, 95% CI = –1.36–0.94; $P = 0.716$), 7-item General Anxiety Disorder (GAD-7) questionnaire scores (–0.89, 95% CI = –1.88–0.11; $P = 0.082$ and –0.14, 95% CI = –1.20–0.92; $P = 0.793$), 5-level European Quality of Life five-dimensional questionnaire (EQ-5D-5L) scores (0.025, 95% CI = –0.015–0.064; $P = 0.218$ and 0.035, 95% CI = –0.001–0.071; $P = 0.059$), ICEpop CAPability measure for older people (ICECAP-O) scores (0.018, 95% CI = –0.009–0.045; $P = 0.185$ and 0.002, 95% CI = –0.024–0.029; $P = 0.856$), 3-item University of California, Los Angeles (UCLA) scores (0.19, 95% CI = –0.09–0.48; $P = 0.186$ and 0.22, 95% CI = –0.07–0.52; $P = 0.140$) (Table 3). Similar results were found using the complete cases (Extended Data Table 6).

Safety

Follow-up data showed that six participants (2.0%) in each arm reported suicidal ideation. The safety protocol described in Methods was applied. Eighteen intervention participants (6.0%) and 20 controls (6.6%) were hospitalized during the trial for reasons not related to mental health. Two deaths (0.7%) were reported in the control arm. All events were considered not related to either the intervention or study participation.

Exploratory analyses

Bearing in mind the number of such tests performed, there was no evidence of differential effects in the prespecified subgroup analyses, with P values ranging from 0.055 to 0.953 (Extended Data Table 7).

The electronic system indicated that 226 (75.8%) intervention participants ‘opened’ at least 36 of the 48 messages (our prespecified minimum therapeutic dose), 237 (79.5%) opened at least 24 messages and 173 (58.1%) opened all 48 messages.

The main complier average causal effect (CACE) analysis did not yield strong evidence of an improvement in mean PHQ-9 scores for participants who opened 36+ of the messages compared with others (Table 4) at either the 3-month or 5-month follow-ups. Nonetheless, the P values for the main and sensitivity CACE analyses at 3 months ($P = 0.054$ for the prespecified threshold) provide marginal evidence that the adjusted effect of the intervention increased as the proportion of opened messages increased (Table 4).

Sensitivity analyses

The sensitivity analyses relating to aspects, such as the timing of the follow-up and altering the thresholds for the CACE analyses, have been

Table 2 | Primary (PHQ-9 <10; improvement) and secondary (PHQ-9 scores reduced by at least 50% between baseline and the first and second follow-up assessments; reduction) outcomes at the 3-month and 5-month follow-ups

Improvement from and reduction in depressive symptomatology at the 3-month and 5-month follow-ups	Intervention	Control	Risk difference in percentage points (95% CI) ^a	OR (95% CI) ^{b,c}	P ^d
	n of total n (%)	n of total n (%)			
Primary outcome: improvement from depressive symptomatology at 3 months ^e	109/257 (42.4)	87/270 (32.2)	10.2 (2.0–18.4)	1.57 (1.07–2.29)	0.019
Secondary outcome: improvement from depressive symptomatology at 5 months ^e	127/250 (50.8)	130/260 (50.0)	0.8 (–7.9–9.5)	1.01 (0.70–1.46)	0.938
Secondary outcome: reduction in depressive symptomatology at 3 months ^e	95/257 (37.0)	73/270 (27.0)	9.9 (2.0–17.9)	1.58 (1.08–2.29)	0.016
Secondary outcome: reduction in depressive symptomatology at 5 months ^e	111/250 (44.4)	112/260 (43.1)	1.3 (–7.3–9.9)	1.04 (0.73–1.48)	0.841

^aThese (post hoc) analyses are based on crude (that is, unadjusted) data on a complete case basis. ^bORs and 95% CIs were calculated using fixed-effects logistic regression models, from which the P values were derived. ^cAll estimates had missing data imputed according to trial arm using the multiple imputation by chained equations (MICE) models, which included stratification (sex, age group and baseline severity of PHQ-9 score (categorical)) and predictors of missingness (extended data). ^dAll P values reported are based on two-sided Wald tests from the relevant logistic regression model, with no adjustments for multiple tests. ^eThe primary and secondary outcomes of improvement from depressive symptomatology were defined as PHQ-9 scores less than 10; the secondary outcome of reduction in depressive symptomatology was defined as PHQ-9 scores reducing by at least 50% between baseline and the first and second follow-up assessments (yes/no).

covered in the relevant sections above. Results from the sensitivity analysis testing the missing at random (MAR) assumption for the primary and secondary outcomes of improvement from depressive symptomatology at 3 and 5 months suggest that estimates moved toward the null hypothesis. In other words, participants who improved from depressive symptomatology were more likely to be missing at the follow-up, compared with participants who did not improve (Extended Data Table 8).

Discussion

This RCT assessed a self-help digital psychosocial intervention delivered by WhatsApp to older adults in socioeconomically deprived areas in Brazil. Compared with a single educational message, at the primary (3-month) follow-up, the 6-week Viva Vida program showed increased odds of improvement from depressive symptomatology and greater likelihood of a reduction in PHQ-9 of at least 50%. No such differences were observed at 5 months. Levels of anxiety symptomatology, loneliness, quality of life and capability well-being were similar in the two arms at both follow-ups. Equally important, this study demonstrated the feasibility and acceptability of a simple and affordable intervention that can provide help to a large proportion of older adults with depressive symptoms in Brazil. This would be a major achievement in a country with a large treatment gap and lack of specialists to cover such a need.

Digital strategies to deliver behavioral activation interventions have achieved small-to-medium effects on depression, anxiety and quality of life in adults¹⁹. Similar to our findings, two trials (CONEMO) conducted in adults with depression and comorbid hypertension or diabetes in Brazil and Peru showed that a 6-week digital intervention minimally supported by nurses or nurse assistants improved depressive symptomatology by at least 50% at 3, but not 6 months, compared with enhanced usual care²⁰. The studies suggest that both the CONEMO and Viva Vida digital interventions, based on behavioral activation, accelerate the improvement from depression in the short-term in middle-to-older adults but the clinical differences with control groups decrease mainly because these comparison groups show improved outcomes through time. Further studies could evaluate if booster messages in the active group (Viva Vida) would help maintain the speed of improvement and increase clinical benefits even further over time, so that the earlier differences with the control remain over a longer time.

Human support, such nurse assistants in CONEMO, has been reported as an important component of effective digital mental health interventions¹⁶. However, the Viva Vida program precluded such clinical support because a self-help mental health intervention could provide a more affordable first step for treating depression in primary care.

According to the updated taxonomy of digital mental health interventions by Pineda et al.¹³, the Viva Vida program is classified as a type 3 intervention because it offered technical support during the trial along with other trial procedures to guarantee the delivery and fidelity of the intervention, and the reliability of the data collected. Nevertheless, it can be turned into a type 4 fully automated intervention in the real world, where such controlled procedures are not necessary. The messages can be initially accessed from anywhere with an internet connection; once the content is downloaded to WhatsApp or to a device, the messages can be accessed at any time even without an internet connection. With such a fully automated system in place, a more affordable program can be delivered to a large number of people with a small incremental cost for each message sent.

The estimated intervention effect (10.2 percentage points in absolute terms) was lower than that specified as the original target difference (15 percentage points). As the target difference was based on collaborative care interventions with the active support of primary care professionals, we contend that our findings are important because a digital intervention would reach many of those who at the moment are receiving no help. In low-resource settings, where access to mental health care is very limited, fully automated programs might be used to cover treatment gaps when there are limited or no other alternatives. We do not expect it to be suitable for everyone and some individuals might need to be referred to a more intense program with professional support. Moreover, while from the various confidence intervals we cannot rule out smaller, arguably unimportant effects (relative or absolute), it is likewise the case that absolute effects at least as large as our target difference are contained in the relevant interval estimates at 3 months. The effect size of our secondary depression outcomes at 3 months are nonetheless similar to those for digital interventions supported by nurse assistants²⁰ or lay counselors²¹.

WhatsApp is widely used by the target population in Brazil. The use of a well-known and accepted messaging application enables the receipt of messages without requiring the download of a new application or the acquisition of additional digital skills, and is consistent with the findings of a systematic review of similar technologies¹⁶. Older adults with relatively low digital literacy level or without access to a high-speed internet connection were able to use this tool, which showed high adherence and acceptability. Women represented 75% of our sample as a reflection of the epidemiological profile of depressive symptoms among the older population^{4,22}. Future studies could explore the effects of personalized programs according to different participants' preferences or characteristics, such as sex and presence of other comorbidities.

Table 3 | Secondary outcomes, including depressive symptomatology (PHQ-9), anxiety symptomatology (GAD-7), health-related quality of life (EQ-5D-5L), capability well-being in older adults (ICECAP-O) and perceived loneliness (3-item UCLA) at the 3-month and 5-month follow-ups

Outcome	Baseline		3-month follow-up				5-month follow-up			
	Intervention	Control	Intervention	Control	Coefficient (95% CI) ^{a,b}	P ^c	Intervention	Control	Coefficient (95% CI) ^{a,b}	P ^c
PHQ-9, <i>n</i>	298	305	257	270			250	260		
Mean (s.d.)	16.39 (4.66)	16.68 (4.83)	11.72 (7.72)	12.92 (7.13)	-1.14 (-2.33–0.05)	0.060	10.01 (7.07)	10.41 (7.01)	-0.21 (-1.36–0.94)	0.716
GAD-7, <i>n</i>	285	283	256	268			250	260		
Mean (s.d.)	14.46 (4.62)	14.36 (4.75)	10.30 (6.62)	11.07 (6.43)	-0.89 (-1.88–0.11)	0.082	9.40 (6.79)	9.49 (6.66)	-0.14 (-1.20–0.92)	0.793
EQ-5D-5L, <i>n</i>	296	305	256	266			247	259		
Mean (s.d.)	0.789 (0.182)	0.794 (0.170)	0.812 (0.189)	0.801 (0.171)	0.025 (-0.015–0.064)	0.218	0.840 (0.164)	0.825 (0.171)	0.035 (-0.001–0.071)	0.059
ICECAP-O, <i>n</i>	295	300	254	266			246	259		
Mean (s.d.)	0.614 (0.172)	0.621 (0.190)	0.655 (0.174)	0.642 (0.194)	0.018 (-0.009–0.045)	0.185	0.665 (0.165)	0.667 (0.172)	0.002 (-0.024–0.029)	0.856
3-item UCLA, <i>n</i>	298	305	255	267			250	259		
Mean (s.d.)	6.04 (1.91)	6.16 (1.92)	5.68 (1.97)	5.54 (1.94)	0.19 (-0.09–0.48)	0.186	5.43 (1.97)	5.25 (1.97)	0.22 (-0.07–0.52)	0.140

PHQ-9 scores range from 0 to 27, with higher scores representing more severe depressive symptomatology. GAD-7 scores range from 0 to 21, with higher scores representing more severe anxiety symptomatology. EQ-5D-5L scores range from -0.264 to 1, with higher scores representing higher quality of life. ICECAP-O scores range from 0 to 1, with higher scores representing greater levels of capability well-being. The 3-item UCLA scores range from 3 to 9, with higher scores representing more severe levels of loneliness. ^aDifferences in means were estimated using linear regression models, adjusted for the baseline assessment of the corresponding outcome and the stratified variable (sex, age group and baseline severity of the PHQ-9 score (categorical)).

^bAll estimates had missing data imputed separately, according to trial arm, using MICE models that included predictors of missingness (extended data), stratification (sex, age group and baseline severity of PHQ-9 score (categorical)) and any imbalances in outcome measure at baseline. ^cAll *P* values reported are based on two-sided *t*-tests from the relevant linear regression model, with no adjustments for multiple tests.

Table 4 | CACE analyses on the mean PHQ-9 scores at the 3-month and 5-month follow-ups using electronic recordings

Threshold	3-month follow-up		5-month follow-up	
	Adjusted difference in mean PHQ-9 (95% CI) ^{a,b}	P ^c	Adjusted difference in mean PHQ-9 (95% CI) ^{a,b}	P ^c
Opened at least 24 messages (<i>n</i> =237, 79.5%)	-1.52 (-3.06–0.02)	0.053	-0.49 (-1.98–0.99)	0.513
Opened at least 36 messages, that is, the minimum therapeutic dose (<i>n</i> =226, 75.8%)	-1.59 (-3.21–0.02)	0.054	-0.52 (-2.08–1.04)	0.513
Opened all 48 messages (<i>n</i> =173, 58.1%)	-2.09 (-4.20–0.03)	0.053	-0.68 (-2.71–1.35)	0.513

PHQ-9 scores range from 0 to 27, with higher scores representing more severe depressive symptomatology. ^aAll models are adjusted for stratification (sex, age group and baseline severity of PHQ-9 score (categorical)). ^bAll estimates had missing data imputed separately, according to trial arm, using MICE models that included predictors of missingness (extended data), stratification (sex, age group and baseline severity of PHQ-9 score (categorical)).

^cAll *P* values reported are based on two-sided *t*-tests from the relevant CACE regression model, with no adjustments for multiple tests.

In addition to the issue of the lower magnitude of the observed difference in absolute terms (and some of the values in the relevant CI) compared with the target difference as discussed above, this study has some important limitations. As assessments were done by phone, simple questions and short interviews were the best option available. Thus, the use of the PHQ-9 for the assessment of depressive symptomatology was a pragmatic choice and the threshold of ten as indicative of clinical depression was based on the literature^{23,24}. Relevant conditions among older adults, such as diabetes and hypertension, were self-reported. We were unable to assess cognition with a standardized assessment over the phone. However, research assistants were trained to identify when the older adult had substantial difficulties understanding

and answering the questions. In these situations, the assessment was interrupted. When possible, the research assistant asked permission to speak to someone in the household about the difficulties presented by the person during the interview. In addition, the research assistant discussed these cases with the study coordinator, who decided whether the person would be included in the study. During the interview, we asked if the person had any vision or hearing problems, and if they had any difficulty in understanding audio or visual messages on a mobile phone. Individuals with a lower digital literacy or without access to WhatsApp were not included in this study because of the design of the digital intervention. Major adaptations are needed to deliver a more appropriate intervention to reach such individuals. Finally, cost-effectiveness and process evaluation remain to be assessed using the economic and qualitative data collected within PRODIGITAL-D. Indeed, the qualitative data may also shed further light on the issue of any effects of opening more rather than fewer of the messages as investigated within the above, essentially exploratory, CACE analyses.

Albeit at about half the level that we accounted for in our sample size calculation for the primary outcome at 3 months (about 12.5% overall compared with the 25% allowed for), a further limitation is that there is a risk of bias through attrition. As well as being low compared with expectations, some reassurance on this issue can be gleaned from the similar amounts of missing data across the two arms and the consistency of results between the analyses using imputed data (albeit under an MAR assumption) and the complete case analyses.

As this was a research project, we used standardized procedures for assessments, including quality control measures. Trained research assistants performed the recruitment and outcome assessments remotely on participants of both study arms. It is always possible that simply asking questions may have a therapeutic effect. In any case, the same procedures were used with participants in both arms. Another procedure common to participants of both arms was the contact from the technical support team if they experienced difficulties accessing messages. Such contacts were not frequent (20 in the intervention arm and 12 in the control arm) and were strictly for solving any technical problems with receiving messages. Participants in the intervention arm could also contact the technical support team during the 6-week intervention, but no contacts were made. In any case, it was not possible to disentangle the effect of the intervention from the possible

therapeutic effects that remote assessments or the technical support contact may have had on participants.

This RCT provides evidence of the feasibility and a small and short-term effectiveness of a simple, brief and well-accepted self-help digital intervention for depressive symptomatology in older adults that does not require the active support of health professionals. Such interventions, which can be accessed from anywhere with (and without) an internet connection, can contribute to mental health care in low-resourced settings where there is little access to treatment for depression.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-024-02864-4>.







References


1. *Population Ages 65 and Above, Total—Low & Middle Income, High Income* (World Bank, 2023); data.worldbank.org/indicator/SP.POP.65UPTO?end=2021&locations=XO-XD&start=1960&view=chart
2. *Censo 2022: Número de Pessoas com 65 Anos ou Mais de Idade Cresceu 57,4% em 12 Anos* (Agência de Notícias, Instituto Brasileiro de Geografia e Estatística, 2023); agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias-noticias/38186-censo-2022-numero-de-pessoas-com-65-anos-ou-mais-de-idade-cresceu-57-4-em-12-anos
3. de Albuquerque Brito, V. C. et al. Prevalence of self-reported depression in Brazil: National Health Survey 2019 and 2013. *Epidemiol. Serv. Saude* **31**, e2021384 (2022).
4. Nakamura, C. A. et al. Depressive and subthreshold depressive symptomatology among older adults in a socioeconomically deprived area in Brazil. *Int. J. Geriatr. Psychiatry* **37**, 34811807 (2021).
5. Karyotaki, E. et al. Association of task-shared psychological interventions with depression outcomes in low- and middle-income countries: a systematic review and individual patient data meta-analysis. *JAMA Psychiatry* **79**, 430–443 (2022).
6. Gilbody, S., Bower, P., Fletcher, J., Richards, D. & Sutton, A. J. Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. *Arch. Intern. Med.* **166**, 2314–2321 (2006).
7. Scazufca, M. et al. A task-shared, collaborative care psychosocial intervention for improving depressive symptomatology among older adults in a socioeconomically deprived area of Brazil (PROACTIVE): a pragmatic, two-arm, parallel-group, cluster-randomised controlled trial. *Lancet Healthy Longev.* **3**, e690–e702 (2022).
8. Unützer, J. et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA* **288**, 2836–2845 (2002).
9. Bosanquet, K. et al. Collaborative care for Screen-Positive ElDeRs with major depression (CASPER plus): a multicentred randomised controlled trial of clinical effectiveness and cost-effectiveness. *Health Technol. Assess.* **21**, 1–252 (2017).
10. Henrique, M. G. et al. Acceptability and fidelity of a psychosocial intervention (PROACTIVE) for older adults with depression in a basic health unit in São Paulo, Brazil: a qualitative study. *BMC Public Health* **21**, 2278 (2021).
11. Ferrinho, P., Sidat, M., Goma, F. & Dussault, G. Task-shifting: experiences and opinions of health workers in Mozambique and Zambia. *Hum. Resour. Health* **10**, 34 (2012).
12. Bigoni, A. et al. Brazil's health system functionality amidst of the COVID-19 pandemic: an analysis of resilience. *Lancet Reg. Health Am.* **10**, 100222 (2022).
13. Pineda, B. S. et al. Updated taxonomy of digital mental health interventions: a conceptual framework. *Mhealth* **9**, 28 (2023).
14. *Self-Care Interventions for Health* (World Health Organization, 2022); who.int/news-room/fact-sheets/detail/self-care-health-interventions
15. Philippe, T. J. et al. Digital health interventions for delivery of mental health care: systematic and comprehensive meta-review. *JMIR Ment. Health* **9**, e35159 (2022).
16. Riadi, I. et al. Digital interventions for depression and anxiety in older adults: a systematic review of randomised controlled trials. *Lancet Healthy Longev.* **3**, e558–e571 (2022).
17. Nakamura, C. A. et al. Digital psychosocial intervention for depression among older adults in socioeconomically deprived areas in Brazil (PRODIGITAL-D): protocol for an individually randomised controlled trial. *Trials* **23**, 761 (2022).
18. Scazufca, M. et al. A collaborative care psychosocial intervention to improve late life depression in socioeconomically deprived areas of Guarulhos, Brazil: the PROACTIVE cluster randomised controlled trial protocol. *Trials* **21**, 914 (2020).
19. Han, A. & Kim, T. H. Effects of internet-delivered behavioral activation on individuals with depressive symptoms: a systematic review and meta-analysis. *J. Psychiatr. Res.* **152**, 104–118 (2022).
20. Araya, R. et al. Effect of a digital intervention on depressive symptoms in patients with comorbid hypertension or diabetes in Brazil and Peru: two randomized clinical trials. *JAMA* **325**, 1852–1862 (2021).
21. Arjadi, R. et al. Internet-based behavioural activation with lay counsellor support versus online minimal psychoeducation without support for treatment of depression: a randomised controlled trial in Indonesia. *Lancet Psychiatry* **5**, 707–716 (2018).
22. *Pesquisa Nacional de Saúde* (Instituto Brasileiro de Geografia e Estatística, 2019); www.pns.icict.fiocruz.br/bases-de-dados/
23. Kroenke, K., Spitzer, R. L. & Williams, J. B. The PHQ-9: validity of a brief depression severity measure. *J. Gen. Intern. Med.* **16**, 606–613 (2001).
24. Levis, B. et al. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. *BMJ* **365**, l1476 (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

Marcia Scazufca ^{1,2,9} , **Carina Akemi Nakamura** ^{2,9}, **Nadine Seward**³, **Thiago Vinicius Nadaletto Didone**^{2,4}, **Felipe Azevedo Moretti**^{2,5}, **Marcelo Oliveira da Costa**², **Caio Hudson Queiroz de Souza** ², **Gabriel Macias de Oliveira**², **Monica Souza dos Santos**², **Luara Aragoni Pereira**¹, **Mariana Mendes de Sá Martins**², **Pepijn van de Ven**⁶, **William Hollingworth**⁷, **Tim J. Peters** ^{8,9} & **Ricardo Araya**^{3,9} 

¹Instituto de Psiquiatria, Hospital das Clínicas HCFMUSP, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil. ²Departamento de Psiquiatria, Faculdade de Medicina FMUSP, Universidade de São Paulo, São Paulo, Brazil. ³Health Service and Population Research, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. ⁴Departamento de Saúde Coletiva, Centro de Ciências da Saúde, Universidade Estadual de Londrina, Londrina, Brazil. ⁵Neuroscience Research Group, Institute D'Or for Research and Teaching, São Paulo, Brazil. ⁶Health Research Institute, University of Limerick, Limerick, Ireland. ⁷Health Economics Bristol, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. ⁸Population Health Sciences, Bristol Medical School, and Bristol Dental School, University of Bristol, Bristol, UK. ⁹These authors contributed equally: Marcia Scazufca, Carina Akemi Nakamura, Tim J. Peters, Ricardo Araya.  e-mail: sczufca@gmail.com; ricardo.araya@kcl.ac.uk

Methods

Study design and participants

PRODIGITAL-D was a pragmatic, single-blind, two-arm individually RCT with 1:1 allocation, conducted in 24 primary care clinics, also called UBSs, in underprivileged areas of the city of Guarulhos, Brazil. This city is part of the metropolitan region of São Paulo with a population of around 1.3 million.

We recruited individuals aged 60+ years, registered with any of the participating 24 UBSs, able to receive WhatsApp messages, and who screened positive for depressive symptomatology (PHQ-9 ≥ 10)²³ with at least one core symptom of depression (low mood or anhedonia, that is, PHQ-2 ≥ 1)²⁵. The PHQ-9 threshold had a sensitivity of 91% and specificity of 88% against diagnostic interviews for adults aged 60+ years²⁴. Exclusion criteria included substantial visual or hearing impairments that would hinder comprehension of messages on a mobile phone, acute suicidal risk identified during the screening assessment, another individual residing in the same household already participating in PRODIGITAL-D or previous participation in the PROACTIVE trial. Informed verbal consent was obtained before the screening assessment and when participants were invited to the trial, both conducted by phone. Consent was audio-recorded after authorization from the older adult. Details of the protocol are published elsewhere¹⁷.

This study was approved by the ethics committee of the Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (Comissão para Análise de Projetos de Pesquisa, ref: 4.097.596, first approved 10 March 2021) and authorized by the Guarulhos Health Secretary. It was registered with the Brazilian Registry of Clinical Trials (ReBEC), [RBR-4c94dtn](https://www.rebec.org.br/registro/4c94dtn), on 22 October 2021. Trial registration materials were submitted to ReBEC on 3 August 2021.

Procedures

A list with all the names and phone numbers of older adults registered with the 24 participating UBSs was provided by the Guarulhos Health Secretariat. After excluding duplicates and individuals without phone numbers, individuals were assigned a random ID for our study, as the original list was in alphabetical order. Research assistants then sent a WhatsApp message to older adults according to the new list sorted by ascending random ID numbers to search for individuals with a valid mobile number and who used WhatsApp. Individuals with a valid WhatsApp number (that is, who received the prescreening WhatsApp message) were approached by telephone to be screened further for depressive symptomatology (using the PHQ-2)²⁵. Up to two call attempts were made on different days. Individuals with a PHQ-2 score of one or greater were then asked the other seven items of the PHQ-9 questionnaire²³. Those with PHQ-9 scores of ten or greater completed a baseline assessment to ascertain the exclusion criteria; eligible individuals were invited to participate in the trial. The screening and baseline assessments together lasted approximately 35 min. The invitation for the trial, which includes reading the full participant information sheet, occurred in the same call, whenever the individual agreed to continue it for around 15 min longer. When this was not possible, an additional call no more than 28 days after the PHQ-9 screening was made to complete the recruitment. Follow-up data were collected by phone at 3 (weeks 12–16) and 5 months (weeks 20–24) after sending the first message (intervention) or the single message (control). A 4-week window was used to maximize response to follow-up considering the logistical realities of conducting phone interviews.

Quality control of a random sample of recorded assessments was conducted by an independent research assistant (not involved in any data collection) to ensure all procedures had been followed and the quality of data collected.

Randomization and masking

Participants were stratified according to age group (60–69 years, 70–79 years, 80+ years), sex (male, female) and severity of depressive

symptoms (baseline PHQ-9 categories: 10–14, 15–19, 20+). The allocation sequence was generated using randomly permuted blocks with random block sizes of six, eight or ten by research team members not involved in data collection (C.A.N. and T.J.P.). The ‘randomization module’ of the Research Electronic Data Capture (REDCap)^{26,27} was used to conceal the allocation sequence and randomize individuals.

Research assistants involved in recruitment and follow-up data collection were blinded to trial allocation. Two different groups of research assistants were responsible for either the first or the second follow-up assessment. Researchers responsible for scheduling the automated delivery of messages were not blinded to trial allocation but had no role in collecting the inclusion or outcome data. Due to the nature of the interventions, it was not possible to mask trial participants.

Interventions

Participants in both arms continued receiving their UBS’s usual primary care. The research team did not interfere with that care (consultation or treatments, including medication) and the UBS health professionals had no role in Viva Vida.

During the first 2 weeks (for the intervention arm) and after the single message (for the control arm), a technical support person identified and contacted individuals who were having issues with receiving or opening the message. This was to help participants with any potential technical issues that were preventing them from accessing the messages.

Viva Vida program

Participants allocated to the intervention received a psychosocial intervention delivered via WhatsApp. A total of 48 audio or visual messages were automatically sent 4 days a week for 6 weeks (one in the morning, one in the afternoon). The audio messages were based on psychoeducation²⁸ or behavioral activation²⁹ principles and lasted on average 3 min using storytelling techniques. In the audio messages, there were two characters (Ms Ana and Mr Leo) reading and commenting on letters sent by fictional older adults participating in the program. Through these letters they shared stories about their lives and symptoms of depression and how the Viva Vida intervention helped them to feel better. Eight visual messages including summary diagrams were used to reinforce the content of the audio messages. Additionally, at the end of each week, participants received an extra message with a question about their opinion of the program that could be responded to using the WhatsApp quick reply tool. Participants also received one message at the beginning of Viva Vida with a support contact number in case they experienced technical issues receiving the messages. As the messages were delivered automatically, this strategy allowed participants to contact us in case they were not receiving the intervention as intended. During the program, participants were also advised to visit the UBS if they needed professional support for any health issues. Extended Data Fig. 1 shows the structure of the Viva Vida program and the content of the audio and visual messages that participants received each week.

The Viva Vida messages were based on the PROACTIVE intervention delivered to older adults^{10,18,30}. The use of WhatsApp as a tool to deliver an intervention was initially discussed in a group with six older adults and one community health worker in one UBS. After developing the messages, selected older adults received a sample and were interviewed to evaluate qualitatively the acceptability of the messages. The automated system and the recruitment by phone were subsequently evaluated in a feasibility study with a random sample of older adults with depressive symptomatology recruited from two UBSs in Guarulhos that did not participate in the study.

Single message control arm

Control arm participants were sent a single and brief (6-min) audio message providing information about depression and simple ways to deal with depressive symptoms. It also offered additional tips encour-

aging participants to live a healthier life, such as improving sleep and maintaining a balanced diet, advising them to seek health care if they felt they needed additional support.

Outcomes

Improvement from depressive symptomatology at 3 months was the primary outcome. We considered improvement when the PHQ-9 scores were below 10 (ref. 23). Secondary outcomes included improvement from depressive symptomatology at 5 months and reduction in PHQ-9 scores by at least 50% between baseline and follow-up visits at 3 and 5 months. We also evaluated as exploratory outcomes the effects on the continuous scores of depressive symptomatology (PHQ-9), anxiety symptomatology with the GAD-7 (ref. 31), loneliness with the 3-item UCLA loneliness scale³², health-related quality of life EQ-5D-5L³³ and capability well-being with the ICECAP-O³⁴.

Safety

As mentioned before, participants continued receiving usual primary care and were advised to contact the primary care team whenever they felt they needed additional support from health professionals. The risk associated with receiving the interventions was considered minimal. Severe adverse events, including acute suicidal risk, hospitalizations and death were collected during the follow-up assessments using a standardized form. Participants who reported suicidal ideation were assessed for suicidal risk with the standardized protocol used in a previous study⁷. UBS managers were contacted to inform about participants with suicidal ideation. We also investigated any hospitalization or death with participants or family members to understand if it was related to study participation.

Data collection and management

All the measures mentioned above were applied at baseline, and at the 3-month and 5-month follow-up assessments. Sociodemographic and clinical characteristics were also collected at baseline. Economic data were collected at each follow-up assessment. Severe adverse events were assessed for all participants at 3 and 5 months.

Data were collected and managed using REDCap^{26,27} hosted at the Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo. REDCap is a secure, web-based software platform designed to support data capture for research studies. A system was developed to collect and store data from the WhatsApp messages. A dedicated server component scheduled the messages for groups of participants (grouped according to trial arm and the start date of the messages based on the date of randomization) and pushed these to the participants using the WhatsApp Business interface. The latter provides time stamps indicating when participants receive and open their messages, as well as the answers chosen using the quick reply tool and the content of spontaneous messages sent by participants.

Sample size

With an assumed attrition of 25%, 440–500 randomized individuals would yield 80–85% power to detect a 15 percentage point difference in depression improvement (PHQ-9 < 10) rates between the control and intervention arms at 3 months (25% versus 40%) using a two-sided 5% alpha. Such a difference in improvement rate is considered clinically meaningful^{8,30} but there were no local studies that had investigated a clinically meaningful threshold.

Statistical analysis

Descriptive statistics were obtained using complete cases. All comparative analyses were performed using intention-to-treat principles with imputed data in regression models, aligned to the SAP contingent on the amount of missing data³⁵. Logistic regression was used for binary outcomes; linear regression was used for continuous outcomes. All regression analyses were adjusted for stratification, and all analyses

with an outcome other than PHQ-9 were also adjusted for the baseline score of the corresponding outcome.

Prespecified subgroup analyses at both follow-up assessments were investigated using likelihood ratio tests for interactions between randomized arm and the following: sex; age; educational level; comorbid physical illness (hypertension, diabetes or both); and baseline PHQ-9 severity categories. We also ran a model for the primary outcome adjusting for elapsed time between consenting to participate in the trial and the first or primary follow-up.

Initially, patterns of missingness were investigated by comparing missing with complete data for the baseline demographic characteristics, according to trial arm, at both follow-up visits. To reduce bias and loss of information, we used MICE with 50 imputations, as implemented in the multiple imputation command in Stata under the assumption that data were MAR³⁶. Variables included in the MICE models consist of the 'improved from depression' outcome, covariates described in the SAP (age group, severity of PHQ-9, sex, treatment arm) and variables found to be predictors of missingness^{37,38}. Predictors of missingness included the following: hypertension at baseline, income and levels of enjoyable activities participants engaged in. Also included were baseline scores for corresponding outcome models, including depressive symptomatology (PHQ-9), anxiety symptomatology (GAD-7), health-related quality of life (EQ-5D-5L), capability well-being (ICECAP-O) and loneliness (3-item UCLA).

To assess the sensitivity of our findings against modest departures from the MAR assumption, a weighted sensitivity analysis using the selection model approach was applied^{39–41}. Briefly, once data had been imputed under MAR, parameter estimates from each imputed dataset were reweighted to allow for the data to be missing not at random. To test the stability of our model, we considered different degrees of departure from the MAR assumption by considering plausible values of δ ranging from 0.10 to 0.40. This range corresponds to ORs for the data being observed when a participant improved from depression compared to when they did not, ranging from 1.11 to 1.50 (that is, the exponential of 0.10 and 0.40 respectively).

CACE analysis⁴² using an instrumental variable estimator and imputed data was applied to estimate the effect of the number of messages electronically recorded as being opened on depression outcomes. The protocol¹⁷ and the SAP³⁵ prespecified the 'minimum therapeutic dose' as listening to 'most of the messages' and 'more than half of the messages' received, respectively because it was initially based on self-reported categories ('none', 'a few', 'at least half', 'most of the messages' and 'all of the messages') of a question from the first follow-up assessment. Because in the event we were able to access the electronic data, we operationalized the prespecified minimum therapeutic dose as opening 36 or more messages (75% of the total of 48 messages) versus opening 35 or fewer messages. The CACE analyses were conducted using the PHQ-9 score at 3 and 5 months, adjusting for stratification. We also conducted sensitivity analyses using thresholds of opening at least half of the total messages (24+) versus 23 messages or fewer, and opening all 48 messages versus not doing so.

Regression diagnostics were run for all regression models. The Box–Tidwell test was run after the logistic regression models to test whether the logit transform was a linear function of the predictors for the different models. Normality assumptions for linear regression models were evaluated through residual plots. Statistical tests were two-sided and all analyses were conducted using Stata v.17 (StataCorp LLC).

Ethics and inclusion statement

Most of the authors (10 of 15) are from Brazil, where the RCT took place, and were funded to work on the development and delivery of the intervention, data collection or study management. The authors based in high-income countries are researchers who contributed to the conception of the study and supported the application to obtain

funding shared between Brazil and the UK. This study is relevant for Brazil and other LMICs. One member of the trial steering committee is from the Guarulhos Health Secretary's office, while two other members are researchers based in other cities of Brazil. Meetings with primary care and UBS managers in Guarulhos were held before, during and after the trial. This study was approved by the ethics committee of the Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (Comissão para Análise de Projetos de Pesquisa, ref. 4.097.596) and authorized by the Guarulhos Health Secretary.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

De-identified individual participant data and the data dictionary will be made available 24 months after publication. Proposals with specific aims and an analysis plan should be directed to the corresponding author (M.S.).

Code availability

The code for the data analysis will also be made available 24 months after publication.

References

25. Kroenke, K., Spitzer, R. L. & Williams, J. B. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med. Care* **41**, 1284–1292 (2003).
26. Harris, P. A. et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **42**, 377–381 (2009).
27. Harris, P. A. et al. The REDCap consortium: building an international community of software platform partners. *J. Biomed. Inform.* **95**, 103208 (2019).
28. Donker, T., Griffiths, K. M., Cuijpers, P. & Christensen, H. Psychoeducation for depression, anxiety and psychological distress: a meta-analysis. *BMC Med.* **7**, 79 (2009).
29. Ekers, D. et al. Behavioural activation for depression; an update of meta-analysis of effectiveness and sub group analysis. *PLoS ONE* **9**, e100100 (2014).
30. Scazufca, M. et al. Pilot study of a two-arm non-randomized controlled cluster trial of a psychosocial intervention to improve late life depression in socioeconomically deprived areas of São Paulo, Brazil (PROACTIVE): feasibility study of a psychosocial intervention for late life depression in São Paulo. *BMC Public Health* **19**, 1152 (2019).
31. Spitzer, R. L., Kroenke, K., Williams, J. B. & Löwe, B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch. Intern. Med.* **166**, 1092–1097 (2006).
32. Hughes, M. E., Waite, L. J., Hawkey, L. C. & Cacioppo, J. T. A short scale for measuring loneliness in large surveys: results from two population-based studies. *Res. Aging* **26**, 655–672 (2004).
33. Devlin, N. J. & Krabbe, P. F. M. The development of new research methods for the valuation of EQ-5D-5L. *Eur. J. Health Econ.* **14**, S1–S3 (2013).
34. Grewal, I. et al. Developing attributes for a generic quality of life measure for older people: preferences or capabilities? *Soc. Sci. Med.* **62**, 1891–1901 (2006).
35. PRODIGITAL-D Team. Statistical analysis plan for the PRODIGITAL-D individually randomised controlled trial. figshare.com/s/33a83a1ad01751dcf38c (2022).
36. Lee, K. J. & Carlin, J. B. Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. *Am. J. Epidemiol.* **171**, 624–632 (2010).
37. Sterne, J. A. C. et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* **338**, b2393 (2009).
38. Rubin, D. B. *Multiple Imputation for Nonresponse in Surveys* (John Wiley & Sons, 1987).
39. Carpenter, J. R., Kenward, M. G. & White, I. R. Sensitivity analysis after multiple imputation under missing at random: a weighting approach. *Stat. Methods Med. Res.* **16**, 259–275 (2007).
40. Carpenter, J., Pocock, S. & Lamm, C. J. Coping with missing data in clinical trials: a model-based approach applied to asthma trials. *Stat. Med.* **21**, 1043–1066 (2002).
41. Héraud-Bousquet, V., Larsen, C., Carpenter, J., Desenclos, J.-C. & Le Strat, Y. Practical considerations for sensitivity analysis after multiple imputation applied to epidemiological studies with incomplete data. *BMC Med. Res. Methodol.* **12**, 73 (2012).
42. Peugh, J. L., Strotman, D., McGrady, M., Rausch, J. & Kashikar-Zuck, S. Beyond intent to treat (ITT): a complier average causal effect (CACE) estimation primer. *J. Sch. Psychol.* **60**, 7–24 (2017).

Acknowledgements

We thank the staff of the UBSs in Guarulhos participating in the study, Maria d.J. Assis Ribeiro, Marcelo Bueno da Silva and other members of the Escola SUS-Guarulhos and the Guarulhos Health Secretariat who supported the study; we also thank D. Glavin for supporting the IT team with the development of the system dashboard and Trial Steering Committee and Data Monitoring Committee members, particularly Simon Gilbody, Roberto Lourenço and Rodrigo F. M. Leite. We thank Prof. Walter Colli for his enthusiasm and support with this project. This study was funded by the São Paulo Research Foundation (process no. 2017/50094-2, awarded to M.S.) and the Joint Global Health Trials initiative, jointly funded by the Department of Health and Social Care, the Foreign, Commonwealth & Development Office, the Medical Research Council and Wellcome (process no. MR/RO06229/1, awarded to R.A., T.J.P. and W.H.). We were supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico-Brazil (no. 307579/2019-0, M.S.) and Fundação de Amparo à Pesquisa do Estado de São Paulo (nos. 2018/19343-9 and 2022/05107-7 to C.A.N.; no. 2021/04493-8 to T.V.N.D.; no. 2020/02272-1, F.A.M.; no. 2020/14768-1 to M.O.d.C.; no. 2020/14504-4 to C.H.Q.d.S.; no. 2021/04230-7 to G.M.d.O.; nos. 2021/10148-1 and 2022/08668-0 to M.S.d.S.; and no. 2021/03849-3 to M.M.d.S.M.). For the purpose of open access, we have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. The funding bodies did not have any role in the design of the study or data collection, analyses, or interpretation of the data or in writing the paper.

Author contributions

M.S., C.A.N., W.H., T.J.P. and R.A. conceived the study. C.A.N. and N.S. had full access to all the data. N.S. and T.J.P. were responsible for the data analyses. M.S., C.A.N., N.S., T.J.P. and R.A. wrote the first draft of the paper. M.S., C.A.N., N.S., T.V.N.D., F.A.M., M.O.d.C., C.H.Q.d.S., G.M.d.O., M.S.d.S., L.A.P., M.M.d.S.M., P.v.d.v., W.H., T.J.P. and R.A. edited subsequent revisions of the draft and approved the final paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41591-024-02864-4>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41591-024-02864-4>.

Correspondence and requests for materials should be addressed to Marcia Scazufca or Ricardo Araya.

Peer review information *Nature Medicine* thanks Raphael Machado Castilhos, Tao Chen and the other, anonymous, reviewer(s) for their

contribution to the peer review of this work. Primary Handling Editor: Ming Yang, in collaboration with the *Nature Medicine* team.

Reprints and permissions information is available at www.nature.com/reprints.

Week	Main topics	Period	Monday	Wednesday	Friday	Saturday
1	Introduction to the Viva Vida program Psychoeducation on depression	M	Message 1 (audio)*	Message 3 (audio)	Message 5 (audio)	Message 7 (audio)
		A	Message 2 (audio) Extra message (visual)	Message 4 (audio)	Message 6 (audio)	Message 8 (visual) Quick reply message 1
2	Introduction to the vicious and virtuous cycle of depression Introduction to behavioral activation	M	Message 9 (audio)	Message 11 (audio)*	Message 13 (audio)	Message 15 (audio)
		A	Message 10 (visual)	Message 12 (visual)*	Message 14 (audio)	Quick reply message 2
3	Behavioral activation	M	Message 16 (audio)	Message 18 (audio)	Message 20 (audio)	Message 22 (audio)
		A	Message 17 (audio)	Message 19 (audio)	Message 21 (audio)	Message 23 (visual) Quick reply message 3
4	Planning activities	M	Message 24 (audio)	Message 26 (audio)	Message 28 (audio)	Message 30 (audio)
		A	Message 25 (audio)	Message 27 (audio)	Message 29 (audio)	Message 31 (visual) Quick reply message 4
5	Review of psychoeducation and behavioral activation	M	Message 32 (audio)	Message 34 (audio)	Message 37 (audio)	Message 39 (audio)
		A	Message 33 (audio)	Message 35 (audio) Message 36 (visual)	Message 38 (audio)	Message 40 (visual) Quick reply message 5
6	Psychoeducation on relapse prevention	M	Message 41 (audio)	Message 43 (audio)	Message 45 (audio)	Message 47 (audio)
		A	Message 42 (audio)	Message 44 (audio)	Message 46 (audio)	Message 48 (visual) Quick reply message 6

Extended Data Fig. 1 | Structure of the Viva Vida program. *Abbreviations:* A: afternoon; M: morning. * Example messages with corresponding links are provided for reference - Message 1 (audio), Message 11 (audio) and Message 12 (visual). Week 1: Introduction to the Viva Viva Vida program. Psychoeducation about depression. Extra message with contact of technical support. Week 2: Introduction to the vicious and virtuous cycle of depression, using the metaphor of the 'Wheel of improvement and worsening'. Introduction to behavioral activation using examples of simple things that can help to feel better about the

common symptoms of depression. Week 3: Continues behavioral activation using examples of simple things that can help you feel better about the common symptoms of depression. Week 4: Planning activities, the importance of choosing pleasant and meaningful activities, and how to organise the activities during the week. Week 5: Review of psychoeducation and behavioural activation contents. Week 6: Psychoeducation on relapse prevention, keeping in the virtuous cycle of depression after the end of the program and the importance of early identification of recurrence of depressive symptoms.

Extended Data Table 1 | Comparison of baseline demographics between trial arms for participants followed up at 3 months

Number/total sample size (%)		
	Intervention (n=257)	Control (n=270)
Female, No. (%)	193/257 (75.1)	198/270 (73.3)
Age group, No. (%)		
60-69 years	214/257 (83.3)	225/270 (83.3)
70+ years	43/257 (16.7)	45/270 (16.7)
Education, No. (%)		
None	39/256 (15.2)	44/267 (16.5)
1-4 years	70/256 (27.3)	77/267 (28.8)
5-8 years	83/256 (32.4)	72/267 (27.0)
>8 years	64/256 (25.0)	74/267 (27.7)
Personal income, No. (%)		
Up to 1 MW	172/257 (66.9)	171/270 (63.3)
>1-2 MW	50/257 (19.5)	47/270 (17.4)
>2 MW	35/257 (13.6)	52/270 (19.3)
Hypertension (self-reported), No (%)	178 (69.3)	202/270 (74.8)
Diabetes (self-reported), No (%)	106/257 (41.3)	108/270 (40.0)
Receiving pharmacological treatment for depression (self-reported), No (%)	33/257 (12.8)	41/270 (15.2)
PHQ-9 severity group, No. (%)		
10-14	100/257 (38.9)	105/270 (38.9)
15-19	91/257 (35.4)	88/270 (32.6)
20+	66/257 (25.7)	77/270 (28.5)

MW, minimum wage (in 2021, the minimum wage in Brazil was R\$1,100 (approximately US\$213)); PHQ-9, 9-item Patient Health Questionnaire.

Extended Data Table 2 | Comparison of baseline demographics between trial arms for participants followed up at 5 months

Number/total sample size (%)		
Baseline variables	Intervention (n=250)	Control (n=260)
Female, No. (%)	186/250 (74.4)	188/260 (72.3)
Age group, No. (%)		
60-69 years	210/250 (84.0)	216/260 (83.1)
70+ years	40/250 (16.0)	44/260 (16.9)
Education, No. (%)		
None	38/249 (15.3)	41/257 (16.0)
1-4 years	63/249 (25.3)	75/257 (29.2)
5-8 years	81/249 (32.5)	70/257 (27.2)
>8 years	67/249 (26.9)	71/257 (27.6)
Personal income, No. (%)		
Up to 1 MW	163/250 (65.2)	162/260 (62.3)
>1-2 MW	50/250 (20.0)	46/260 (17.7)
>2 MW	37/250 (14.8)	52/260 (20.0)
Hypertension (self-reported), No (%)	168/250 (67.2)	193/260 (74.2)
Diabetes (self-reported), No (%)	106/250 (42.4)	102/260 (39.2)
Receiving pharmacological treatment for depression (self-reported), No (%)	33/250 (13.2)	39/260 (15.0)
PHQ-9 severity group, No. (%)		
10-14	102/250 (40.8)	104/260 (40.0)
15-19	87/250 (34.8)	87/260 (33.5)
20+	61/250 (24.4)	69/260 (26.5)

MW, minimum wage (in 2021, the minimum wage in Brazil was R\$1,100 (approximately US\$213)); PHQ-9, 9-item Patient Health Questionnaire.

Extended Data Table 3 | Comparison between participants with complete data and participants with missing data at the 3-month follow-up assessment

	Number/total sample size (%)			
	Intervention		Control	
	Complete data 257/298 (86.2)	Missing PHQ-9 41/298 (13.8)	Complete data 270/305 (88.5)	Missing PHQ-9 35/305 (11.5)
Female, No. (%)	193/257 (75.1)	32 (78.1)	198/270 (73.3)	28/35 (80.0)
Age group, No. (%)				
60-69 years	214/257 (83.3)	37/41 (90.2)	225/270/270 (83.3)	26/35 (74.3)
70+ years	43/257 (16.7)	4/41 (9.8)	45 (16.7)	9/35 (25.7)
Formal education, No. (%)	224/257 (87.2)	36/41 (87.8)	231/270 (85.9)	29/35 (82.9)
Education (level), No. (%)				
None	39/256 (15.2)	6/41 (14.6)	44/267 (16.5)	7/35 (20.0)
1-4 years	70/256 (27.3)	13/41 (31.7)	77/267 (28.8)	7/35 (20.0)
5-8 years	83/256 (32.4)	11/41 (26.8)	72/267 (27.0)	13/35 (37.1)
>8 years	64/256 (25.0)	11/41 (26.8)	74/267 (27.7)	8/35 (22.9)
Job, No. (%)	68/257 (26.5)	17/41 (41.5)	74/270 (27.4)	10/35 (28.6)
Personal income, No. (%)				
Up to 1 MW	172/257 (66.9)	25/41 (61.0)	171/270 (63.3)	20/35 (57.1)
>1-2 MW	50/257 (19.5)	10/41 (24.4)	47/270 (17.4)	9/35 (25.7)
>2 MW	35/257 (13.6)	6/41 (14.6)	52/270 (19.3)	6/35 (17.1)
Household income, No. (%)				
Up to 1 MW	92/249 (37.0)	17/41 (41.5)	94/260 (36.2)	11/34 (32.4)
>1-2 MW	75/249 (30.1)	10/41 (24.4)	78/260 (30.0)	10/34 (29.4)
>2 MW	82/249 (32.9)	14/41 (34.2)	88/260 (33.9)	13/34 (38.2)
Hypertension (self-reported), No. (%)	178/257 (69.3)	29/41 (70.7)	202/270 (74.8)	20/35 (57.1)
Diabetes (self-reported), No. (%)	106/257 (41.3)	17/41 (41.5)	108/270 (40.0)	13/35 (37.1)
PHQ-9 severity group, No. (%)				
10-14	100/257 (38.9)	20/41 (48.8)	105/270 (38.9)	14/35 (40.0)
15-19	91/257 (35.4)	8/41 (19.5)	88/270 (32.6)	12/35 (34.3)
20+	66/257 (25.7)	13/41 (31.7)	77/270 (28.5)	9/35 (25.7)
Engagement in enjoyable, meaningful activities over the past two weeks, No. (%)				
Not at all	44/257 (17.1)	11/40 (27.5)	56/270 (20.7)	10/35 (28.6)
Several days	40/257 (15.6)	8/40 (20.0)	49/270 (18.2)	2/35 (5.7)
Over half the days	57/257 (22.2)	4/40 (10.0)	50/270 (18.5)	2/35 (5.7)
Nearly every day	116/257 (45.1)	17/40 (42.5)	115/270 (42.6)	21/35 (60.0)

MW, minimum wage (in 2021, the minimum wage in Brazil was R\$1,100 (approximately US\$213)); PHQ-9, 9-item Patient Health Questionnaire.

Extended Data Table 4 | Comparison between participants with complete data and participants with missing data at the 5-month follow-up assessment

	Number/total sample size (%)			
	Intervention		Control	
	Complete data 250/298 (83.9)	Missing PHQ-9 48/298 (16.1)	Complete data 260/305 (85.2)	Missing PHQ-9 45/305 (14.8)
Female, No. (%)	186/250 (74.4)	39/48 (81.3)	188/260 (72.3)	38/45 (84.4)
Age group, No. (%)				
60-69 years	210/250 (84.0)	41/48 (85.4)	216/260 (83.1)	35/45 (77.8)
70+ years	40/250 (16.0)	7/48 (14.6)	44/260 (16.9)	10/45 (22.2)
Formal education, No. (%)	218/250 (87.2)	42/48 (87.5)	223/260 (86.1)	37/45 (82.2)
Education (level), No. (%)				
None	38/249 (15.3)	7/48 (14.6)	41/257 (16.0)	10/45 (22.2)
1-4 years	63/249 (25.3)	20/48 (41.7)	75/257 (29.2)	9/45 (20.0)
5-8 years	81/249 (32.5)	13/48 (27.1)	70/257 (27.2)	15/45 (33.3)
>8 years	67/249 (26.9)	8/48 (16.7)	71/257 (27.6)	11/45 (24.4)
Job, No. (%)	66/250 (26.4)	19/48 (39.6)	69/260 (26.5)	15/45 (33.3)
Personal income, No. (%)				
Up to 1 MW	163/250 (65.2)	34/48 (70.8)	162/260 (62.3)	29/45 (64.4)
>1-2 MW	50/250 (20.0)	10/48 (20.8)	46/260 (17.7)	10/45 (22.2)
>2 MW	37/250 (14.8)	4/48 (8.3)	52/260 (20.0)	6/45 (13.3)
Household income, No. (%)				
Up to 1 MW	89/242 (36.8)	20/48 (41.7)	90/250 (36.0)	15/44 (34.1)
>1-2 MW	70/242 (28.9)	15/48 (31.3)	77/250 (30.8)	11/44 (25.0)
>2 MW	83/242 (34.3)	13/48 (27.1)	83/250 (33.2)	18/44 (40.9)
Hypertension (self-reported), No. (%)	168/250 (67.2)	39/48 (81.3)	193/260 (74.2)	29/45 (64.4)
Diabetes (self-reported), No. (%)	106/250 (42.4)	17/48 (35.4)	102/260 (39.2)	19/45 (42.2)
PHQ-9 severity group, No. (%)				
10-14	102/250 (40.8)	18/48 (37.5)	104/260 (40.0)	15/45 (33.3)
15-19	87/250 (34.8)	12/48 (25.0)	87/260 (33.5)	13/45 (28.9)
20+	61/250 (24.4)	18/48 (37.5)	69/260 (26.5)	17/45 (37.8)
Engagement in enjoyable, meaningful activities over the past two weeks, No. (%)				
Not at all	42/249 (16.9)	13/48 (27.1)	56/260 (21.5)	10/45 (22.2)
Several days	41/249 (16.5)	7/48 (14.6)	49/260 (18.9)	2/45 (4.4)
Over half the days	54/249 (21.7)	7/48 (14.6)	43/260 (16.5)	9/45 (20.0)
Nearly every day	112/249 (45.0)	21/48 (43.8)	112/260 (43.1)	24/45 (53.3)

MW, minimum wage (in 2021, the minimum wage in Brazil was R\$1,100 (approximately US\$213)); PHQ-9, 9-item Patient Health Questionnaire.

Extended Data Table 5 | Comparison of adjusted estimates from multiple imputed data using the MICE models, with estimates from complete case analysis for the improvement from depressive symptomatology (PHQ-9 scores <10) at 5 months, and for a reduction in PHQ-9 scores by at least 50% between baseline and 3 and 5 months (secondary outcome)

Model	Complete case analysis ^c		Imputed data from MICE analyses ^{c,d}	
	Odds ratio (95% CI)	P Value ^e	Odds ratio (95% CI)	P Value ^e
Improvement from depressive symptomatology at five months ^a (n=510)	1.02 (0.71, 1.46)	.911	1.01 (0.70, 1.46)	.938
Reduction in depressive symptomatology at three months ^b (n=527)	1.59 (1.10, 2.31)	.015	1.58 (1.08, 2.29)	.016
Reduction in depressive symptomatology at five months ^b (n=510)	1.06 (0.75, 1.51)	.740	1.04 (0.73, 1.48)	.841

MICE, multiple imputation by chained equations; OR, odds ratio; PHQ-9, 9-item Patient Health Questionnaire ^aThe secondary outcome of improvement from depressive symptomatology was defined as PHQ-9 scores less than ten. ^bThe secondary outcome of reduction in depressive symptomatology was defined as PHQ-9 scores reducing by at least 50% between baseline and the first and second follow-up assessments (yes/no). ^cOdds ratios and 95% CIs were calculated using logistic regression models adjusted for sex, age group and baseline severity of PHQ-9 (categorical). ^dAll estimates had missing data imputed by trial arm using MICE models that included sex, age group, baseline severity of PHQ-9 (categorical) and predictors of missingness. ^eAll P values reported are based on two-sided Wald tests from the relevant logistic regression model, with no adjustments for multiple tests.

Extended Data Table 6 | Comparison of estimates from multiple imputed data using MICE models, with estimates from complete case analysis for secondary outcomes at 3 and 5 months

Model	Complete case analysis ^a		Imputed data from MICE analyses ^{a,b}	
	Difference in means (95% CI)	P Value ^c	Difference in means (95% CI)	P Value ^c
PHQ-9 scores at three months	-1.09 (-2.28, 0.09)	.070	-1.14 (-2.33, 0.05)	.060
GAD-7 scores at three months	-0.78 (-1.78, 0.23)	.130	-0.89 (-1.88, 0.11)	.082
EQ-5D-5L scores at three months	0.015 (-0.010, 0.039)	.246	0.025 (-0.015, 0.064)	.218
ICECAP-O scores at three months	0.021 (-0.006, 0.048)	.132	0.018 (-0.009, 0.045)	.185
3-item UCLA scores at three months	0.18 (-0.10, 0.46)	.213	0.19 (-0.09, 0.48)	.186
PHQ-9 scores at five months	-0.34 (-1.50, 0.83)	.567	-0.21 (-1.36, 0.94)	.716
GAD-7 scores at five months	-0.25 (-1.33, 0.84)	.657	-0.14 (-1.20, 0.92)	.793
EQ-5D-5L scores at five months	0.020 (-0.003, 0.043)	.084	0.035 (-0.001, 0.071)	.059
ICECAP-O scores at five months	0.003 (-0.022, 0.029)	.804	0.002 (-0.024, 0.029)	.856
3-item UCLA scores at five months	0.22 (-0.07, 0.50)	.142	0.22 (-0.07, 0.52)	.140

CI, confidence interval; EQ-5D-5L, 5-level EuroQol health-related quality of life questionnaire; GAD-7, 7-item General Anxiety Disorder questionnaire; ICECAP-O, ICEpop CAPability measure for older people; PHQ-9, Patient Health Questionnaire-9; 3-item UCLA, University of California, Los Angeles. PHQ-9 scores range from 0 to 27, with higher scores representing more severe depressive symptomatology. GAD-7 scores range from 0 to 21, with higher scores representing more severe anxiety symptomatology. EQ-5D-5L scores range from -0.264 to 1, with higher scores representing higher quality of life. ICECAP-O scores range from 0 to 1, with higher scores representing greater levels of capability well-being. 3-item UCLA scores range from 3 to 9, with higher scores representing more severe levels of loneliness. ^aDifferences in means were estimated using linear regression models, adjusted for relevant baseline assessment of corresponding outcome, sex, age group and baseline severity of PHQ-9 (categorical). ^bAll estimates had missing data imputed separately, according to trial arm, using MICE models that included predictors of missingness, sex, age group and baseline severity of the PHQ-9 score (categorical). ^cAll P values reported are based on two-sided t-tests from the relevant linear regression model, with no adjustments for multiple tests.

Extended Data Table 7 | Results from subgroup analyses testing for the effect modification between prespecified variables and treatment arm at 3 and 5 months for the primary and secondary outcomes of improvement from depressive symptomatology using complete data only

Interaction term with treatment allocation	<i>P</i> Value of Wald test ^a	
	Three-month follow-up ^b	Five-month follow-up ^b
Gender	.953	.055
Age	.660	.351
Education	.723	.284
Baseline PHQ-9 scores	.113	.565
Hypertension	.061	.069
Diabetes	.505	.440
Comorbid hypertension and diabetes	.119	.445

OR, odds ratio; PHQ-9: 9-item Patient Health Questionnaire. ^aAll *P* values reported are based on two-sided Wald tests, with no adjustments for multiple tests. ^bORs and 95% CIs were calculated using logistic regression models adjusted for sex, age group and baseline severity of PHQ-9 (categorical).

Extended Data Table 8 | Adjusted ORs (95% CIs) for different departures from the missing at random assumption, for the outcome variable of improvement from depressive symptomatology (PHQ-9 scores <10) at 3 and 5 months assuming greater probability of the outcome being missing when depression was present

Delta	Adjusted Odds Ratio (95% CI)	
	Three-month follow-up	Five-month follow-up
0.1	1.56 (1.07, 2.27)	1.05 (0.74, 1.52)
0.2	1.54 (1.06, 2.24)	1.06 (0.74, 1.51)
0.3	1.53 (1.06, 2.22)	1.05 (0.73, 1.49)
0.4	1.52 (1.06, 2.20)	1.03 (0.73, 1.46)

CI, confidence interval; OR, odds ratio; PHQ-9, 9-item Patient Health Questionnaire.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

De-identified individual participant data and data dictionary will be made available 24 months after publication. Proposals with specific aims and an analysis plan should be directed to the corresponding author.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	Gender was self-reported. Subgroup analyses, including for gender, are also reported in the study.
Reporting on race, ethnicity, or other socially relevant groupings	This study does not report any information on race or ethnicity.
Population characteristics	The majority of participants were women (75%), aged 60-69 year-old (83%), with less than eight years of education (74%) and were earning minimum wage (64%). Severity of depressive symptomatology was considered moderate (PHQ-9 scores between 10 and 14) for 40% of the participants.
Recruitment	Participants were recruited by phone following a list of individuals registered with the participating primary care clinics. Individuals with an active WhatsApp number were approached by telephone to be screened for depressive symptomatology using the Patient Health Questionnaire (PHQ). Those with PHQ-2 scores ≥ 1 were administered the full PHQ-9 and if they scored ≥ 10 as well as met our other eligibility criteria they were invited to participate in the trial. The list of individuals were provided by the Guarulhos Health Secretariat and we pre-screened 95% of all registered individuals. Around 72% of those with a valid WhatsApp number could not be included as they did not answer either of the two calls made by the research assistants (these calls were on different days but logistical constraints precluded more than two such attempts). Possible reasons for failure to answer the call include that the telephone was turned off, the call was not heard, or the individual did not answer calls from unknown numbers. There is therefore potential for an element of self-selection in this part of the recruitment process although this would not have impacted appreciably on the internal validity of the cluster randomised trial.
Ethics oversight	Ethics Committee of the Hospital das Clínicas da Faculdade de Medicina da Universidade de Sao Paulo – HCFMUSP

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	PRODIGITAL-D was a two-arm individually-randomized controlled trial with 1:1 allocation. This study reports the effectiveness outcomes of the trial.
Research sample	Older adults (60+ years) registered with a primary care clinic in Guarulhos, Brazil, and with a valid phone and WhatsApp number. The majority of participants were women (75%) and aged 60-69 year-old (83%). The majority had less than eight years of education (74%) and were earning minimum wage (64%). The sample is representative of the older population with depressive symptoms in Brazil. The Viva Vida program was developed to reach older adults registered with primary care clinics, commonly located in socioeconomically deprived areas of the country.
Sampling strategy	Random sampling was employed to recruit participants. Sample size was calculated considering that 440-500 randomized individuals would yield 80-85% power to detect a 15-percentage point difference in depression recovery rates between the control and intervention arms at three months (25% versus 40%) using a two-sided 5% alpha, and assuming a 25% attrition.
Data collection	The following instruments were used during the baseline and follow-up assessments: 9-item Patient Health Questionnaire (PHQ-9), Generalized Anxiety Disorder-7 (GAD-7), 3-item University of California, Los Angeles (UCLA) Loneliness scale (3-item UCLA), European Quality of Life five-dimensional questionnaire, five-level version (EQ-5D-5L); ICEpop CAPability measure for Older people (ICECAP-O). All data were collected by phone calls and managed using REDCap. Research assistants involved in recruitment and follow-up data collection were masked to trial allocation. Two different groups of research assistants were responsible for either the first or the second follow-up assessment.
Timing	Data collection took place between 8 September 2021 and 26 September 2022.
Data exclusions	No data were excluded from the analyses.
Non-participation	No participants dropped out/declined participating after randomisation.

Randomization

Participants were stratified by age group (60-69 years, 70-79, 80+ years), gender (male, female) and depression severity (PHQ-9 categories: 10-14, 15-19, 20+). The allocation sequence was generated using random-permuted blocks with random block sizes by research team members not involved in data collection. The 'randomization module' of the Research Electronic Data Capture (REDCap) was used to conceal the allocation sequence and randomize individuals.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	Brazilian Registry of Clinical Trials (ReBEC), RBR-4c94dtn.
Study protocol	Full trial protocol is part of the supplemental material.
Data collection	Participants were recruited by phone from a randomly ordered list of 24 primary care clinics, named Unidades Básicas de Saúde (UBSs), in Guarulhos, Brazil. The city is part of the metropolitan region of Sao Paulo with a population of around 1.4 million. A total of 603 participants were recruited between September 2021 and April 2022, and followed up for five months.
Outcomes	Recovery from depression (PHQ-9 score<10) at three months was the primary outcome. Secondary outcomes included recovery from depression at five months, and reduction in PHQ-9 scores by at least 50% between baseline and follow-up visits at three and five months. We also evaluated effects on the continuous scores of: anxiety symptomatology with the Generalized Anxiety Disorder-7 (GAD-7); loneliness with the 3-item University of California, Los Angeles (UCLA) loneliness scale (3-item UCLA); quality of life with the European Quality of Life five-dimensional questionnaire, five-level version (EQ-5D-5L); and capability wellbeing with the ICEpop CAPability measure for Older people (ICECAP-O).