JAMA Internal Medicine | Original Investigation

Interactive Voice Response–Based Self-management for Chronic Back Pain
The COPES Noninferiority Randomized Trial

Alicia A. Heapy, PhD; Diana M. Higgins, PhD; Joseph L. Goulet, PhD; Kathryn M. LaChappelle, MPH; Mary A. Driscoll, PhD; Rebecca A. Czlapinski, MA; Eugenia Buta, PhD; John D. Piette, PhD; Sarah L. Krein, PhD; Robert D. Kerns, PhD

➕ Supplemental content

IMPORTANCE Recommendations for chronic pain treatment emphasize multimodal approaches, including nonpharmacologic interventions to enhance self-management. Cognitive behavioral therapy (CBT) is an evidence-based treatment that facilitates management of chronic pain and improves outcomes, but access barriers persist. Cognitive behavioral therapy delivery assisted by health technology can obviate the need for in-person visits, but the effectiveness of this alternative to standard therapy is unknown. The Cooperative Pain Education and Self-management (COPES) trial was a randomized, noninferiority trial comparing IVR-CBT to in-person CBT for patients with chronic back pain.

OBJECTIVE To assess the efficacy of interactive voice response–based CBT (IVR-CBT) relative to in-person CBT for chronic back pain.

DESIGN, SETTING, AND PARTICIPANTS We conducted a noninferiority randomized trial in 1 Department of Veterans Affairs (VA) health care system. A total of 125 patients with chronic back pain were equally allocated to IVR-CBT (n = 62) or in-person CBT (n = 63).

INTERVENTIONS Patients treated with IVR-CBT received a self-help manual and weekly prerecorded therapist feedback based on their IVR-reported activity, coping skill practice, and pain outcomes. In-person CBT included weekly, individual CBT sessions with a therapist. Participants in both conditions received IVR monitoring of pain, sleep, activity levels, and pain coping skill practice during treatment.

MAIN OUTCOMES AND MEASURES The primary outcome was change from baseline to 3 months in unblinded patient report of average pain intensity measured by the Numeric Rating Scale (NRS). Secondary outcomes included changes in pain-related interference, physical and emotional functioning, sleep quality, and quality of life at 3, 6, and 9 months. We also examined treatment retention.

RESULTS Of the 125 patients (97 men, 28 women; mean [SD] age, 57.9 [11.6] years), the adjusted average reduction in NRS with IVR-CBT (−0.77) was similar to in-person CBT (−0.84), with the 95% CI for the difference between groups (−0.67 to 0.80) falling below the prespecified noninferiority margin of 1 indicating IVR-CBT is noninferior. Fifty-four patients randomized to IVR-CBT and 50 randomized to in-person CBT were included in the analysis of the primary outcome. Statistically significant improvements in physical functioning, sleep quality, and physical quality of life at 3 months relative to baseline occurred in both treatments, with no advantage for either treatment. Treatment dropout was lower in IVR-CBT with patients completing on average 2.3 (95% CI, 1.0-3.6) more sessions.

CONCLUSIONS AND RELEVANCE IVR-CBT is a low-burden alternative that can increase access to CBT for chronic pain and shows promise as a nonpharmacologic treatment option for chronic pain, with outcomes that are not inferior to in-person CBT.

TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT01025752

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Alicia A. Heapy, PhD, VA Connecticut Healthcare System (11ACLGS), 950 Campbell Ave, West Haven, CT 06516 (alicia.heapy@va.gov).

Chronic pain affects more adults in the United States than diabetes, heart disease, and cancer combined and is associated with limitations in physical and emotional functioning, and quality of life.[1] The dramatic increase in the use of opioid medications to treat chronic pain has led to concern about the harms associated with long-term opioid therapy (eg, addiction, overdose, and death) and lack of evidence for long-term benefits.[2-4] As a result, the Institute of Medicine,[1] US Department of Health and Human Services,[5] and Centers for Disease Control and Prevention[6] all advise against opioid medications as a first-line treatment or monotherapy for chronic pain and instead emphasize multimodal care that includes evidence-based nonpharmacologic approaches. Cognitive behavioral therapy (CBT) is a low-risk psychological intervention that is effective in reducing pain and improving function for numerous pain complaints, including back pain,[7] osteoarthritis,[8] and fibromyalgia.[9] Cognitive behavioral therapy assists patients with reconceptualizing pain as influenced by not only biological, but psychological, behavioral, and social factors. Patients learn cognitive (eg, reframing catastrophic thoughts) and behavioral (eg, relaxation techniques) pain coping skills over 6 to 12 treatment sessions.

Although CBT is efficacious, many patients experience travel and scheduling barriers to attending in-person sessions. Cherkin et al[10] found that both CBT and mindfulness-based stress reduction produced clinically significant improvements in pain and functioning relative to usual care; however, only half of participants attended 6 of the 8 sessions. In addition, trained therapists are scarce and concentrated in urban areas or near academic medical centers,[11] further hindering access. Using mobile health technology, such as interactive voice response (IVR), may improve access to CBT for chronic pain.

Interactive voice response is an automated telephonic technology that allows patients to report symptoms, functioning, and pain coping skill use and receive prerecorded information and feedback via their telephone. Interactive voice response–based interventions have been used successfully to provide education, peer support, and tailored messages to enhance treatment adherence and maintenance of gains.[12-16] Naylor and colleagues[14,15,17] found maintenance or improvements in CBT treatment effects and reduction in prescription opioid use when IVR monitoring was used to reinforce in-person CBT. Despite these promising results, to our knowledge there have been no trials of a solely IVR-delivered treatment for chronic pain.

We hypothesized that IVR-based CBT (IVR-CBT) would not be significantly less efficacious than in-person treatment and would provide greater access to care with less patient burden. Specifically, IVR-CBT would produce reductions in pain intensity that were not meaningfully inferior to in-person CBT. We also compared pain-related physical and emotional functioning and health-related quality of life for participants in both conditions. Finally, we examined IVR-CBT relative to in-person CBT in terms of treatment retention.

Key Points

Question What is the efficacy of interactive voice response–based cognitive behavioral therapy (IVR-CBT) relative to in-person CBT for chronic back pain?

Findings In this noninferiority randomized clinical trial of 125 patients with chronic back pain, the average pain reduction was similar for both conditions with the difference between groups falling below the noninferiority margin of 1, indicating that IVR-CBT is noninferior to in-person CBT for chronic back pain.

Meaning IVR-CBT is a low-burden alternative that can increase access to CBT for chronic pain and shows promise as a nonpharmacologic treatment option, with outcomes that are not inferior to in-person CBT.

Methods

Design, Study Setting, and Participants

The Cooperative Pain Education and Self-management (COPES) trial was a randomized, noninferiority trial comparing IVR-CBT to in-person CBT for patients with chronic back pain. A description of the COPES trial has been published elsewhere[18] (see Supplement 1). Participants were veterans from 1 VA health care system. Eligibility screening was conducted using an electronic health record search for back pain–related diagnoses and at least moderate pain intensity (≥4 on the 0-10 numeric rating scale or NRS) reported during the most recent primary care visit. Potentially eligible patients were mailed an opt-out recruitment letter. In addition, recruitment flyers were posted in the medical center. Eligibility criteria included having an electronic health record–verified back condition, average pain score of at least 4 on the NRS for at least 3 months, self-reported ability to walk 1 block, absence of planned surgical intervention for pain, access to a touchtone telephone, and absence of medical or psychiatric conditions that could impair participation (eg, terminal cancer, severe depression [Beck Depression Inventory-II, BDI-II, score >28],[19] dementia based on the St Louis University Mental Status [SLUMS]) score,[20] or active psychosis or substance use disorder based on the Mini-International Neuropsychiatric Interview[21]).

After providing written informed consent, eligible patients were enrolled by the study research assistant (RA) (K.M.L.), completed baseline survey measures, and were given an Omron GoSmart HJ-112N pedometer. Patients were compensated for completing the assessments: $20 for baseline, $30 for the 3-month assessment, $40 for the 6-month assessment, and $50 for the 9-month assessment, total of $140 if all assessments were completed. The study was approved by the VA Connecticut Healthcare System and Yale School of Medicine institutional review boards. Recruitment occurred from June 2012 to July 2015. All follow-up assessments were completed by April 1, 2016.

Randomization

Following the baseline assessment, participants were assigned in equal proportions, in block sizes of 4, and stratified by distance from the medical center, and back pain type (nonspecific, radicular, associated with other spinal causes).[22] The

randomization schedule was created by an independent biostatistician using SAS statistical software (version 8.2; SAS Institute Inc) and concealed in the study database until assignment to an intervention. After randomization, participants were assigned to a therapist by the RA based on therapist availability. Therapists were doctoral-level clinical psychologists or advanced-practice psychiatric nurse practitioners and followed treatment manuals to deliver both interventions.

Intervention

All participants received a treatment manual specific to their intervention that had been adapted from prior trials and that was designed to be delivered over 10 weeks.[23] Each manual included an introductory module that presented the rationale for CBT, 8 pain coping skill modules, and a relapse prevention module. All participants received 11 weeks of daily IVR assessment of pain, sleep, step count, and pain coping skill practice, engaged in a progressive walking program, and continued to receive care from their primary care clinician.

In-person CBT participants attended weekly in-person 30- to 40-minute treatment sessions. Therapists reviewed the IVR reports and provided feedback during sessions.

The IVR-CBT therapists reviewed the IVR-reported information and recorded 2- to 5-minute personalized feedback patterned after prior studies using IVR.[14,17] Feedback was delivered weekly during participants' regular IVR assessments. Details of the intervention are published elsewhere.[18]

Measures

We followed recommendations from the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials, for domains to assess and selection of measures.[24] Unblinded participants independently completed measures in person, via mail, or online at baseline, 3, 6, and 9 months after baseline, with 3 months being the primary end point. Completion method was selected by the participant. Therapists did not participate in outcome data collection. Prior to treatment, participants in both groups underwent daily IVR assessment of pain intensity, sleep duration, and pedometer-measured step counts for 7 days to establish a baseline of function.[25]

Primary Outcome

Average pain intensity over the past week was assessed using an 11-point NRS (0, no pain; 10, worst pain imaginable).[26] Originally, we planned to use the mean of NRS least, average, and worst pain intensity over the past week. Instead, we elected to use average NRS alone because it is more commonly used and therefore promotes cross-study comparisons. The correlation between the 2 measures is 0.95 and noninferiority results using the original outcome did not differ from current findings (eTable 1 in Supplement 2).

Secondary Outcomes

The Interference subscale of the West Haven-Yale Multidimensional Pain Inventory[27] and the Roland and Morris Disability Questionnaire (RMDQ)[28] were used to assess pain-related interference and physical functioning, respectively. Higher scores represent greater interference and poorer function. Depressive symptom severity was assessed using the BDI-II.[29] Higher scores indicate more depressive symptomatology. Sleep quality was assessed with the Pittsburgh Sleep Quality Index.[30] Scores of 5 or greater indicate poor sleep quality. The Veterans 36-item short-form (SF-36) questionnaire was used to assess physical and mental health–related quality of life.[31] Lower scores represent poorer quality of life. All measures have strong psychometric properties.[19,27-29,31,32]

Treatment Feasibility and Fidelity

Treatment weeks completed (maximum, 10 weeks) was defined as in-person CBT attendance or completion of at least 1 IVR call during the treatment week.

Treatment fidelity was assessed using a revised version of the Yale Adherence and Competence Scale.[33] A clinician experienced in chronic pain CBT rated a convenience sample of 30% of the audiotaped in-person CBT sessions for adherence and competence. Average scores of 4 or greater out of 7, indicating adequate adherence (99% of sessions) and competence (96% of sessions), were achieved. In IVR-CBT, therapists used templated scripts to enhance treatment fidelity when developing prerecorded IVR feedback.

Adverse Events

Walking-related adverse events were collected weekly during IVR assessment calls. Additional adverse events were identified by medical record review at follow-up.
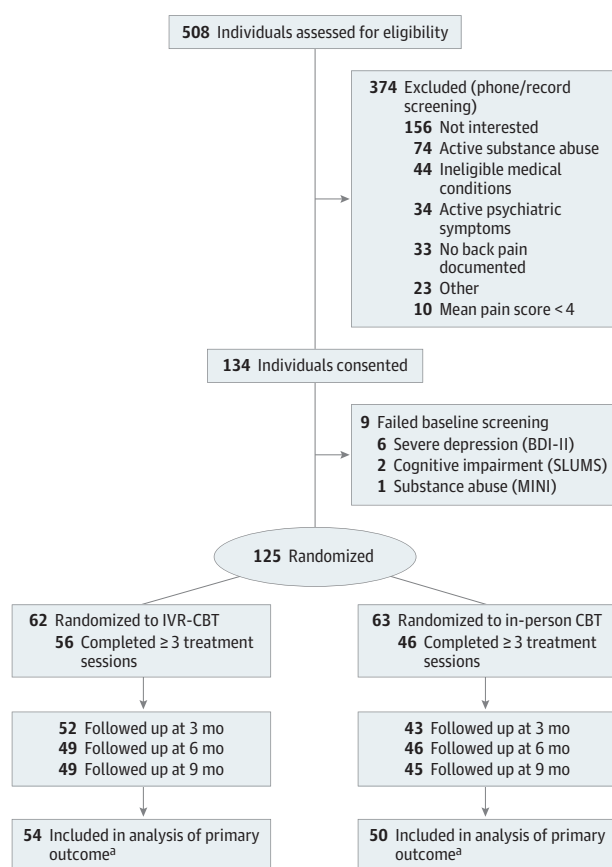
The sample size was based on a noninferiority test comparing IVR-CBT with in-person CBT for NRS pain intensity, with 80% power, type I error (1-sided) of .025, and assuming the true difference between means is 0. In preliminary data from similar participants,[23] the baseline mean (SD) NRS was 7.0 (2.4). Since a 20% (1.4%) reduction in NRS from baseline was considered minimally clinically important,[26] the noninferiority margin was set at 1 NRS unit. If the true mean NRS for IVR-CBT was less than 1 unit higher than in-person CBT, IVR-CBT would be deemed noninferior. This hypothesis can be tested by comparing the limit of the 95% CI for the mean difference between IVR-CBT and in-person CBT with the noninferiority margin 1; if this upper limit falls below 1, IVR-CBT is noninferior. Including a 15% inflation for losses, the required sample size was 230.

Because enrollment was slower than expected and the SD in the primary outcome was less than the initial power calculation (1.6 vs 2.4), we conducted a sample size recalculation based on observed data to determine if enrollment could be stopped prior to meeting the initial target of 230. Using the baseline SD (1.6), we determined that we would need 42 completers per group for 80% power. Enrollment was stopped at the end of the no-cost extension period with 125 randomized participants.

Statistical Analysis

Randomization adequacy was assessed by comparing baseline demographic and clinical characteristics between groups. Between-group comparisons of changes from baseline in primary and secondary outcomes used mixed-effects models for repeated measurements. The original primary analysis pro-

Figure 1. Participant Flow Through Trial Comparing IVR-CBT
With In-Person CBT for Back Pain

508 Individuals assessed for eligibility

374 Excluded (phone/record
     screening)
156 Not interested
 74 Active substance abuse
 44 Ineligible medical
     conditions
 34 Active psychiatric
     symptoms
 33 No back pain
     documented
 23 Other
 10 Mean pain score < 4

134 Individuals consented

9 Failed baseline screening
6 Severe depression (BDI-II)
2 Cognitive impairment (SLUMS)
1 Substance abuse (MINI)

125 Randomized

62 Randomized to IVR-CBT
56 Completed ≥ 3 treatment
   sessions

63 Randomized to in-person CBT
46 Completed ≥ 3 treatment
   sessions

52 Followed up at 3 mo
49 Followed up at 6 mo
49 Followed up at 9 mo

43 Followed up at 3 mo
46 Followed up at 6 mo
45 Followed up at 9 mo

54 Included in analysis of primary
   outcome[a]

50 Included in analysis of primary
   outcome[a]

[a] Patients who had at least 1 postbaseline assessment. BDI-II indicates Beck
Depression Inventory II; CBT, cognitive behavioral therapy; IVR, interactive
voice response; MINI, Mini International Neuropsychiatric Interview Version
6.0.0; SLUMS, Saint Louis University Mental Status examination.

posed a 1-sided *t* test and mixed models for secondary analyses; before analyzing the data, we changed all analyses to mixed models because these models better accommodate missing data and allow adjustment for stratification variables. The models included treatment, time (3, 6, and 9 months), treatment-by-time interaction terms, baseline value of the outcome, and stratification variables (distance from VA facility and back pain cause). An unstructured variance-covariance matrix was used to model the within-subject correlations. Results were summarized as least-squares (LS) means (95% CI) within and between groups at each time point. Following CONSORT recommendations,[34] noninferiority analyses were conducted on an intent-to-treat (ITT) and a per-protocol (completed ≥3 treatment weeks) basis at α = .025. All other analyses were ITT at α = .05.

We conducted a responder analysis of the 2 most commonly examined outcomes in pain treatment trials, pain intensity and physical functioning. A response was defined as having a clinically meaningful improvement of at least 30%[35] at 3 months in NRS and RMDQ.

Sensitivity analyses were conducted that included baseline variables associated with primary outcome absence as co-

variates in models evaluating intervention effect. Additional sensitivity analyses used multiple imputation for missing variables.[36] The imputation model included the primary and secondary outcomes measured at all time points, treatment, stratification, and baseline variables identified as predictive of absence. Results were based on 100 imputations. Analyses were conducted using SAS software (version 9.4) and R software (version 3.2.3; R Foundation).

Results

Among 508 individuals assessed for eligibility, 134 were consented, 9 failed a baseline screening measure, and 125 were randomized (IVR-CBT, 62; in-person CBT, 63) (Figure 1). Eighteen randomized participants withdrew prior to attending any treatment sessions (IVR-CBT, 4; in-person CBT, 14; *P* = .02).

The groups were similar at baseline with 80 (64.5%) white patients and 28 (22.4%) women. The mean age of the sample was 57.9 years, and the average NRS pain rating indicated moderate pain (mean, 5.58) (Table 1).

Of those randomized, 102 (82%) completed at least 3 sessions, which constituted receiving a per-protocol "dose" of treatment (IVR-CBT, 56; in-person CBT, 46; *P* = .02). The IVR-CBT participants attended more sessions than those in in-person CBT (mean of 8.9 vs 6.6; difference, 2.3; 95% CI, 1.0-3.6; *P* < .001) (Figure 2). Follow-up response rate was 76% at 3 months (IVR-CBT 84% vs in-person CBT 68%), 76% at 6 months (79% vs 73%), and 75% at 9 months (79% vs 71%). The primary analysis included 104 participants (87% IVR-CBT participants; 79% in-person CBT participants) who had at least 1 postbaseline assessment of past-week NRS pain intensity rating (see eTable 2 in Supplement 2 for baseline characteristics of these participants).

Primary Outcome

Immediately posttreatment (3-month postbaseline), the adjusted change from baseline in NRS was −0.77 (95% CI, −1.39 to −0.29) for IVR-CBT and −0.84 (95% CI, −1.29 to −0.26) for in-person CBT (Table 2). IVR-CBT was noninferior to in-person CBT in posttreatment NRS: mean difference between groups: 0.07; 95% CI, −0.67 to 0.80, with an upper limit (0.80) below the noninferiority margin of 1 (Figure 3). Noninferiority was sustained at subsequent time points. Noninferiority was also demonstrated in the per-protocol population (eTable 3 in Supplement 2).

Secondary Outcomes

There were no significant differences between IVR-CBT and in-person CBT in any secondary outcomes at any of the follow-up assessments (Table 2). Participants in both conditions had significant and similar improvements relative to baseline in physical functioning, sleep quality, and SF-36 Physical Component score (PCS) at posttreatment.

Responder Analysis

Posttreatment, 33% of participants (14 of 42) receiving in-person CBT reported clinically meaningful improvement in

Table 1. Baseline Characteristics of Participants by Treatment Group

Characteristic	All (n = 125)	In-Person (n = 63)	IVR (n = 62)
Age, mean (SD), y	57.9 (11.6)	56.7 (11.5)	59.2 (11.8)
Female, No. (%)	28 (22.4)	15 (23.8)	13 (21.0)
Race/ethnicity, No. (%)			
Black, non-Hispanic	32 (25.8)	18 (29.0)	14 (22.6)
Hispanic	9 (7.3)	6 (9.7)	3 (4.8)
White, non-Hispanic	80 (64.5)	37 (59.7)	43 (69.4)
Other	3 (2.4)	1 (1.6)	2 (3.2)
Education, mean (SD), y	13.9 (2.1)	14.1 (2.2)	13.7 (2.0)
Employment, No. (%)			
Full-time	25 (20.0)	16 (25.4)	9 (14.5)
Part-time	17 (13.6)	10 (15.9)	7 (11.3)
Unemployed	19 (15.2)	6 (9.5)	13 (21.0)
Retired	36 (28.8)	16 (25.4)	20 (32.3)
Student	5 (4.0)	2 (3.2)	3 (4.8)
Disabled	23 (18.4)	13 (20.6)	10 (16.1)
Relationship status, No. (%)			
Single	26 (19.3)	14 (22.2)	12 (19.4)
Married	56 (44.8)	31 (49.2)	25 (40.3)
Significant other (if >10 mo)	1 (0.8)	0 (0.0)	1 (1.6)
Divorced/separated	35 (28.0)	16 (25.4)	19 (30.6)
Widowed	7 (5.6)	2 (3.17)	5 (8.06)
Distance to VA, miles, No. (%)			
<10	26 (20.8)	13 (20.6)	13 (21.0)
10-25	44 (35.2)	23 (36.5)	21 (33.9)
>25	55 (44.0)	27 (42.9)	28 (45.2)
History of substance abuse, No. (% yes)	32 (25.6)	14 (22.2)	18 (29.0)
Pain characteristics			
Back pain intensity, mean (SD)[a]	6.46 (1.6)	6.67 (1.7)	6.26 (1.5)
Back pain duration, median (IQR), y	11 (5.0-25.0)	10 (5.0-22.5)	12 (6.0-28.8)
Back pain cause, No. (%)			
Nonspecific	69 (55.2)	34 (54.0)	35 (56.5)
Radiculopathy or spinal stenosis	54 (43.2)	28 (44.4)	26 (41.9)
Other cause	2 (1.6)	1 (1.6)	1 (1.6)
Pain sites, mean (SD), No	2.92 (1.9)	2.86 (1.7)	2.98 (2.2)
Site, No. (% yes)			
Leg	74 (59.2)	37 (58.7)	37 (59.7)
Foot	42 (33.6)	25 (39.7)	17 (27.4)
Arm	32 (25.6)	14 (22.2)	18 (29.0)
Shoulder	57 (45.6)	29 (46.0)	28 (45.2)
Neck	55 (44.0)	25 (39.7)	30 (48.4)
Primary and secondary outcome scores, mean (SD)			
Pain NRS average[b,c]	5.58 (1.6)	5.71 (1.7)	5.45 (1.5)
RMDQ[d]	13.0 (4.8)	14.0 (4.7)	12.1 (4.7)
BDI II[e]	10.3 (7.6)	10.2 (7.9)	10.4 (7.5)
PSQI global[f]	10.3 (4.4)	11.4 (4.1)	9.26 (4.4)
WHYMPI interference subscale[g]	3.03 (1.3)	3.32 (1.3)	2.74 (1.2)
SF-36V[h]			
Physical	35.2 (7.4)	33.5 (7.8)	36.9 (6.7)
Mental	49.0 (8.6)	49.0 (8.3)	48.9 (9.0)
Other outcomes			
Opioid prescription at baseline, No. (% yes)	15 (12.0)	9 (14.3)	6 (9.7)

Abbreviations: BDI II, Beck Depression Inventory II; IQR, interquartile range; IVR, interactive voice response; NRS, numeric rating scale; PCS, Pain Catastrophizing scale; PSQI, Pittsburg Sleep Quality Index; RMDQ, Roland Morris Disability Questionnaire; SF-36V, 36-Item Short Form Health Survey for Veterans; VA, Veterans Administration; WHYMPI, West-Haven Yale Multidimensional Pain Inventory.

[a] Denotes rating of back pain intensity only over the past week.

[b] Primary outcome.

[c] Average pain NRS score (over past week) range is 0 to 10. Higher scores indicate more pain.

[d] RMDQ score range is 0 to 24. Higher scores indicate worse function.

[e] Beck Depression Inventory II score range is 0-63. Higher scores indicate more depression symptoms.

[f] PSQI global score range is 0 to 21. Higher scores indicate more sleep disturbance.

[g] WHYMPI interference subscale score range is 0 to 6. Higher scores indicate more inference.

[h] SF-36V physical and mental component scores range from 0 to 100. Lower scores indicate poorer health status and higher scores indicate higher function.

Figure 2. Number of Treatment Weeks by Condition
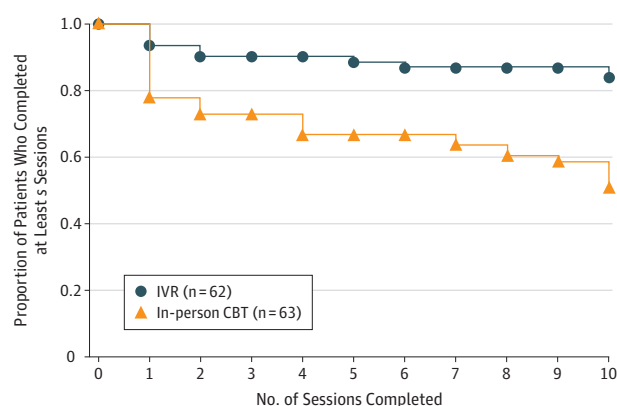
CBT indicates cognitive behavioral therapy; IVR, interactive voice response.

pain intensity of at least 30%, compared with 19% (10 of 52) receiving IVR-CBT (*P* = .19). For RMDQ, 35% of participants (15 of 43) in in-person CBT and 45% (23 of 51) in IVR-CBT reported a reduction of at least 30% (*P* = .43).

Missing Data Sensitivity Analyses

Participants with missing data on the primary outcome posttreatment were more likely to be black (45% [14 of 32] vs 19% [18 of 93]; *P* = .01), had lower average SLUMS dementia screen score (26.9 vs 25.9; *P* = .04), and shorter pain duration (11.3 vs 17.6 years; *P* = .01) than their counterparts. Conclusions remained substantively the same after including these variables as covariates in the models and after multiply imputing missing data (eTable 4 in Supplement 2).

Adverse Events

Forty-six participants experienced 92 related and unrelated adverse events (AEs) (IVR-CBT, 40; in-person CBT, 52). Most related AEs were increased pain from exercise. The number of AEs was not significantly different by treatment group (*P* = .44). Two serious AEs were reported, but judged by the institutional review board to be unrelated to study participation.

Discussion

In this noninferiority randomized trial, IVR-CBT was noninferior to in-person CBT for reduction in pain intensity at all assessment points. Participants in both groups demonstrated statistically significant reductions in average pain intensity at 3 and 6 months postbaseline, but not after 9 months. Examination of secondary outcomes revealed no evidence of a statistically significant difference between groups. Participants in both groups demonstrated significant improvements from pretreatment to posttreatment in physical functioning, physical quality of life, and sleep quality. A greater proportion of in-person CBT participants obtained a clinically meaningful improvement in pain intensity, whereas a larger proportion of participants in IVR-CBT obtained a clinically meaningful

improvement in physical functioning, although neither comparison reached statistical significance. While neither treatment was associated with significant improvements in mood or mental health–related quality of life, mean baseline scores on these measures were near the normal range. Overall, both treatments were associated with a statistically and clinically meaningful change for participants with no meaningful differences between treatments.

Although we found significant improvements in pain intensity and other pain-relevant variables, the effects were modest. Our findings are similar to those of prior studies with veterans, which found significant but modest improvements in response to psychological[23] and collaborative care interventions.[37] Like these prior trials, participants in our study had multiple pain sites, low employment rates, and one-quarter had a history of substance abuse. Both treatments were associated with improved outcomes despite participants' burden of medical and psychiatric illness.

Consistent with our hypothesis that IVR-CBT would be less burdensome than in-person CBT, participants receiving IVR-CBT completed more treatment weeks than those receiving in-person CBT, were significantly less likely to withdraw prior to treatment, and were more likely to obtain a minimal treatment dose. Our findings speak to the appeal, feasibility, and acceptability of IVR-CBT as an alternative to face-to-face therapy and confirm a primary justification for providing treatment using technology; it promotes convenient, patient-centered treatment.

The benefit-to-burden ratio for evaluating any treatment modality depends on each patient's circumstances and preferences. Particularly, among those for whom in-person treatment is not practical, IVR-CBT represents a viable, efficacious, and well-tolerated option. In the presence of barriers to nonpharmacologic options, prescription medications including opioids become the default for many patients. Treatment impact may be captured most accurately by evaluating efficacy and reach.[38,39] Given that IVR-CBT can be offered more widely and at lower burden to patients, it shows promise as a scalable and population-based treatment.

Limitations

We acknowledge several limitations to the study. Because a third placebo arm was not included, we cannot definitively show that either treatment is noninferior to placebo. The trial also did not meet initial enrollment targets; however, we still had adequate power to reject the null hypothesis that the difference in pain intensity between the treatments was less than 1 point. We changed the original eligibility criteria, removing the medical clearance requirement for walking because primary care physicians believed it was unnecessary given the safety of walking. In addition, we did not conduct urine toxicology screens to identify substance misuse because this was perceived as a burden to participants. We do not believe these changes increased the risk of the trial or biased the results and likely increased the real-world applicability of the trial. The participants were all US military veterans, most were male, and all were diagnosed as having painful back conditions. Similar findings in patients with different demographic characteristics and conditions should not be assumed. Participants were slightly older and had longer pain duration than participants

Table 2. Primary and Secondary Outcomes Change From Baseline and Between-Group Differences[a]

Follow-up, mo	No.	In-Person	No.	IVR	Difference IVR-CBT vs In-Person CBT, Mean (95% CI)	P Value for Difference Between Groups
				Change From Baseline, Mean (95% CI)		
Average Pain Intensity NRS (past week) (n = 104)						
3	42	−0.84 (−1.39 to −0.29)	52	−0.77 (−1.29 to −0.26)	0.07 (−0.67 to 0.80)	NA
6	45	−1.00 (−1.52 to −0.48)	49	−1.23 (−1.73 to −0.72)	−0.23 (−0.94 to 0.49)	NA
9	45	−0.44 (−1.01 to 0.14)	49	−0.51 (−1.06 to 0.04)	−0.08 (−0.86 to 0.71)	NA
WHYMPI Total (n = 104)						
3	43	−0.04 (−0.35 to 0.26)	51	−0.37 (−0.66 to −0.08)	−0.33 (−0.74 to 0.09)	.12
6	46	−0.02 (−0.40 to 0.35)	48	−0.37 (−0.73 to 0.00)	−0.34 (−0.87 to 0.18)	.20
9	45	−0.09 (−0.48 to 0.30)	48	−0.12 (−0.50 to 0.26)	−0.02 (−0.57 to 0.52)	.93
RMDQ Total (n = 104)						
3	43	−2.42 (−3.74 to −1.11)	51	−2.92 (−4.16 to −1.69)	−0.50 (−2.29 to 1.29)	.58
6	46	−1.86 (−3.25 to −0.46)	48	−3.38 (−4.75 to −2.02)	−1.53 (−3.46 to 0.41)	.12
9	45	−2.02 (−3.32 to −0.71)	48	−2.63 (−3.90 to −1.35)	−0.61 (−2.42 to 1.20)	.51
PSQI Total (n = 97)						
3	38	−1.09 (−2.06 to −0.11)	47	−2.03 (−2.94 to −1.12)	−0.94 (−2.28 to 0.40)	.17
6	42	−1.21 (−2.12 to −0.29)	45	−1.33 (−2.21 to −0.44)	−0.12 (−1.39 to 1.16)	.86
9	42	−1.25 (−2.38 to −0.11)	44	−0.87 (−1.99 to 0.24)	0.37 (−1.22 to 1.97)	.64
BDI-II Total (n = 105)						
3	43	−1.25 (−3.14 to 0.65)	52	−1.00 (−2.79 to 0.78)	0.24 (−2.32 to 2.80)	.85
6	46	0.78 (−1.40 to 2.96)	49	−0.49 (−2.61 to 1.64)	−1.27 (−4.27 to 1.73)	.40
9	45	1.62 (−0.97 to 4.21)	49	0.89 (−1.62 to 3.40)	−0.74 (−4.30 to 2.83)	.68
SF-36V PCS (n = 100)						
3	40	1.91 (0.01 to 3.81)	47	2.20 (0.43 to 3.96)	0.29 (−2.30 to 2.87)	.83
6	42	0.90 (−1.15 to 2.95)	45	2.05 (0.09 to 4.02)	1.15 (−1.68 to 3.98)	.42
9	41	2.09 (0.03 to 4.14)	45	1.50 (−0.44 to 3.45)	−0.58 (−3.40 to 2.24)	.68
SF-36V MCS (n = 100)						
3	40	0.42 (−1.74 to 2.59)	47	1.67 (−0.36 to 3.71)	1.25 (−1.66 to 4.16)	.40
6	42	1.01 (−1.33 to 3.35)	45	1.39 (−0.86 to 3.64)	0.38 (−2.82 to 3.58)	.81
9	41	−1.47 (−3.96 to 1.02)	45	0.96 (−1.41 to 3.33)	2.43 (−0.96 to 5.82)	.16

Abbreviations: BDI II, Beck Depression Inventory II; IVR, interactive voice response; NA, not applicable; NRS, numeric rating scale; PCS, Pain Catastrophizing scale; PSQI, Pittsburg Sleep Quality Index; RMDQ, Roland Morris Disability Questionnaire; SF-36V, 36-Item Short-Form Health Survey for Veterans; WHYMPI, West Haven-Yale Multidimensional Pain Inventory.

[a] The means are least-squares mean estimates from mixed models adjusting for treatment, time, time-by-treatment interactions, baseline outcome value and stratification variables.

enrolled in trials used to establish the efficacy of CBT.[7] We specified an a priori noninferiority margin for the primary outcome only and could not determine if IVR-CBT is noninferior in other trial outcomes. Future larger trials across multiple VA sites or including non-VA sites are needed to validate and generalize the results from this trial. Finally, there was a relatively large amount of missing data. However, the sensitivity analyses conducted to assess the influence of missing data supported the main analysis conclusions.

Conclusions

Despite these limitations, this trial suggests that IVR-CBT may address national recommendations to promote pain self-management and reduce barriers to care by providing a scalable, low-burden alternative to standard CBT. Given the large number of Americans living with chronic pain, IVR-CBT and similar approaches are promising strategies for improving the availability of nonpharmacologic care and health outcomes, despite constraints on health system resources.

Figure 3. Possible Scenarios and Observed Results of the Noninferiority Test
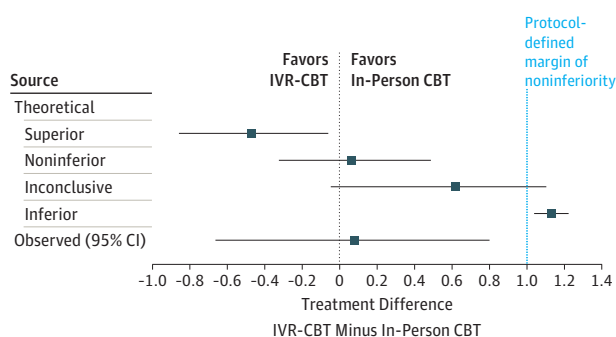
Error bars represent 2-sided 95% CIs. A CI that lies entirely to the left of zero indicates the new treatment (CBT, cognitive behavioral therapy; IVR, interactive voice response [IVR-CBT]) is superior. A CI that lies to the left of the noninferiority margin of 1 indicates that the new treatment is noninferior. A CI that includes the noninferiority margin indicates that the result regarding noninferiority is inconclusive. A CI that is entirely above the noninferiority margin indicates the new treatment is inferior. Figure and explanation adapted from Piaggio et al.[34]

REFERENCES

1. Institute of Medicine. *Relieving Pain in America: A Blueprint For Transforming Prevention, Care, Education, and Research*. Washington, DC: National Academies Press; 2011.

2. Daubresse M, Chang HY, Yu Y, et al. Ambulatory diagnosis and treatment of nonmalignant pain in the United States, 2000-2010. *Med Care*. 2013;51 (10):870-878.

3. Chang HY, Daubresse M, Kruszewski SP, Alexander GC. Prevalence and treatment of pain in EDs in the United States, 2000 to 2010. *Am J Emerg Med*. 2014;32(5):421-431.

4. Martell BA, O'Connor PG, Kerns RD, et al. Systematic review: opioid treatment for chronic back pain: prevalence, efficacy, and association with addiction. *Ann Intern Med*. 2007;146(2):116-127.

5. Department of Health and Human Services. National Pain Strategy: a comprehensive population health-level strategy for pain. 2016. https://iprcc .nih.gov/docs/HHSNational_Pain_Strategy.pdf. Accessed August 1, 2016.

6. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain—United States, 2016. *MMWR Recomm Rep*. 2016;65(1):1-49.

7. Hoffman BM, Papas RK, Chatkoff DK, Kerns RD. Meta-analysis of psychological interventions for chronic low back pain. *Health Psychol*. 2007;26(1): 1-9.

8. Dixon KE, Keefe FJ, Scipio CD, Perri LM, Abernethy AP. Psychological interventions for

arthritis pain management in adults: a meta-analysis. *Health Psychol*. 2007;26(3):241-250.

9. Glombiewski JA, Sawyer AT, Gutermann J, Koenig K, Rief W, Hofmann SG. Psychological treatments for fibromyalgia: a meta-analysis. *Pain*. 2010;151(2):280-295.

10. Cherkin DC, Sherman KJ, Balderson BH, et al. Effect of mindfulness-based stress reduction vs cognitive behavioral therapy or usual care on back pain and functional limitations in adults with chronic low back pain: a randomized clinical trial. *JAMA*. 2016;315(12):1240-1249.

11. Kazdin AE, Blase SL. Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspect Psychol Sci*. 2011;6(1):21-37.

12. Piette JD, Rosland AM, Marinec NS, Striplin D, Bernstein SJ, Silveira MJ. Engagement with automated patient monitoring and self-management support calls: experience with a thousand chronically ill patients. *Med Care*. 2013;51 (3):216-223.

13. Piette JD. Interactive behavior change technology to support diabetes self-management: where do we stand? *Diabetes Care*. 2007;30(10): 2425-2432.

14. Naylor MR, Keefe FJ, Brigidi B, Naud S, Helzer JE. Therapeutic interactive voice response for chronic pain reduction and relapse prevention. *Pain*. 2008;134(3):335-345.

15. Naylor MR, Naud S, Keefe FJ, Helzer JE. Therapeutic Interactive Voice Response (TIVR) to reduce analgesic medication use for chronic pain management. *J Pain*. 2010;11(12):1410-1419.

16. Heapy A, Sellinger J, Higgins D, Chatkoff D, Bennett TC, Kerns RD. Using interactive voice response to measure pain and quality of life. *Pain Med*. 2007;8(suppl 3):S145-S154.

17. Naylor MR, Helzer JE, Naud S, Keefe FJ. Automated telephone as an adjunct for the treatment of chronic pain: a pilot study. *J Pain*. 2002;3(6):429-438.

18. Heapy AA, Higgins DM, LaChappelle KM, et al. Cooperative pain education and self-management (COPES): study design and protocol of a randomized non-inferiority trial of an interactive voice response-based self-management intervention for chronic low back pain. *BMC Musculoskelet Disord*. 2016;17:85.

19. Arnau RC, Meagher MW, Norris MP, Bramson R. Psychometric evaluation of the Beck Depression Inventory-II with primary care medical patients. *Health Psychol*. 2001;20(2):112-119.

20. Tariq SH, Tumosa N, Chibnall JT, Perry MH III, Morley JE. Comparison of the Saint Louis University mental status examination and the mini-mental state examination for detecting dementia and mild neurocognitive disorder: a pilot study. *Am J Geriatr Psychiatry*. 2006;14(11):900-910.

21. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for *DSM-IV* and *ICD-10*. *J Clin Psychiatry*. 1998;59(suppl 20):22-33.

22. Chou R, Qaseem A, Snow V, et al; Clinical Efficacy Assessment Subcommittee of the American College of Physicians; American College

of Physicians; American Pain Society Low Back Pain Guidelines Panel. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med*. 2007;147(7):478-491.

23. Kerns RD, Burns JW, Shulman M, et al. Can we improve cognitive-behavioral therapy for chronic back pain treatment engagement and adherence? a controlled trial of tailored versus standard therapy. *Health Psychol*. 2014;33(9):938-947.

24. Dworkin RH, Turk DC, Farrar JT, et al; IMMPACT. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005;113 (1-2):9-19.

25. Heapy A, Dziura J, Buta E, Goulet J, Kulas JF, Kerns RD. Using multiple daily pain ratings to improve reliability and assay sensitivity: how many is enough? *J Pain*. 2014;15(12):1360-1365.

26. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001;94(2):149-158.

27. Kerns RD, Turk DC, Rudy TE. The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). *Pain*. 1985;23(4):345-356.

28. Roland M, Morris R. A study of the natural history of back pain, part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)*. 1983;8(2):141-144.

29. Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: 25 Years of Evaluation. *Clin Psychol Rev*. 1988;8(1):77-100.

30. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193-213.

31. Kazis LE. The veterans SF-36 health status questionnaire: development and application in the Veterans Health Administration. *Monitor*. 2000;5 (1):1-14.

32. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine (Phila Pa 1976)*. 2000;25(24): 3115-3124.

33. Carroll KM, Nich C, Sifry RL, et al. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug Alcohol Depend*. 2000;57(3):225-238.

34. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA*. 2012;308(24):2594-2604.

35. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain*. 2008;9(2):105-121.

36. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York, NY: John Wiley & Sons; 1987.

37. Dobscha SK, Corson K, Perrin NA, et al. Collaborative care for chronic pain in primary care: a cluster randomized trial. *JAMA*. 2009;301(12): 1242-1252.

38. Glasgow RE, McKay HG, Piette JD, Reynolds KD. The RE-AIM framework for evaluating interventions: what can it tell us about approaches to chronic illness management? *Patient Educ Couns*. 2001;44(2):119-127.

39. Glasgow RE, Bull SS, Piette JD, Steiner JF. Interactive behavior change technology: a partial solution to the competing demands of primary care. *Am J Prev Med*. 2004;27(2)(suppl):80-87.