Self-Administered Psychotherapy for Depression Using a Telephone-Accessed Computer System Plus Booklets: An Open U.S.-U.K. Study

Deborah J. Osgood-Hynes, Psy.D.; John H. Greist, M.D.; Isaac M. Marks, M.D., F.R.C.Psych.; Lee Baer, Ph.D.; Susan W. Heneman; Keith W. Wenzel; Peter A. Manzo, M.S.W.; J. Richard Parkin, M.B.Ch.B., M.R.C.Psych.; Christopher J. Spierings; Susan L. Dottl, Ph.D.; and Helene M. Vitse, B.S.N.

Objective: To evaluate the efficacy and acceptability of a self-help program for mild-to-moderate depression that combined treatment booklets and telephone calls to a computer-aided Interactive Voice Response (IVR) system.

Method: In an open trial, 41 patients from Boston, Massachusetts; Madison, Wisconsin; and London, England, used COPE, a 12-week self-help system for depression. COPE consisted of an introductory videotape and 9 booklets accompanied by 11 telephone calls to an IVR system that made self-help recommendations to patients based on information they entered.

Results: All 41 patients successfully completed the self-assessment in the booklets and telephone calls; 28 (68%) also completed the 12-week self-help program. Hamilton Rating Scale for Depression (HAM-D) and Work and Social Adjustment scores improved significantly (41% and 42% mean reduction in the intent-to-treat sample, respectively, $p < .001$). Eighteen (64%) of the 28 completers were considered responders on the basis of ≥ 50% reduction in their HAM-D scores. There was a higher percentage of completers in the pooled U.S. sites (82% vs. 43%), and U.S. completers improved more than those in the United Kingdom (73% vs. 43% were responders). Most (68%) of the calls were made outside usual office hours, Monday–Friday, 9:00 a.m. to 5:00 p.m. Expectation of effectiveness and time spent making COPE calls (more treatment modules) correlated positively with improvement over 12 weeks. Mean call length for completers was 14 minutes.

Conclusion: A self-help system comprised of a computer-aided telephone system and a series of booklets was used successfully by people with mild-to-moderate depression. These preliminary results are encouraging for people who cannot otherwise access ongoing, in-person therapy.

(J Clin Psychiatry 1998;59:358–365)

Received Jan. 16, 1998; accepted April 29, 1998. *From the Massachusetts General Hospital and Harvard Medical School, Boston (Drs. Osgood-Hynes and Baer and Mr. Manzo), Healthcare Technology Systems, LLC, Madison, Wis. (Dr. Greist, Ms. Heneman, and Mr. Wenzel), the Dean Foundation for Health, Research, and Education, Inc., Middleton, Wis. (Dr. Dottl, Mr. Spierings, and Ms. Vitse), and the Institute of Psychiatry, Bethlem-Maudsley Hospital, London, England (Drs. Marks and Parkin).*

Supported in part by a grant from Pfizer Pharmaceuticals, Inc. The authors gratefully acknowledge the contributions of Terese M. Bailey; Curt Hefty; Henry J. Henk, M.S.; M. Joyce Hirsch, M.A.; Julia M. Mantle, R.N., M.S.; William C. Martin, M.S.; Scott K. Wardle; and Kim M. Ziegler.

COPE™ and Parameter Driven Interview Driver™ are trademarks of Healthcare Technology Systems.

Reprint requests to: Deborah Osgood-Hynes, Psy.D., Massachusetts General Hospital, Charlestown Navy Yard, 13th St., Building 149, 9th Floor, OCD Unit, Charlestown, MA 02129 (e-mail: os-hynes@psych.mgh.harvard.edu).

Depression is a common and recurrent disorder that has a lifetime prevalence of 10% to 25%.[1] Untreated major depression can last 6 to 24 months[1] or even longer for some patients. After the first depressive episode, the average number of recurrent episodes in a lifetime is 4.[1] Untreated depression impacts patients' health functioning substantially, causing physical complaints and poorer health,[2,3] increased health care utilization,[4] frequent disability days from work,[5] poorer intimate relationships, and reduced satisfaction with social interactions.[6,7]

Depression often improves within 12 weeks of starting treatment with antidepressant medication, psychotherapy, or their combination. Because of side effects, some patients discontinue medication prematurely and then relapse.[1] Cognitive-behavior therapy, behavior therapy, interpersonal therapy, and problem-solving techniques are often effective for mild-to-moderate depression.[8–15] Gains frequently persist after formal therapy ends, perhaps because patients use skills learned in therapy if and when depression returns.[16,17]

Despite the value of these particular psychotherapies and evidence that appropriate therapist training improves patient outcome,[18,19] demand for these therapies exceeds

availability. In addition, many depressed people cannot afford therapy. The use of Interactive Voice Response (IVR) telecommunications technology may improve depressed people's access to effective psychological treatment at reasonable cost and with less chance of stigma. IVR involves a centralized computer that presents information and questions to the user by playing prerecorded voice files over the telephone. Patients respond to questions, make choices from menus, and request information by pressing numbers and symbols on the keypad of a Touch-Tone telephone. IVR systems are cost-effective and very convenient. Because the system is run by a computer, patients can call from any Touch-Tone telephone, day and night, 7 days a week (except for short periods of time when regular system backup is performed).

This study evaluated the effectiveness of COPE, a 12-week self-help psychotherapy program for people with mild-to-moderate depression that was delivered by booklets and toll-free telephone calls to a centralized computer-aided therapy system. COPE applied previous research that found computer-instructed assessment to be valid and reliable[20–23] and computer-aided therapy for depression, obsessive-compulsive disorder, and phobias to be effective and well-accepted by users (references 24–28 and Bachofen M, Nakagawa A, Marks IM, et al., manuscript submitted; and Nakagawa A, Marks IM, Park JM, et al., manuscript submitted). This open trial assessed the feasibility, efficacy, and patient acceptance of using booklets plus IVR to deliver self-help guidance for depression.

METHOD

Subjects

Patients were 21 to 75 years old and met DSM-IV[29] criteria for major depression and/or dysthymia, as determined by the clinician-administered Structured Clinical Interview for the DSM-IV (SCID).[30] If another Axis I disorder was present, then depression onset must have preceded it by at least 1 month. Patients with mild-to-moderate depression, defined as a 17-item Hamilton Rating Scale for Depression (HAM-D)[31] score of 12 to 20 were included. Patients were excluded if they had a current or lifetime psychotic disorder, a personality disorder likely to interfere with study participation, or a substance abuse disorder (in the past 6 months). They were also excluded if they were a serious suicide risk or were currently undergoing cognitive-behavior therapy.

Forty-one patients were recruited in Boston, Massachusetts (N = 12); Madison, Wisconsin (N = 15); and London, England (N = 14) through referrals from mental health and primary care professionals and through newspaper advertisements. Twenty-nine patients (71%) were women. The patients' mean ± SD age was 42 ± 13 years. Eighteen (44%) were married or cohabiting, 14 (34%) had never been married, 8 (20%) were divorced, and 1 (2%) was widowed.

Major depression was the most common diagnosis (single or recurrent; N = 25, 61%), followed by dysthymia (N = 11, 27%) and double depression (i.e., both major depression and dysthymia; N = 5, 12%). The mean time since onset of the first episode of depression was 5 years (range, < 1 to 22 years). More patients were diagnosed with dysthymia in London (N = 7) than in Boston (N = 3) or Madison (N = 1) ($\chi^2 = 7.83$, df = 2, $p < .02$), and the London patients were significantly younger (35.4 ± 7.3 years) than the U.S. patients (45.2 ± 14.5 years; $t = 2.85$, df = 39, $p < .01$), but the patients at the 3 sites did not differ significantly on other demographic variables ($p > .05$).

Eight patients (20%) were on antidepressant medications, all on a stable dose for at least 3 months prior to study entry which did not change during the study. Seventeen patients (42%) had taken antidepressant medications in the past, and 18 (44%) had previously had psychotherapy for depression. About a third (N = 13, 32%) said their depression could be due, in part, to the death of someone they knew.

Study Design

The study was a 12-week open trial. The primary outcomes were changes in depression and in work and social function from baseline to end of treatment at week 12.

Measures

All measures were self-rated. Four scales were adapted for administration by the IVR telephone system. The primary measure was a computer-administered 6-item version of the HAM-D[32] based on the 17-item HAM-D.[31,33] For ease of comparison with results from other studies, results on the 6-item HAM-D were converted to a 17-item HAM-D score.[34] The possible HAM-D score range was 0–52, with higher scores indicating more depression.

Other computer-administered IVR measures included the Patient Global Impression (PGI) of Improvement scale[35] and the Work and Social Adjustment (WSA) scale.[36] The PGI is a single item with a score range from 1 (very much improved) to 7 (very much worse). The WSA has 5 items, each scored on a 0 (not at all impaired) to 8 (very severely impaired) scale, with a total score range of 0 to 40. The WSA items address functional impairment at work, in home management, in social leisure activities, in private leisure activities, and in close relationships. All computer-administered measures were given at baseline (except PGI) and at weeks 4, 8, and 12. The HAM-D and PGI were also given at weeks 1 and 2. Benefits of IVR assessment include (1) saving the time and costs of human raters (including costs of establishing and maintaining interrater reliability); (2) errors due to data entry are reduced, and scores are immediately available for analyses; (3) patients are often more honest with the computer or prefer it when disclosing information of a sensitive nature; (4) conscious or unconscious rating bias to meet

enrollment criteria is minimized; (5) the potential confounding of the experimental treatment with the therapeutic effects of the evaluation process is overcome; and (6) when scales are administered by computer over the telephone using IVR, patients can be evaluated off-site from any Touch-Tone telephone, thus enabling remote evaluation of treatment response 24 hours a day.[34]

Patients also completed 2 paper-based scales, patient treatment expectations (at baseline and week 12) and satisfaction (at week 12). Expectation contained 3 questions: (1) How logical does this type of treatment seem to you? (2) How successful do you think this program will be (was) in reducing your depression? and (3) How confident would you be in recommending this program to a friend who experiences depression? These questions were rated from 1 (very logical/successful/confident) to 5 (not at all logical/successful/confident). The 24-item patient satisfaction measure assessed satisfaction with treatment and acceptance of the IVR computer-aided telephone system and COPE program.

COPE System

The COPE self-help system included an introductory videotape, 9 booklets, and 11 toll-free calls to the IVR system, which tailored treatment recommendations and gave individualized feedback based on what the patient entered during the calls. The 13-minute videotape described common symptoms of depression, reviewed self-help treatment options, then briefly discussed the components of the COPE self-help system. Procedures for cognitive restructuring, assertive communication, and increasing pleasant activities were adapted for instruction by booklet and telephone (IVR). COPE focused on these 3 core modules—Constructive Thinking, Pleasant Activities, and Assertive Communication—with additional information available on specific topics (e.g., grieving).

During the "Getting Started" call, patients answered a series of triage questions. Based on their responses, the program recommended they start with a particular module; however, patients were free to choose another module if they wished. They were encouraged to work in 1 module at a time and to complete as many as they wanted, in any order. The 9 booklets, described in Table 1, taught skills to combat depression, suggested homework exercises, offered troubleshooting ideas, and instructed patients when and how to make the telephone calls. The calls reinforced the homework suggestions in the booklets and gave auditory practice/role-play opportunities to strengthen new skills. Treatment recommendations and feedback were individualized based on responses given by patients during the calls. Patients worked at their own pace, selected the topics of interest to them, and were able to repeat calls or entire modules if they chose to do so. A personal message system allowed patients to record up to 3 messages during the 12-week trial. They were able to request clinical feedback or clarification on directions for specific problems. Clinicians monitored the messages and recorded responses, which the patients heard at the beginning of their next call to COPE.

Table 1. The 9 COPE Booklets*

Title of Booklet	Contents
Getting Started	Instructions on how to use the booklets and telephone system, how to recognize depression; a review of the assessments given
Down With Gloom	In-depth discussion of depression symptoms
Constructive Thinking	Suggestions to replace negatively biased, dysfunctional thinking with constructive, supportive self-talk
Constructive Thinking Examples	Examples of constructive thinking to supplement the *Constructive Thinking* booklet
Pleasant Activities	Suggestions on how to monitor daily activities, reduce excess napping, increase activities that are pleasant or give a sense of accomplishment, and balance required tasks/chores with pleasant activities
Assertive Communication	Distinction between assertive and nonassertive communication; suggestions on ways to express positive and negative feelings, accept compliments, and stand up for one's rights
Grieving	Methods to cope with loss and bereavement
Progress Chart Booklet	Charts for recording assessment scores and tracking progress while using COPE
Maintaining Gains	Suggestions on consolidating coping strategies and reducing relapse
Patients were also given a table of antidepressant medications that listed generic and brand names, usual starting doses, usual effective daily doses, and the frequency of common side effects	

*Booklets were written at approximately a seventh-grade reading level.

The computer program was based on predetermined call-flow algorithms that gathered data and gave customized feedback to participants. For example, if patients indicated that they were not engaging in many activities that brought them pleasure, the program would suggest they work through the *Pleasant Activities* booklet and calls. The COPE IVR system contained over 700 text segments, recorded in the first author's voice and presented by the Parameter Driven Interview Driver™ (Healthcare Technology Systems, LLC, Madison, Wis.). Patients' responses in both present and previous calls determined which segments were heard. Patients responded to questions and made menu choices by pressing numbers and symbols (*, #) on the telephone keypad. COPE was available toll-free, from any Touch-Tone telephone, 24 hours a day, 7 days a week (except during routine maintenance). Information collected from patients by the IVR system was stored in a central computer database and used for immediate individualized feedback and treatment recommendations during the calls and for data analyses. Confidentiality was maintained by using assigned 5-digit ID numbers and self-selected, 4-digit passwords.

Table 2. HAM-D and PGI Scores for Intent-to-Treat (ITT) and Completers*

		HAM-D											PGI			
		Baseline		Week 12		Decrease					Responders		Week 12		Responders	
Group	N	Mean	SD	Mean	SD	Mean	SD	t[a]	df	p	N	(%)	Mean	SD	N	(%)
All patients																
ITT	41	18.9	6.0	11.1	8.2	7.8	7.2	6.95	40	< .001	20	(49%)	2.5	1.3	19	(46%)
Completers	28	18.3	4.6	8.8	7.6	9.6	7.0	7.21	27	< .001	18	(64%)	2.1	1.2	18	(64%)
US patients only																
ITT	27	17.7	6.2	8.3	7.7	9.4	7.2	6.78	26	< .001	18	(67%)	2.2	1.3	16	(59%)
Completers	22	18.0	4.9	7.7	7.5	10.3	6.6	7.37	21	< .001	16	(73%)	2.0	1.3	15	(68%)
UK patients only																
ITT	14	21.2	5.1	16.5	6.3	4.7	6.3	2.81	13	.015	2	(14%)	3.2	0.9	3	(21%)
Completers	6	19.5	3.3	12.7	7.1	6.8	8.6	1.95	5	NS	2	(33%)	2.5	0.6	3	(50%)

*Abbreviations: HAM-D = Hamilton Rating Scale for Depression, PGI = Patient Global Impression of Improvement. HAM-D responders were those who had at least a 50% decrease in their HAM-D scores from baseline to week 12. PGI responders had a score of 1 (very much improved) or 2 (much improved) by week 12.
[a]Two-tailed dependent sample t test.

Although COPE was designed primarily for patients who were not severely depressed and/or actively suicidal, it asked suicide-assessment questions when patients indicated suicidal ideation or suicide plans. COPE urged them to contact their doctor immediately and would not permit them to proceed until they reported that they and their doctor had decided they would be safe.

Procedures

The study protocol was approved by each site's institutional review board. Eligible patients visited the hospital/clinic at their site at baseline to enroll. After giving written informed consent, they completed a desktop computer version of the HAM-D and were evaluated by a clinician to determine suitability for the study. The SCID[30] was administered to verify diagnoses and to detect possible comorbid conditions. Patients then viewed the videotape, heard an explanation of the COPE materials, and completed the first 2 telephone calls from the office (Welcome Call and Getting Started Call, which included the baseline assessments). At the end of the calls they completed the treatment-expectation questionnaire. Subsequent calls, which lasted 8 to 23 minutes, were made from outside the hospital/clinic. The computer administered interim assessments at weeks 1, 2, 4, and 8 when patients called in to the IVR system. Patients returned to the hospital/clinic at week 12 to complete endpoint questionnaires and to provide feedback on their use of COPE.

Statistical Analysis

Means and standard deviations were calculated for continuous variables, and frequencies and percentages for categorical variables. Group comparisons were made using 2-tailed dependent and independent t tests, analysis of variance, and chi-square. The probability level for all tests was $p < .05$; statistical power was > .80.

An intent-to-treat (ITT) analysis included all study participants who completed the enrollment visit (N = 41). The last data point (endpoint data) for those who did not complete the week-12 office visit was carried forward. "Completers" were study participants who completed the office visit at week 12. "Responders" were defined in 2 ways: those subjects whose HAM-D scores improved ≥ 50% from baseline to week 12 (HAM-D responders) and those whose final PGI-Improvement score was 2 or 1 (indicating much or very much improved, respectively).

RESULTS

All 41 people in the ITT analysis completed the enrollment assessment and the initial IVR telephone assessment which consisted of the HAM-D and WSA scale. Twenty-eight (68%) were completers; 2 (5%) did not use COPE after the enrollment visit. The remaining 11 patients used COPE after the enrollment visit but stopped using COPE at various points in the program, did not attend the week-12 office visit, and were considered noncompleters.

Changes in Outcome Scores

Baseline and week-12 HAM-D and PGI scores are presented in Table 2. Patients' HAM-D scores improved significantly by the end of the study. According to the HAM-D criterion, 15 (37%) participants in the ITT group were responders by week 8, and 20 (49%) by week 12. Response rates were similar across diagnostic categories ($\chi^2 = 0.46$, df = 2, p = .80): 13 (52%) of 25 patients with major depression, 5 (46%) of 11 patients with dysthymia, and 2 (40%) of 5 patients with double depression met the HAM-D response criterion. Nineteen patients (46%) had a final HAM-D score less than 10. Nineteen patients (46%) met the PGI response criterion. There was also a significant improvement in patients' work and social function (Table 3). WSA scores improved significantly in each of the 5 domains (e.g., work, relationships).

Of the 28 completers, 18 improved markedly in their HAM-D scores (see Table 2) with a mean decrease of 9.6 ± 7.0 points. Over two thirds (N = 20, 71%) met HAM-D (N = 18, 64%) and/or PGI (N = 18, 64%) re-

Table 3. Mean ± SD Work and Social Adjustment (WSA) Scores at Baseline and Week 12 (intent-to-treat; N = 41)

Item Number	Item Construct	Baseline Mean	Baseline SD	Week 12 Mean	Week 12 SD	Change From Baseline to Week 12 Mean	Change From Baseline to Week 12 SD	t[a]	df	p
1	Work	3.5	2.1	1.9	2.2	1.7	1.9	5.59	40	< .001
2	Home	3.9	1.6	2.1	2.0	1.8	1.6	7.10	40	< .001
3	Social leisure	4.3	1.9	2.6	2.3	1.7	2.2	5.01	40	< .001
4	Private leisure	3.5	2.0	2.0	1.7	1.5	1.8	5.12	40	< .001
5	Family	3.7	2.1	2.3	1.9	1.4	2.2	4.05	40	< .001
Total		19.0	7.9	10.9	9.1	8.1	8.2	6.27	40	< .001

[a]Two-tailed dependent sample t tests.

sponse criteria. In addition, completers' work and social function improved significantly, with a fall in mean WSA total scores from 18.0 (SD = 7.2) at baseline to 7.8 (SD = 7.0) at week 12 (t = 6.58, df = 27, p < .001).

Significantly more U.S. patients (22 of 27; 82%) than U.K. patients (6 of 14; 43%) completed the study ($\chi^2 = 6.35$, df = 1, p < .02). U.S. patients also improved more (see Table 2): their mean HAM-D scores in the ITT analysis fell 9.4 points, compared to a fall of 4.7 points among all U.K. patients (t = 2.06, df = 39, p = .046). Eighteen (67%) of the 27 U.S. patients were HAM-D responders, and 16 (59%) were PGI responders. In comparison, 2 (14%) of the 14 U.K. patients were HAM-D responders, and 3 (21%) were PGI responders. Among the completers, 16 (73%) of U.S. participants were HAM-D responders and 15 (68%) were PGI responders, compared with 2 (33%) U.K. participants who met HAM-D criterion and 3 (50%) who met PGI criterion. U.K. patients' WSA total scores dropped significantly from baseline (mean ± SD = 22.1 ± 8.3) to endpoint (16.6 ± 9.0; t = 2.48, df = 13, p = .028).

Modules, Number and Length of Calls, and Improvement

Thirty-five (85%) of the 41 participants completed at least 1 of the 3 treatment modules (constructive thinking, pleasant activities, assertive communications). Eighteen patients (44%) completed 2 treatment modules, 8 (20%) completed 3 treatment modules, and 5 (12%) completed the 3 treatment modules and the Maintaining Your Gains module. Four people (10%) repeated a previously completed treatment module. Patients chose to complete the Assertive Communication module most frequently (N = 27, 66%), followed by Constructive Thinking (N = 19, 46%), and Pleasant Activities (N = 13, 32%).

Between baseline and week 12, patients spent a mean of 156 ± 94 minutes making COPE telephone calls. Each ITT patient completed a mean of 13 calls (range, 2 to 49). The mean call length for ITT patients was 12.7 minutes and for completers was 14.4 minutes. U.S. patients spent significantly more time making calls than U.K. patients: they completed a mean of 177 ± 100 minutes, compared to 115 ± 67 minutes by the U.K. patients (t = 2.08, df = 39, p < .05). Of the 391 calls made after the initial office visit, 126 calls (32%) were made on weekdays from 9 a.m. to 5 p.m. (local time). The remaining 265 calls (68%) were made outside usual office hours. Patients were given the option to continue to call past the end of the trial at week 12 if they desired. Of the 41 participants, 36 people completed their calls within week 12 and 5 people made calls after week 12; there were no calls made after the 14th week.

Patients who made the most COPE calls improved the most. Of the 18 people who made 10 or more calls, 13 (72%) were HAM-D responders, compared with only 7 (30%) of the 23 who made fewer than 10 calls (Fisher exact test, p = .01). The number of minutes patients spent on the telephone with the IVR system correlated with change in HAM-D scores (r = .46, df = 40, p = .002) and with PGI scores (r = –.56, df = 37, p < .001).

Fourteen participants (34%) left a total of 19 messages on the personal message system: 10 people left 1 message, 3 left 2 messages, and 1 left 3 messages. Participants from each site used the voicemail option (Boston, N = 6; Madison, N = 5; London, N = 3). Patients used the message system to tell us why they were not calling (e.g., out of town), to ask for help with specific questions (e.g., how to deal with numerous negative thoughts), to clarify instructions in the COPE booklets, to give feedback on how they were doing (both positive and negative), and, in 1 case, to complain about the limits on automated feedback.

Predictors of Success With COPE

The strongest predictor of response using the HAM-D criterion was the patients' answers at baseline to the single question, "How logical does this type of treatment seem to you?" Of the 15 people who said COPE seemed a "very logical" treatment, 12 (80%) were HAM-D responders. Answers to the question "How successful do you think this program will be in reducing your depression?" added no further predictive ability when the answers to the "logic" question were controlled for. A regression model containing logicality and the number of booklets completed significantly predicted HAM-D outcome ($R^2 = .40$; F = 12.3, df = 2,37; p < .001) and final PGI scores ($R^2 = .43$, F = 12.8, df = 2,34; p < .001).

Table 4. Satisfaction Results of the 28 Completers

Satisfaction Item	Response, N (%)				
How comfortable were you using the IVR system?	Very comfortable 19 (68%)	Somewhat comfortable 6 (21%)	Just okay 1 (4%)	Somewhat uncomfortable 1 (4%)	Very uncomfortable 1 (4%)
How well could you describe how you felt?	Very well 12 (43%)	Somewhat well 10 (36%)	Just okay 4 (14%)	Poorly 1 (4%)	Very poorly 1 (4%)
How sensitive were IVR calls to your needs?	Extremely sensitive 3 (11%)	Very sensitive 13 (46%)	Somewhat sensitive 9 (32%)	Insensitive 1 (4%)	Very insensitive 0
How easy was it to answer IVR questions?	Very easy 16 (57%)	Easy 9 (32%)	Neither easy nor hard 1 (4%)	Hard 1 (4%)	Very hard 0
How easy was it to use the COPE booklets?	Very easy 21 (75%)	Easy 6 (21%)	Neither easy nor hard 0	Hard 1 (4%)	Very hard 0
How useful were the COPE booklets?	Very useful 18 (64%)	Somewhat useful 6 (21%)	Just okay 3 (11%)	Not very useful 0	Not useful at all 1 (4%)

Positive outcome was not predicted by education, age, diagnosis (major depression versus dysthymia versus double depression), initial depression severity, which treatment module participants did first, whether the Constructive Thinking, Pleasant Activities, Assertive Communication, and/or Maintaining Your Gains calls were completed, use of a "buddy" while doing COPE, nor response on any individual WSA item (for all partial r values, $p > .05$).

Follow-up telephone contact by site coordinators with participants who were considered noncompleters (did not complete the week-12 office visit) indicated that noncompletion did not necessarily indicate nonuse of the COPE system or ineffectiveness of treatment; 8 of 13 noncompleters used the COPE program beyond the initial enrollment visit. Of the 41 study participants, 5 (18%) of 27 U.S. subjects and 8 (57%) of 14 U.K. subjects were noncompleters. We reached 9 of the 13 noncompleters for follow-up; 6 had used the IVR portion of the COPE program beyond the initial visit and 3 had not. Of these, 4 felt they needed more human interaction, 4 stopped calling when they started feeling better, and 1 increased medication due to worsening depression and was removed from the study.

The Patient Satisfaction Scale was filled out by the 28 completers (Table 4). Overall, patients felt comfortable using the IVR system, found it easy to use, and found the booklets helpful. Twenty-one (75%) of the 28 felt that COPE had improved the quality of their lives.

DISCUSSION

The results of this open trial are encouraging: mildly to moderately depressed participants improved significantly in mood and in work and social function over 12 weeks of treatment with the COPE computer-aided IVR self-help system. It allowed easy access to depression self-treatment methods and was perceived as helpful and easy to use by most patients.

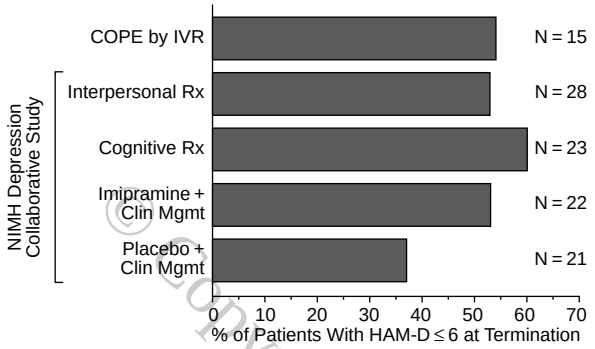
No treatment module stood out as any more effective than another (i.e., Pleasant Activities versus Constructive Thinking versus Assertive Communication). Expectation of effectiveness predicted outcome, and time spent doing COPE calls correlated positively with improvement. Most patients completed COPE within the 12 weeks of the trial. For all 41 participants, there was a 41% reduction in HAM-D scores. This was comparable to the 1997 controlled trial of cognitive therapy and pharmacotherapy[37] that found at week 12 a HAM-D reduction of 41% for the cognitive therapy group and 36% and 31% reductions for 2 antidepressant groups.

Every responder with an endpoint HAM-D score of 0 ($N = 6$) had indicated prior to starting treatment that COPE seemed moderately or very logical to them. If the predictive power of this question is confirmed in future trials of COPE, it may be useful for initial screening of mildly or moderately depressed patients. Identifying patients who may need additional attention and education before starting treatment may improve compliance and outcome. Patient education (e.g., causes of depression, symptoms, treatment options, mechanisms of action, anticipated outcomes, compliance difficulties, and early warning signs of relapse) can enhance treatment adherence.[38,39]

The reasons for the lower completion and response rates in the United Kingdom, compared to the United States, are not obvious. U.K. patients did not have significantly higher levels of depression at baseline and were not significantly different on pretreatment expectations. They did, however, have more impairment in their work due to depression than the U.S. patients, were significantly younger, were more likely to have dysthymia than the U.S. patients, and among noncompleters expressed an increased desire to have more personal interaction. Anecdotally, it appeared that U.K. patients had less familiarity with IVR systems. They also spent one third less time making calls to the IVR. These factors may have contributed to the lower completion and response rates. On the other hand, U.K. patients suffering from OCD improved as much as U.S. patients when treating themselves with BT STEPS™, a self-help IVR-based program for that disorder.[28]

Of the 41 participants, 13 (32%) did not attend the week-12 office visit. This noncompletion rate closely mirrors findings that up to 40% (range reported of 4% to

Figure 1. Comparison of Recovery Rates of Completers in COPE and "Less Severely Depressed Patients" (HAM-D ≥ 14 < 20) in the NIMH Depression Collaborative Study*



*NIMH Depression Collaborative Study data from reference 8.

90%)[17] of patients do not complete face-to-face psychotherapy[40] or antidepressant medication.[41]

One limitation of this open trial is the absence of a standard treatment or placebo group as control. While attention may have accounted for decreased scores and increased functioning for some, it should be noted that there was minimal contact with research coordinators and only 34% of patients (14 of 41) received extra attention through the personal message system. Of these, 10 (71%) of the 14 made just 1 call and received a brief (< 3 minute) prerecorded return message. Initial results are promising, but could be due to a placebo effect or spontaneous improvement; a controlled trial could address this issue. In the absence of a controlled trial, we compared recovery rates for COPE with those in the National Institute of Mental Health (NIMH) Depression Collaborative Research Study[8] (Figure 1). Fifteen (54%) of COPE's 28 patients in the completer analysis met the NIMH response criterion (i.e., posttreatment HAM-D ≤ 6). Drawing further comparisons, according to the intent-to-treat meta-analyses conducted by the Agency for Health Care Policy and Research (AHCPR) in their 1993 Clinical Practice Guidelines for Depression,[19] the anticipated/estimated percentage of adult outpatients who will respond to antidepressant medication treatment (tricyclics, heterocyclics, SSRIs, MAOIs, anxiolytics) is 54% ± 8.5%. The psychotherapy responder estimate is 50% ± 5.3%. These figures are comparable to COPE's intent-to-treat responder rate (50% or greater decrease in HAM-D scores), which was 49% for U.K. and U.S. sites (pooled). With site breakdown, ITT responder rates are 67% and 14%, U.S. and U.K. respectively.

Another limitation is that we assessed only short-term effectiveness. Since this was an early trial of telephone-accessed computer treatment for depression, further evaluation is needed of COPE's ability to alter diagnostic status, maintain improvement, and prevent relapse (particularly among dysthymics) over longer periods of time. While absence of depression disorder diagnosis after treatment is another indication of remission,[42] the correlation between HAM-D scores and diagnosis is high.[43]

COPE customizes self-help recommendations to reduce mild-to-moderate depression. Participants have access to support and to individualized suggestions of ways to reduce their depression nearly 24 hours a day. Over two thirds (68%) of the calls to COPE were made outside regular office hours, highlighting a major benefit of IVR systems—access at the patients' convenience. At debriefing, many patients commented on the convenience of being able to make calls at any time.

Several U.K. patients asked to use COPE on the express condition that their general practitioner *not* be told of their participation. They appeared to appreciate COPE's confidentiality.

IVR technology can be applied effectively to a broad spectrum of mental and physical health problems. It is especially valuable because it can be used at any time, from any location, by any patient with access to a Touch-Tone telephone. In addition, because the system is conducted by a computer program and data are automatically collected, it is possible to manipulate and evaluate the effectiveness of the various components of the therapy systematically. While it is possible to systematically manipulate treatment variables in therapies administered by humans, such manipulations cannot be standardized as well as they can be by using IVR computer administration.

COPE is not intended to replace professional, clinician-administered treatment of depression. It can be used alone as a self-help program or in combination with medication and/or face-to-face psychotherapy.

In an era where emphasis is placed on the delivery of cost-effective treatments, this computer-aided telephone technology, combined with self-help booklets, offers hope of making effective treatments available to many people who would not otherwise have access to them. For those who do not improve significantly after using COPE, the assessments and education COPE offers could facilitate subsequent face-to-face treatment.

REFERENCES

1. Clinical Practice Guideline Number 5: Depression in Primary Care, vol 1. Detection and Diagnosis. Rockville, Md: US Dept Health Human Services, Agency for Health Care Policy and Research; 1993. AHCPR publication 93-0550
2. Wells KB, Golding JM, Burnham MA. Psychiatric disorder and limitations in physical functioning in a sample in Los Angeles general population. Am J Psychiatry 1988;145:712–717
3. Coulehan JL, Schulberg HC, Block MR, et al. Medical comorbidity of major depressive disorder in a primary medical practice. Arch Intern Med 1990;150:2363–2367
4. Regier DA, Hirschfeld MA, Goodwin FK, et al. The NIMH depression awareness, recognition and treatment program: structure, aims and scientific basis. Am J Psychiatry 1988;145:1351–1357
5. Broadhead WE, Blazer DG, George LK, et al. Depression, disability days, and days lost from work in a prospective epidemiologic survey. JAMA 1990;264:2524–2528

6. Fredman L, Weissman MM, Leaf PJ, et al. Social functioning in community residents with depression and other psychiatric disorders: results of the New Haven Epidemiologic Catchment Area study. J Affect Disord 1988;15:103–112
7. Johnson J, Weissman MM, Klerman GL. Service utilization and social morbidity associated with depressive symptoms in the community. JAMA 1992;267:1478–1483
8. Elkin I, Shea MT, Watkins JT, et al. NIMH Treatment of Depression Collaborative Research Program: general effectiveness of treatments. Arch Gen Psychiatry 1989;46:971–982
9. Hersen M, Bellack AS, Himmelhoch JM, et al. Effects of social skills training, amitriptyline, and psychotherapy in unipolar depressed women. Behav Ther 1984;15:21–40
10. Beck AT, Hollon SD, Young JE, et al. Treatment of depression with cognitive therapy and amitriptyline. Arch Gen Psychiatry 1985;42:142–148
11. Shea MT, Elkin I, Imber SD, et al. Course of depressive symptoms over follow-up: findings from the NIMH Treatment of Depression Collaborative Research program. Arch Gen Psychiatry 1992;49:782–787
12. Lewinsohn PM, Antonuccio DA, Steinmetz J, et al. The Coping With Depression Course: A Psychoeducational Intervention for Unipolar Depression. Eugene, Ore: Castilia Press; 1984
13. Weissman MM, Pruscoff BA, DiMasco A, et al. The efficacy of drugs and psychotherapy in the treatment of acute depressive episodes. Am J Psychiatry 1979;136:555–558
14. Mynors-Wallis LM, Davies I, Gray A, et al. A randomized controlled trial comparing problem-solving treatment with amitriptyline and placebo for major depression in primary care. BMJ 1995;310:441–445
15. Mynors-Wallis LM, Gath DH, Loyd-Thomas AR, et al. A randomized controlled trial and cost analysis of problem-solving treatment for emotional disorders given by community nurses in primary care. Br J Psychiatry 1997;170:113–119
16. Simons AD, Murphy GE, Levine JL. Cognitive therapy and pharmacotherapy for depression: sustained improvement over one year. Arch Gen Psychiatry 1986;43:43–48
17. Gallagher DE, Thompson LW. Treatment of major depressive disorder in older adult outpatients with brief psychotherapies. Psychotherapy: Theory, Research, and Practice. 1982;19(4):482–490
18. Shaw BF, Olmstead M. Competency ratings in relation to protocol adherence and clinical outcome. Presented at The Society for Psychotherapy Research; June 1989; Toronto, Ontario, Canada
19. Clinical Practice Guideline Number 5: Depression in Primary Care, vol 2. Treatment of Major Depression. Rockville, Md: US Dept Health Human Services, Agency for Health Care Policy and Research; 1993. AHCPR publication 93-0551
20. Baer L, Jacobs DG, Cukor P, et al. Automated telephone screening survey for depression. JAMA 1995;273:1943–1944
21. Baer L, Brown-Beasley W, Sorce J, et al. Computer-assisted telephone administration of a structured interview for obsessive-compulsive disorder. Am J Psychiatry 1993;150:1737–1738
22. Kobak KA, Reynolds WM, Greist JH. Development and validation of a computer-administered version of the Hamilton Anxiety Scale. Psychological Assessment 1993;5(4):487–492
23. Kobak KA, Reynolds WM, Rosenfeld R, et al. Development and validation of a computer-administered version of the Hamilton Depression Rating Scale. Psychological Assess 1990;2(1):56–63
24. Ghosh A, Marks IM, Carr AC. Therapist contact and outcome of self-exposure treatment for phobias: a controlled study. Br J Psychiatry 1988; 33:112–117
25. Selmi PM, Klein MH, Greist JH, et al. Computer-administered cognitive-behavioral therapy for depression. Am J Psychiatry 1990;147:51–56
26. Greist JH, Baer L, Marks IM, et al. Computer-assisted behavior therapy for obsessive-compulsive disorder. In: New Research Program and Abstracts of the 149th Annual Meeting of the American Psychiatric Association; May 9, 1996; New York, NY. Abstract NR759:277
27. Marks IM, Baer L, Greist JH, et al. Home self-assessment of obsessive-compulsive disorder; use of a manual and a computer-conducted telephone interview: two UK-US studies. Br J Psychiatry 1998;172:406–412
28. Greist JH, Marks IM, Baer L, et al. Home self-exposure therapy of obsessive compulsive disorder using a manual and a computer-conducted telephone interview: a US-UK study. MD Comput. In press
29. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. Washington, DC: American Psychiatric Association; 1994
30. First MB, Spitzer RL, Gibbon M, et al. Structured Clinical Interview for DSM-IV Axis I Disorders, Patient Edition (SCID-I/P, Version 2.0). New York, NY: Biometric Research, New York State Psychiatric Institute; 1995
31. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62
32. Bech P, Allerup P, Gram LF, et al. The Hamilton Depression Scale: evaluation of objectivity using logistic models. Acta Psychiatr Scand 1981;63: 290–299
33. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 1967;6:278–296
34. Mundt JC, Kobak KA, Taylor LVH, et al. Administration of the Hamilton Depression Rating Scale using interactive voice response technology. MD Comput 1997;15:1–9
35. Guy W. ECDEU Assessment Manual for Psychopharmacology. US Dept Health, Education, and Welfare publication (ADM) 76-338. Rockville, Md: National Institute of Mental Health; 1976:218–222
36. Marks IM. Behavioural Psychotherapy. Bristol: John Wright; 1986:46. Now published by I. Marks, Institute of Psychiatry, London SE58AF, England
37. Blackburn IM, Moore RG. Controlled acute and follow-up trial of cognitive therapy and pharmacotherapy in out-patients with recurrent depression. Br J Psychiatry 1997;171:328–334
38. Peet M, Harvey NS. Lithium maintenance, I: a standard education program for patients. Br J Psychiatry 1991;158:197–200
39. Anderson CM, Griffin S, Rossi A, et al. A comparative study of the impact of education vs process groups for families of patients with affective disorders. Fam Process 1986;25(2):185–205
40. Persons JB, Burns DD, Perloff JM. Predictors of dropout and outcome in cognitive therapy for depression in a private practice setting. Cognitive Ther Res 1988;12(6):557–575
41. Katzelnick DJ, Kobak KA, Jefferson JW, et al. Prescribing patterns of antidepressant medications for depression in a HMO. Formulary 1996;31: 374–388
42. Harrison WM, Stewart JW. Pharmacotherapy of dysthymia. Psychiatr Ann 1993;23(11):638–648
43. Kobak KA, Taylor LV, Dottl SL, et al. A computer-administered telephone interview to identify mental disorders. JAMA 1997;278:905–910