Never too old to learn : the effectiveness of the Coping with Depression course for elderly

Haringsma, R.

3

Effectiveness of the Coping With Depression course for older adults provided by the community-based mental health care system in the Netherlands; a randomized controlled field trial

Haringsma, R., Engels, G. I., Cuijpers, P., & Spinhoven, Ph. (2006). Effectiveness of the Coping with Depression course for older adults provided by the community-based mental health care system in the Netherlands; a randomized controlled field trial. *International Psychogeriatrics,18*, 307 - 325
.

Abstract

Background.
The Dutch version for older adults of the Coping With Depression (CWD) course has been implemented in the prevention arm of the community-based mental health care system in the Netherlands. The study group included older adults with sub-clinical depression as well as with a major depressive disorder; all were enrolled into the course by mental health care professionals. The effectiveness (immediate and long-term) of the course for this heterogeneous population was studied in an effectiveness trial.

Method.
Participants were self-referred, responding to media announcements. A total of 119 participants aged 55-85 (69% female), with sub-clinical depression and major depression were randomized to either the CWD course (n = 61) or waiting list (n =58).

Results.
Nine participants dropped out of the course. According to a diagnostic interview based on the DSM-IV, 39% had a major depressive disorder (MDD), 69% had had a previous MDD, and 45% had an anxiety disorder. Older adults in the intervention group showed a significant decrease in depression symptoms. Gains were maintained over 14 months. In the intervention condition, 83% had a pretreatment score ≥16 on the Center for Epidemiologic Studies Depression scale (CES-D), at posttreatment 62% still scored ≥16.

Conclusions.
The course was beneficial for participants with mild or severe depression, and treatment acceptability was high. It should be fitted into a stepped-care protocol that varies intervention intensity according to clinical needs, using the posttreatment level of functioning as an indication for the next step.

.

Introduction

Depression – major depression and subclinical depression - is a common psychiatric disturbance in late-life. Prevalence for major depression in elderly people is around 3%, and 8 – 15 % have subclinical or minor depression (Beekman et al., 1999; Karel and Hinrichsen, 2000). Extensive research has shown the efficacy of psychosocial interventions for late-life depression (Bartels et al., 2002; Engels and Vermey, 1997; McCusker et al., 1998; Scogin and McElreath, 1994). Various outreach programs have been developed to reach the community-living senior citizen with depression who may not seek treatment. In a meta-analysis of the effects of such programs carried out in controlled research settings, a mean effect size (ES) of 0.77 was found, which is comparable with effect sizes found in younger adults. Most effective were programs based on cognitive behavioral therapy (Cuijpers, 1998c). However, the extent to which results obtained in controlled research settings can and should be used to inform the treatment of patients in the community is still unknown (Stirman et al., 2003). There is a need to establish more firmly whether the efficacy of psychosocial interventions, as found in randomized controlled trials in research settings, can be generalized to the usual care setting, with the typical consumers of community mental health services and typical community staff (Street et al., 2000).

The Dutch version for older adults of the Coping With Depression Course (CWD) (Lewinsohn et al., 1984) was implemented in the prevention arm of the community mental health system In the Netherlands in the nineties. The CWD course has proved an efficacious group treatment for unipolar depression and has been adapted for different populations (Cuijpers, 1998a). At present, 60% of the Prevention departments of the Community Mental Health Centers (CMHCs) offer this course regularly to older adults with mild depressive complaints (Voordouw and Kramer, 2001). Because depression is often an episodic disorder, with periods of remission and exacerbation, the CMHCs accept both individuals who are self-described depressed and those who report minimal difficulties. In their study on the criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D), Haringsma et al. (2004) found that participants of the course varied widely in their level of depression from exhibiting only slight symptoms to being severely depressed. The vast majority of the participants had a lifetime major depressive disorder (MDD), and 40% met the criteria for an MDD. This means that for the latter subgroup, the course can be considered treatment with remission as the goal; for the other 60% the prevention of a new major depressive episode is the objective. A decrease in depressive symptomatology is the desired outcome in both groups.

A randomized controlled field trial was carried out to analyze the immediate effectiveness of this program as it is provided by the Dutch mental health care system. The long-term effectiveness was explored in the experimental group only. As the study was embedded in the procedures employed by the CMHCs, the results will give an externally valid picture of the immediate and long-term effects of the program.

Method

Participants and Recruitment Procedures

Ten CMHCs, both urban and rural, and from all the regions in the Netherlands that provide the course, participated in this study with 13 courses. The individuals in the study were older adults taking part in the CWD courses during the years 2000 and 2001. Participants were self-referred, responding to media announcements. Accepted were older adults (minimum age around 55) with current or a history of depressive symptoms. Exclusion criteria were: cognitive impairment, current bipolar disorder, schizophrenia, acute substance disorder, recent bereavement, hearing impairment, and insufficient command of the Dutch language. These are the standard procedures used by the CMHCs.

Not all the individuals enrolled into the course by the CMHCs were accepted for the study. Study participants had to be at a minimum age 55 years and receiving no other form of psychotherapy. Use of psychotropic medication was accepted. All CWD participants were provided with a complete description of the study, and written informed consent was obtained before enrollment into the study. Participation in the study was voluntary; consequently not all the seniors attending the CWD course took part in the study. Those who joined the study and completed the posttreatment assessment were reimbursed for the small contribution ($15 – 40) imposed by most of the CMHCs. The study was approved by the Medical Ethics Committee of the Leiden University Medical Centre.

Study Design

The present study was a randomized controlled field trial. Randomization was by block design to ensure that participants with and without a current MDD according to the criteria of the diagnostic and statistical manual of mental health - 4th edition (DSM-IV; American Psychiatric Association, 1994) were equally divided. Randomization to one of the two conditions took place after the diagnostic interview. Assessments of the effect of the course took place posttreatment; follow-up (FU) was at 2 and 14 months.

In a meta-analysis of the effects of outreach programs offered to depressed elders in the community a large mean *ES* of 0.77 was found (Cuijpers, 1998c). To be able to detect a large between-group difference in effectiveness of at least .75 with a power of 90% and an alpha (α) set at .05 (two-sided) at least 39 subjects in each condition are needed. With an expected drop-out rate of 25% (Cuijpers, 1998c) and including about 104 patients, around 78 subjects will complete the course or waiting list (39 in each condition). In an intention-to-treat analysis with 52 or more subjects in each condition and an α of .05 (two-sided) the power to detect a medium *ES* of at least 0.50 is 71% or higher.

Interventions

Intervention condition (IC)

The intervention consisted of 10 weekly sessions of two hours each in groups of six to 13 participants (not all of whom were participating in the study) following a protocol (the Dutch version of the CWD course for elders, an adaptation from Lewinsohn's CWD course (Cuijpers, 1998b; Lewinsohn et al., 1984). The contents are based on a social learning view of depression; the skills taught are relaxation, increasing pleasant activities, constructive thinking, improving social skills, and maintaining treatment gains. The instructors were two health professionals trained in conducting this course.

Waiting list condition (WLC)

The people in the WLC received the course directly after the intervention had finished. For ethical reasons, the participants in the control condition were not kept waiting until the intervention group had completed the 14 months FU. A reminder from the CMHC about the start of their course was the only attention they received; they did not receive any other forms of psychological treatment.

Treatment Integrity

All sessions were audio taped; two tapes of each course (one for each half of the course) were randomly selected for review of adherence to the therapy manual. One of the two tapes was randomly selected for listening. Adherence was judged on the basis of a checklist covering the topics that should be addressed. The second tape was assessed when (a) deviations from the manual were observed – this happened once, and (b) the quality of the tape was poor – this occurred three times. No deviations from the manual were found in the assessments of the four second tapes and we therefore concluded that all groups in the study adhered to the manual.

Assessment Procedures

All self-report questionnaires used at the various assessments were completed at home. Missing or incorrect items were treated as follows: (a) following submission, participants were contacted by telephone to discuss incorrect answers and missing items, which were then remediated and (b) if this failed, the values of the missing items were imputed, but only if the number of missing items did not exceed 15% of the items of a particular measure.

Pretreatment questionnaires were distributed in the two weeks prior to the start of the course. In this period, diagnoses by the researchers took place at the CMCH with a structured interview based on the DSM-IV (1994). Participants in both conditions received the posttreatment questionnaires in the week following the tenth session in the IC. If they were not returned within two weeks, participants were contacted once by telephone or mail to urge submission.

Measures

Clinical diagnoses

The Dutch translation (Overbeek et al., 1999) of the clinician-rated (CR) version of the Mini International Neuropsychiatric Interview (M.I.N.I.; Sheehan, et al., 1998a) was used. It assesses the most prevalent DSM-IV (1994) axis I disorders: (a) affective disorders - MDD, (current and lifetime), dysthymia (current), and manic disorder; (b) anxiety disorders - panic disorder, agoraphobia, social phobia, obsessive-compulsive disorder, posttraumatic stress disorder, and general anxiety disorder; (c) alcohol and drug dependency; (d) psychotic disorders; (e) eating disorders - anorexia and bulimia; and (f) somatization disorder. Validation of the M.I.N.I.-CR against the Structured Clinical Interview DSM-IIIR - patient version (SCID-P) and the Composite International Diagnostic Interview for ICD-10 (CIDI) showed good to very good kappa values (Sheehan, et al., 1998b). In the present study, the interviews were conducted by trained interviewers at the CMHC where the course was scheduled. Interrater reliability (Kappa) between the interviewers and first author was .95 for MDD, and .61 for previous MDD.

Depression

Primary outcome measure was the Dutch version of Center for Epidemiologic Studies Depression scale (CES-D), a 20-item self-report questionnaire on depressive symptomatology experienced during the past week (Bouma et al., 1995; Radloff, 1977). The total scores range from 0-60. A score of $\geq$ 16 indicates the presence of clinically relevant depression (Beekman *et al*, 1997; Radloff and Teri, 1986). The psychometric properties are generally good, with Crohnbach's α's ranging from .80-.90 in a sample of Dutch seniors (Beekman et al., 1994), and because of the absence of physical symptoms of depression, the questionnaire is considered very suitable for the elderly (Beekman, et al., 1994; Lewinsohn, et al., 1997; Radloff & Teri, 1986).

Depression and Anxiety

The Dutch version of the Hospital Anxiety and Depression Scale (HADS) was used. The HADS was developed as a brief self-assessment scale to detect states of depression and anxiety in medical outpatients (Zigmond & Snaith, 1983). The HADS consists of a sum scale and two subscales that differentiate between depression and anxiety. Both subscales consist of seven items with a score range of 0-21. Validation studies by Flint and Rifat (1996) and Spinhoven et al. (1997) confirmed the two-factor structure and showed α's ranging from .71-.90 for the sum scale and both subscales.

Health related quality of life

The Dutch version of the Medical Outcome Study Short Form General Health Survey (MOS-SF-20) was used. This self-report questionnaire (20 items) consists of six scales covering mental health, perceived health (mental and physical), social and role functioning, physical functioning and pain. A high score means a high level of

functioning except for the pain scale, where a high score indicates a high level of pain. The α of the scales varied between .80 and .91 (Kempen, 1992; Stewart & Ware, 1988).

Data Analyses

Preliminary analyses

Because of the group format of the intervention, the participants in the IC could not be considered independent observations. To assess the amount of variance attributable to group differences a random coefficient regression model (RCRM) was used to estimate the intra-class correlation in the IC over the four assessment times. Preliminary analyses included checks for normality and the computation of descriptive statistics. All variables were distributed acceptably close to normal. They were described computing frequencies, means and standard deviations.

Analysis of variance (ANOVA) and $\chi 2$-tests were used to compare the following groups on baseline characteristics: (a) the randomized participants and those who refused to be randomized (but had no objections to being interviewed), (b) participants randomized to IC and to WLC, (c) those dropping out of the intervention and those who complied with the course, and (d) the participants who left the study after the posttreatment assessment and those who completed both FU assessments.

Pre- and posttreatment - controlled data

The effects of the intervention were assessed on the completers sample by using a 2 x 2 x 2 split-plot design, using the presence of MDD (Depression) and IC as between-subject variables and time as within-subject variable (Depression x Condition x Time). Main and interaction effects were tested using the mulitvariate criterion of Wilks' lambda (Λ). The analyses were repeated (a) with comorbid anxiety disorder as a third between-subjects factor and (b) with the CES-D on the intention-to-treat (ITT) sample, which included the subjects who dropped out of the IC. A last-value-carried-forward procedure was used to provide data for missing values that occurred because of dropout.

To assess the clinical significance of change on an outcome measure in a clinical population Jacobson and Truax (1991) proposed two criteria: (a) the change should move the individual outside the range of the dysfunctional population (referred to as change in status) and (b) the change should be statistically reliable and exceed the measurement error (referred to as reliable change). We did not assess clinical diagnoses at posttreatment, but used the score on the CES-D as an indication of functional status: those with a score ≥ 16, the recommended cut point (Beekman et al., 1997; Radloff and Teri, 1986) were considered to be dysfunctional. A change in status was defined as a change from a pretreatment score on the CES-D of ≥ 16 (dysfunctional status) to a posttreatment score of < 16 (functional status) or vice versa. The significance of the changes in status within each condition was analyzed with McNemar tests. In our sample a reliable change (RC) was a change of more than |8.6|

on the CES-D, resulting in three RC categories: a reliable improvement, a change that is uncertain (falling within the boundaries of measurement error), and a reliable worsening; χ^2-tests were used to compare the distribution of the RC categories over both conditions.

The between-group effect sizes for all the outcome variables were calculated with Cohen's *d (*Cohen, 1988). In the ITT sample this was done only for the CES-D, HADS and MOS-SF20.

Follow-up measurements – uncontrolled data

The maintenance effects of the intervention at the two follow-up measurements were assessed with a 2 x 3 splitplot design, using the presence of MDD at baseline as between-subject variable and posttreatment and the two follow-up measurements as within-subject variables (Depression x Time). Wilks' Λ was used as the multivariate criterion for significance. McNemar tests were performed to assess changes in functional status between posttreatment and 14 month FU. The within-group effect sizes (d_{impr}) were also calculated with Cohen's *d*.

RCRMs were fitted using the multilevel analysis software MLwiN 1.10 (Rasbach et al., 2000). For all data analyses the Statistical package for Social Sciences version 11.5 package (SPSS, Chicago, IL, U.S.A.) was used.

Results

The CMHCs accepted 246 persons into the program. Overall, 109 (44%) participants were excluded by the health care professionals because (a) they did not meet the inclusion criteria of the study (age and no concurrent therapy) or (b) refused to join the study; so altogether 137 participants were randomly allocated to either the IC or the WLC. We removed people from the analysis if they received psychotherapy during the course or waiting period (six in each), five others withdrew from the study, (two in the IC and three in the WLC). Pretreatment data of one participant were incomplete. In sum, pretreatment data of 119 participants (referred to as the ITT sample) were available for the analyses: 61 in the IC and 58 in the WLC. To analyze possible selection bias, we compared the pretreatment data of this sample with the data of 69 individuals who refused randomization, but consented to the administration and use of the pretreatment measures.

In the IC, nine participants (13%) dropped out of the course; the reasons were medical (two), course not suitable (two), improvement (one), deterioration (one), and unknown reasons (three). There were no drop-outs in the WLC. Participation rate for the completers was high with a mean number of nine sessions (*SD* =1.0). The completer comprised 110 participants. Their mean age was 64.2 years (*SD* =7.2; range 55-85), the majority was female (69%) and of Dutch origin (90%); 55% was married or cohabiting and 52% reported less than 11 years of formal education. The majority (72%) reported the presence of at least one chronic medical condition.

Table 1. *Mental health characteristics of completers sample*

Variable	Intervention (*n*= 52) *n* (%)	Waiting List (*n* = 58) *n* (%)
Antidepressants and/or sedatives[a]	26 (50)	27 (47)
Axis I Disorders:		
No axis 1 disorder	17 (33)	22 (38)
Anxiety disorders [b]	14 (27)	14 (24)
MDD	9 (17)	12 (21)
MDD + anxiety disorders	12 (23)	10 (17)
MDD history		
Never an MDD	6 (12)	6 (10)
Remission	25 (48)	30 (52)
First episode	2 (4)	4 (7)
Recurrent	11 (21)	10 (17)
Chronic	8 (15)	8 (14)

MDD= Major Depressive Disorder based on DSM-IV criteria.

[a] includes St John's Wort; [b] can be more than one anxiety disorder: panic disorder, agoraphobia, social phobia, obsessive-compulsive disorder, post traumatic stress disorder or general anxiety disorder.

Preliminary Analyses

At pretreatment the mean CES-D score for the samples was 24.71 (*SD* = 9.82). Mental health characteristics are summarized in Table 1. In both IC and WLC approximately two-thirds of the participants met the criteria for at least one axis I disorder. The most prevalent disorders were MDD (39%) and anxiety disorders (45%). Previous MDD was reported by 69% of the participants. There were no significant differences (all $p > .10$) in sociodemographic characteristics, presence of MDD, or co-morbid anxiety disorders, or the use of psychotropic medication between (a) the those who refused randomization ($n = 69$) and those who participated in the study ($n = 119$), (b) participants randomized to IC ($n = 61$) and those randomized to WLC ($n = 58$), (c) individuals who completed the intervention ($n = 52$) and those who dropped out ($n = 9$), and (d) individuals who left the study after post-treatment ($n = 10$) and those who completed both follow-up assessments ($n = 42$).

The intraclass correlation of .055 wass not significant, $F(13, 211) = 0.72$, $p > .10$. Hence we decided to ignore the amount of variance attributable to group differences[1].

[1] The IC sample with 13 groups is small, therefore restricted iterative generalized least square (RIGLS) algorithms were used (Kreft and de Leeuw, 1998).

Table 2. *Outcome measures of completers sample and between-groups effect sizes for both the completers and intention-to-treat (ITT) samples*

	Intervention Condition (n = 52) (MDD n = 21)				Waiting List Condition (n = 58) (MDD n = 22)				Effect Size	
Variable	Pre		Post		Pre		Post		Compl	ITT
	M	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	n=110	n=119
CES-D [a]										
Sum	25.84	(10.00)	18.77	(9.23)	23.69	(9.62)	23.63	(10.85)	0.49	0.37
MDD	31.95	(8.26)	21.50	(9.66)	30.91	(8.14)	29.91	(8.62)	0.92	0.77
No MDD	21.71	(9.00)	16.93	(8.60)	19.28	(7.63)	19.80	(10.36)	0.30	0.21
HADS-S[a]										
Sum	19.88	(6.70)	16.04	(6.79)	20.31	(7.03)	19.30	(7.78)	0.45	
MDD	23.65	(6.27)	18.67	(7.58)	25.0	(6.16)	24.36	(5.71)	0.85	
No MDD	17.33	(5.78)	14.26	(5.66)	17.44	(5.96)	16.21	(7.27)	0.30	
HADS-D[a]										
Sum	9.65	(3.81)	7.56	(4.08)	9.97	(4.32)	9.34	(4.30)	0.42	
MDD	11.43	(4.25)	9.05	(4.65)	12.45	(4.19)	11.59	(3.83)	0.60	
No MDD	8.44	(2.99)	6.55	(3.36)	8.44	(3.68)	7.97	(4.03)	0.38	
HADS-A[a]										
Sum	10.22	(4.34)	8.48	(3.59)	10.34	(3.89)	9.95	(4.32)	0.37	
MDD	12.21	(4.27)	9.62	(3.93)	12.55	(3.88)	12.77	(2.94)	0.91	
No MDD	8.87	(3.90)	7.71	(3.19)	9.00	(3.28)	8.23	(4.15)	0.14	
MOS-MH[b]										
Sum	45.77	(14.44)	54.38	(16.47)	46.55	(18.44)	46.76	(20.08)	- 0.41	
MDD	37.9	(11.77)	50.67	(18.60)	33.27	(14.63)	34.36	(13.5)	-1.00	
No MDD	51.1	(13.77)	56.9	(14.63)	54.67	(15.71)	54.33	(19.79)	- 0.15	
MOS-PH[b]										
Sum	48.27	(24.59)	52.88	(24.68)	51.14	(23.18)	48.95	(24.92)	- 0.16	
MDD	38.57	(25.26)	46.67	(26.52	42.95	(23.18)	41.14	(20.58)	- 0.23	
No MDD	54.84	(22.19)	57.1	(22.83)	56.29	(24.80)	53.86	(26.40)	- 0.13	
MOS-P[c]	50.0	(32.84)	49.52	(29.49)	42.67	(36.88)	48.28	(34.69)	- 0.04	
MOS-PF[b]	55.13	(32.51)	55.13	(32.6)	59.48	(33.97)	60.49	(31.20)	0.17	
MOS-SF[b]	66.15	(24.59)	68.46	(25.47)	66.90	(28.42)	63.45	(26.79)	- 0.19	
MOS-RF[b]	48.53	(43.99)	46.57	(46.91)	57.14	(46.15)	54.91	(45.58)	0.18	

MDD = Major Depressive Disorder (DSM-IV criteria.) CES-D = Center for Epidemiologic Studies Depression scale; HADS-S = HADS sum scale; HADS-D = HADS depression scale; HADS-A = HADS anxiety scale; MOS-MH = MOS mental health scale; MOS-PH = MOS perceived health scale; MOS-P = MOS pain scale; MOS-PF = MOS physical functioning scale; MOS-SF = MOS social functioning scale; MOS-RF = MOS role functioning scale. Pre = pre-treatment; Post = post-treatment. Compl = completed course; ITT = intention-to-treat.

[a] high score = more symptoms; [b] high score = better health; [c] high score = more pain.

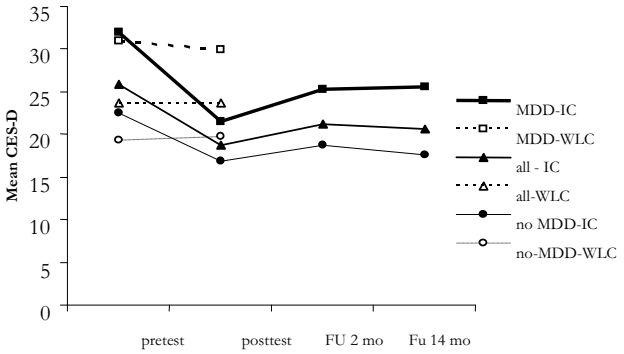

Figure 1. Development of depression over time on the CES-D in the intervention condition (IC) and the waiting list condition (WLC). Further breakdown of the IC and WLC into subgroups: with and without major depressive disorder (MDD, no-MDD).

Comparison of intervention and control subjects

All of the results on the outcome measures for both conditions and separately for those with and without MDD at pretreatment are presented in Table 2, the effect sizes for both the completers sample as the ITT sample are presented in the last column. Multivariate tests of the 2 x 2 x 2 repeated measures ANOVA indicated a significant main effect of Depression [Λ=0.70, F(10, 97) = 4.12; p <.001], but not of Condition [Λ = 0.90, F(10, 97) = 1.11, p = .363], or of Condition x Depression [Λ = 0.94, F(10, 97)=0.63, p=.788]. We found a significant main effect for Time [Λ=0.75, F(10, 97) = 3.32, p =.001], and a significant interaction effect for Time x Condition [Λ = 0.81, F(10, 97) = 2.29, p =.019]. The interaction effects Time x Depression [Λ = 0.92, F (10, 97) = 0.85, p = .585] or Time x Condition x Depression [Λ = 0.93, F(10, 97) = 0.73, p =.691] were not significant.

Further univariate tests indicated significant (all p values < 0.05) main effects of Time for CES-D [F(1,106) = 22.95], the HADS sum scale [F(1,106) = 18.67], the HADS depression scale [F(1,106) = 16.25], the HADS anxiety scale [F(1,106) = 10.59] and the MOS-mental health scale [F(1,106) = 11.71]. Significant (all p values < 0.05) interaction effects of Time x Condition for CES-D [F(1,106) = 20.22], the HADS sum scale [F(1,106) = 7.23], the HADS depression scale [F(1,106) = 4.46], the HADS anxiety scale [F(1,106) = 5.92], MOS-mental health scale [F(1,106) = 9.95], and the MOS-perceived health scale [F(1,106) = 5.67]. A significant Time x Depression interaction effect was found only for the CES-D [F(1,106) = 4.80], the primary outcome measure. The results suggest that with regard to depression and anxiety symptomatology, the IC subjects who completed the course improved

significantly compared with the WLC subjects, regardless of their diagnostic status at pre-treatment. In the IC, those with an MDD at pre-treatment showed the greatest decrease in the score of the CES-D; however, the three-way interaction effect was not significant, this could be a result of low power. Repeating the analyses with comorbid anxiety disorder as a third between-subject factor, which had a significant main effect too, yielded similar results. For the ITT sample the results with the CES-D were also similar. Figure 1 shows the pre- and post treatment changes in both groups.

Clinical significant change

Table 3 shows, for both the IC and the WLC, the changes in functional status from pre- to posttreatment and the RC categories (improvement, uncertain change and worsening). The changes in functional status were significant in the IC [McNemar $\chi^2(1, n = 52) = 8.64, p = .002$], but not in the WLC [McNemar $\chi^2(1, n = 58) = 0.17, p = .10$]. The IC differed significantly from the WLC with regard to the proportion of RC categories, with a higher proportion of improvement in the IC [$\chi^2(2, n = 110) = 10.29, p = .006$].

Table 3. *Assessment of clinical relevance of change – completers sample*

	Intervention Condition *n* = 52	Waitlist Condition *n* = 58
CES-D ≥ 16		
Pre-treatment	44 (83%)	45 (78%)
Post-treatment	32 (62%)	45 (78%)
Change functional status *		
Stayed <16	7 (13%)	6 (10%)
Stayed ≥ 16	31 (60%)	38 (66%)
Became functional	13 (25%)	7 (12%)
Became dysfunctional	1 (2%)	7 (12%)
Reliable Change **		
Improvement (RC ≥ 8.6)	22 (42%)	11 (19%)
No reliable change	29 (56%)	39 (67%)
Worsening (RC ≤ 8.6)	1 (2%)	8 (14%)

CES-D = Center for Epidemiologic Studies Depression scale; Functional status = CES-D <16; Dysfunctional status = CES-D ≥ 16.

* McNemars test for significance of change in IC $p = .002$; ** χ^2 test for significance of RC between conditions $p = .006$.

Maintenance of intervention effect over 2 months and 14-month FU

Complete datasets of 42 participants were available for assessment of the maintenance of the effects. The changes over time on the CES-D are presented in Table 5. Multivariate tests of the 2 x 3 repeated measures ANOVA (posttreatment, and two

follow-ups) showed that neither the main effect of Depression [$\Lambda = 0.86$, $F(5, 36) = 1.15$, $p = .353$], nor the main effect of Time [$\Lambda = 0.75$, $F(10, 31) = 1.02$, $p = .447$], nor the interaction effect of Depression x Time [$\Lambda = 0.78$, $F(10, 31) = 0.87$, $p = .569$] were significant. Furthermore, post-treatment classification of functional status was maintained at 14 months FU by 86% [McNemars $\chi^2(1, n = 42) = 0.17$, $p = .10$]. In sum, there was no significant change over time between post-treatment and 14-month FU.

Effect Sizes

In completers, the overall between-group *ES* was 0.49 with the CES-D (ES_{MDD}=0.92; $ES_{no\text{-}MDD}$ was 0.30) (see Table 2 for the between-group effect sizes and the secondary outcome measures). Compared with pre-treatment, the within-group *ES* ($\underline{d}_{impr}$) for the total group was 0.70 at post-treatment and 0.54 at 14-month FU (Table 4).

Discussion

The main finding is that the CWD course for older adults is effective for older people with subclinical depression as well as for those with a current MDD. The results showed significant improvement for the participants in the IC compared to those in the WLC on the primary outcome measure (CES-D) and on the secondary outcome measures related to mental health (HADS anxiety and depression scales and MOS mental health and perceived health scales). The control group showed no evidence of spontaneous improvement on any of the outcome measures during the 10-week waiting period, a finding also reported in community-based studies on the course of depression in older adults (Beekman et al., 1995).

There were no differences in the use of psychotropic medication between the two groups, and we excluded all participants who followed concurrent psychotherapy. Therefore, the differences in outcome between conditions can not be attributed to differences in the use of psychotropic medication or concurrent psychotherapy.

Our follow-up measurements were restricted to the subjects of the IC and limited to 14 months after post-measurement. The treatment benefits were maintained over this period of time. This is in line with the literature on efficacy studies of psychotherapy for late-life depression (Cuijpers, 1998c; Gallagher-Thompson et al., 1990; Karel and Hinrichsen, 2000).

The between-group effect sizes were consistent with the effect sizes reported in other studies. The ES of 0.30 for the non-MDD subgroup is in keeping with the weighted mean ES of 0.24 reported by Jané-Llopis et al. (2003) in their meta-analyses of prevention studies in elder subjects at risk for the development of depression. The ES of 0.92 for the MDD subgroup is comparable with effect sizes reported in studies in older subjects with clinically relevant levels of depression (Cuijpers, 1998c; Engels and Verwey, 1997; Scogin and McElreath, 1994).

Table 4. *Course of depression with CES-D, caseness (CES-D ≥ 16) and within-group effect size (d_{impr})*

Group	Pretreatment		Posttreatment			FU 2 months			FU 14 months		
	M (*SD*)	≥16	*M* (*SD*)	≥16	d_{impr}	*M* (*SD*)	≥16	d_{impr}	*M* (*SD*)	≥16	d_{impr}
Total group *n* =42	26.55 (10.42)	83%	19.57 (9.62)	64%	0.70	21.2 (10.03)	69%	0.52	20.62 (11.41)	64%	0.54
MDD *n* =16	33.06 (8.36)	94%	22.25 (10.56)	75%	1.14	25.20 (11.63)	75%	0.80	25.56 (11.74)	81%	0.74
No MDD *n* =26	22.54 (9.60)	77%	17.92 (8.79)	58%	0.50	18.73 (8.20)	65%	0.43	17.58 (10.28)	54%	0.50

Data of cases that completed four measurements; CES-D = Center for Epidemiologic Studies Depression scale;
MDD = Major Depressive Disorder based on DSM-IV criteria.

The dropout rate of 15% in the present study was fairly low, compared with the mean dropout rate of 23%, found by Cuijpers (1998c), and the participation rate in the course was high, indicating that the course is an acceptable intervention for this group.

Stepped-care models have been propagated recently with the aim to maximize efficiency of treatment by stepping up the intensity of the intervention according to individual need. Interventions of mild intensity are tried first and, depending on the effect, treatment is continued or stopped (Davison, 2000; Haaga, 2000; Sobell and Sobell, 2000). The CWD course studied here could be fitted into such a stepped-care framework as an intervention of mils intensity. Taking the posttreatment level of functioning as the criterion to determine whether depression was likely, we found that 38% (20 participants) had a CES-D score < 16, and for these individuals the course was probably adequate. However, 62% of the participants scored above the cutoff on the CES-D. The high prevalence of previous episodes of depression, together with the high score on the CES-D, indicates that either they were likely to relapse or they still had an MDD (Haringsma et al., 2004). This conclusion is underscored by the finding that the posttreatment level of functioning had not changed a year later. In this group 13 participants can be classified as improved but not recovered, the course may have been too brief and more sessions may be needed to reach the functional status. Nineteen participants can be classified as unchanged or deteriorated. This can be seen as an indication that a different type of treatment is needed for these participants, which focuses on their specific problems. For instance, a great many participants reported anxiety disorders (46%). For these participants the anxiety symptomatology might be the first treatment target.

We recognize a number of limitations in this study. First, the large proportion of elders not willing to participate in the study could have resulted in selection bias threatening the external validity of the present study results. However, comparison of the elders that joined the study with the 69 elders who refused randomization showed that there were no significant differences between these two groups in socio-demographic characteristics or mental health. Second, because we did not measure the incidence of MDD using a structured psychiatric interview at all the assessments, our results pertain mainly to the level of depressive symptoms experienced by the participants. In order to classify the clinical level of depression we used the cut-off score of the CES-D to describe the individual course of depression over time, a method also described in other studies (Beekman et al., 1995). A third limitation is the lack of a control condition for the follow-up part of the study. This naturalistic follow-up does not allow a definitive conclusion that the course of depression during the follow-up period may be related to the intervention. In addition, the sample size for the follow-up assessment was small. Replication in larger sample sizes is necessary to validate our preliminary conclusion, that in general, treatment gains were maintained over time. Fourth, follow-up was only conducted in the first year following the course. Longer follow-up periods of the subjects are needed to determine how long the protection holds. Fifth, our sample of 13 interventions groups is relatively small for the

use of random coefficient regression modeling (Kreft & de Leeuw; 1998). Replication in a larger sample is needed to confirm our conclusion that the variance due to group differences can be ignored.

The strengths of this study are, first, that this empirically supported depression intervention program, which is provided by the mental health care system on a national scale, was studied in its natural setting. Second, the study incorporated desirable features of both efficacy and effectiveness research, as it is prospective, randomized and focused on a replicable intervention. Third, the study included enough participants to detect a large between-group difference in effectiveness.

Since its introduction in 1995, the CWD course is well embedded in the mental health care system in the Netherlands. This study highlights the fact that it is a valuable intervention, well accepted by the target group. The CWD course was beneficial for all, regardless of clinical diagnosis. However, the post-treatment level of functioning indicated that for 62% of the participants treatment should be continued. The modest effectiveness is not surprising, considering the short duration and the broad character of the program. We recommend that the course is fitted into a stepped-care protocol that varies intervention intensity according to clinical needs; the post-treatment level of functioning can be used to indicate the next step. The course format makes it an attractive, low threshold intervention for a cohort known to be shy of psychological treatment and mental health institutions.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bartels, S. J. *et al.* (2002). Evidence-based practices in geriatric mental health care. *Psychiatric Services*, 53, 1419-1431.

Beekman, A. T. F., Copeland, J. M., & Prince, M. J. (1999). Review of community prevalence of depression in later life. *British Journal of Psychiatry,* 174*,* 307-311.

Beekman, A. T. F., Deeg, D. J. H., Braam, A. W., & DeVries, M. Z. (1997). Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): Results from a community-based sample of older subjects in the Netherlands. *Psychological Medicine, 27,* 231-235.

Beekman, A. T. F., Deeg, D. J. H., Tilburg, T. van, Smit, J. H., & Hooijer, C. (1995). Major and minor depression in later life: A study of prevalence and risk factors. *Journal of Affective Disorders,* 36*,* 65-75.

Beekman, A. T. F., Limbeek, J. van, Deeg, D. J. H., Wouters, L., & Tilburg, W. van (1994). Screening for depression in the elderly in the community: using the Center for Epidemiological Studies Depression scale (CES-D) in the Netherlands. *Tijdschrift voor Gerontologie en Geriatrie,* 25*,* 95-103.

Blazer, D. G. (1998). Late life affective disorders. *Archives of Gerontology and Geriatrics, Suppl.* 6, 43-47.

Bouma, J., Ranchor, A. V., Sanderman, R., & Sonderen, E. van (1995). *Het meten van depressie met de CES-D. Een handleiding.* [Measuring depression with the CES-D. A manual.] Noordelijk Centrum voor gezondheidsvraagstukken, Rijksuniversisteit Groningen, Groningen.

Cohen, J. (1988). *Staistical power analyses for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cuijpers, P. (1998a). A psycho-educational approach to the treatment of depression; a meta-analysis of Lewinsohn's 'Coping With Depression" course. *Behavior Therapy,* 29*,* 521-533.

Cuijpers, P. (1998b). *Kleur geven aan een grijs bestaan* [Coping with depression for older adults]. Baarn: Uitgeverij Intro.

Cuijpers, P. (1998c). Psychological outreach programmes for the depressed elderly: a meta-analysis of effects and dropout. *International Journal of Geriatric Psychiatry,* 13, 41-48.

Davison, G. C. (2000). Stepped care: doing more with less? *Journal of Consulting and Clinical Psychology,* 68*,* 580-585.

Engels, G. I. & Vermey, M. (1997). Efficacy of nonmedical treatments of depression in elders; a quantitave analysis. *Journal of Clinical Gerontology,* 3, 17-35.

Flint, A. J. & Rifat, S. L. (1996). Validation of the Hospital Anxiety and Depression scale as a measure of severity of geriatric depression. *International Journal of Geriatric Psychiatry,* 11*,* 991-994.

Gallagher-Thompson, D., Hanley-Peterson, P., & Thompson, L. W. (1990). Maintenance of gains versus relapse following brief psychotherapy for depression. *Journal of Consulting and Clinical Psychology,* 58*,* 371-374.

Haaga, D. A. F. (2000). Introduction to the special section on stepped care models in psychotherapy. *Journal of Consulting and Clinical Psychology,* 68*,* 547-548.

Haringsma, R., Engels, G.I., Beekman, A.T.F. & Spinhoven, Ph (2004). The criterion validity of the Center for Epidemiological Studies Depression Scale (CES-D) in a sample of self-referred elders with depressive symptomatology. *International Journal of Geriatric Psychiatry,* 19, 558-563.

Jacobson, N. S. & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.

Jané-Llopis, E., Hosman, C. M. H., Jenkins, R., & Anderson, P. (2003). Predictors of efficacy in depression programmes. Meta-analysis*. British Journal of Psychiatry,* 183, 384-397.

Karel, M. J. & Hinrichsen, G. (2000). Treatment of depression in late life: psychotherapeutic interventions. *Clinical Psychology Review,* 20, 707-729.

Kempen, G. IJ. M. (1992). Het meten van de gezondheidstoestand van ouderen. Een toepassing van de Nederlandse versie van de MOS-schaal. [Measuring health in the elderly; an application of the Dutch version of the MOS-scale]. *Tijdschrift voor Gerontologie en Geriatrie,* 23*,* 132-140.

Kreft, I. & Leeuw, J. de (1998). *Introducing multilevel modelling*. London: Sage Publications.

Lewinsohn, P. M., Seeley, J. R., Roberts, R. E., & Allen, N. B. (1997). Center for epidemiologic studies depression scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychology and Aging,* 12*,* 277-287.

Lewinsohn, P. M., Steinmetz, J. L., Antonuccio, D., & Teri, L. (1984). Group therapy for depression: The Coping with Depression course. *International Journal of Mental Health,* 13*,* 8-33.

McCusker, J., Cole, M., Keller, E., Bellavance, F., & Berard, A. (1998). Effectiveness of treatments of depression in older ambulatory patients. *Archives of Internal Medicine,* 158, 705-712.

Overbeek, T., Schruers, K., & Griez, E. (1999). *MINI- The International Neuropsychiatric Interview. Dutch version 5.0.0 DSM-IV*. Maastricht, Netherlands: University of Maastricht.

Radloff, L. S. (1977). The CES-D Scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement,* 1*,* 385-401.

Radloff, L. S. & Teri, L. (1986). Use of Center for Epidemiological Studies-Depression scale with older adults. *Clinical Gerontologist, 5,* 119-136.

Rasbach, J. *et al.* (2004). MLwiN 1.10. Centre for Multilevel Modelling, Institute of Education, University of London, Great Britain. http//multilevel.ioe.ac.uk/.

Scogin, F. & McElreath, L. (1994). Efficacy of psychosocial treatments for geriatric depression: a quantitative review. *Journal of Consulting and Clinical Psychology,* 62*,* 69-74.

Sheehan, D. V., Janavs J, Baker, R., Harnett-Sheehan, K., Knapp E, Sheehan, M. F., *et al.* (1998a). MINI - Mini International Neuropsychiatric Interview - English Version 5.0.0 - DSM-IV. *Journal of Clinical Psychiatry, 59,* 34-57.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., *et al.* (1998b). 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. *Journal of Clinical Psychiatry,* 59*,* 22-33.

Sobell, M. B. & Sobell, L. C. (2000). Stepped care as a heuristic approach to the treatment of alcohol problems. *Journal of Consulting and Clinical Psychology,* 68*,* 573-579.

Spinhoven, Ph., Ormel, J., Sloekers, P. P. A., Kempen, G. IJ. M., Speckens, A. E. M., & Hemert, A. M. van, (1997). A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychological Medicine,* 27*,* 363-370.

Stewart, A. L. & Ware jr, J. E. (1988). De Medical Outcomes Study Short-Form general Health Survey (MOS SF-20). In E.C.Konig-Zahn, J. W. Furer, and B. Tax (Eds.), *Het meten van de gezondheidstoestand, deel 1* (pp. 77-85) [Measuring health status, part 1]. Assen: van Gorcum.

Stirman, S. W., DeRubeis, R.J., Brody, P. E., & Crits-Cristoph, P. (2003). Are samples in randomized controlled trials of psychotherapy representative of community outpatients? A new methodology and initial findings. *Journal of Consulting and Clinical Psychology*, 71, 963-972.

Street, L. L., Niederehe, G., & Lebowitz, B. D. (2000). Toward greater public health relevance for psychotherapeutic intervention research: an NIMH workshop report. *Clinical Psychology: Science and Practice*, 7, 127-137.

Voordouw, I. & Kramer, J. (2001). *Implemetatie van de cursus Omgaan met depressie in de preventieve geestelijk gezondheidszorg. Resultaten van de tussentijdse evaluatie*. [Implementation of the Coping with Depression course in the preventive mental health care. Interim results.] Utrecht: Trimbos Instituut.

Zigmond, A. S. & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica,* 67*,* 361-370.