



UvA-DARE (Digital Academic Repository)

The ten-year course of depression in primary care and long-term effects of psychoeducation, psychiatric consultation and cognitive behavioral therapy

Conradi, H.J.; Bos, E.H.; Kamphuis, J.H.; de Jonge, P.

DOI

[10.1016/j.jad.2017.03.064](https://doi.org/10.1016/j.jad.2017.03.064)

Publication date

2017

Document Version

Final published version

Published in

Journal of Affective Disorders

License

Article 25fa Dutch Copyright Act

[Link to publication](#)

Citation for published version (APA):

Conradi, H. J., Bos, E. H., Kamphuis, J. H., & de Jonge, P. (2017). The ten-year course of depression in primary care and long-term effects of psychoeducation, psychiatric consultation and cognitive behavioral therapy. *Journal of Affective Disorders*, 217, 174-182. <https://doi.org/10.1016/j.jad.2017.03.064>

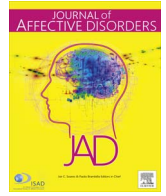
General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)



Research paper

The ten-year course of depression in primary care and long-term effects of psychoeducation, psychiatric consultation and cognitive behavioral therapy



Henk Jan Conradi^{a,*}, Elisabeth H. Bos^b, Jan H. Kamphuis^a, Peter de Jonge^b

^a University of Amsterdam, Department of Clinical Psychology, The Netherlands

^b University of Groningen, Department of Developmental Psychology, The Netherlands

A B S T R A C T

Background: While the majority of depressed patients are treated in primary care, long-term follow-up data on the naturalistic course of depression and treatment effectiveness in this setting are scarce. This study examined the ten-year course of depression in primary care patients who had participated in a randomized clinical trial aiming at enhancement of depression outcomes.

Methods: Of the original sample ($n=267$), 166 patients participated in the ten-year follow-up; missingness was random. Four treatments were compared: (1) Care As Usual (CAU; $n=51$); (2) a Psychoeducational Prevention program (PEP; $n=68$); (3) Psychiatric Consultation followed by PEP (PC+PEP; $n=21$); and (4) brief Cognitive Behavioral Therapy followed by PEP (CBT+PEP; $n=26$). During the first three years interviews based on the Composite International Diagnostic Interview (CIDI) were three-monthly applied, the seven years thereafter were assessed with a once applied CIDI and a face-to-face life chart-based interview.

Results: During the ten-year follow-up 76.5% of the patients developed a new depressive episode, 83.4% used antidepressants (median usage 3.1 years), median depression diagnosis-free time was 9.0 years, and median residual symptom-free time 3.8 years. Treatments did not significantly differ on these outcomes, only trends appeared for lower depression severity for CBT+PEP, and, along with PEP, a higher proportion of symptom-free time.

Limitations: Assessment with the once applied life chart interview (a valid and reliable instrument) is less precise than the three-monthly assessments during the first three years.

Conclusions: The long-term course of depression in primary care is unfavorable, whereas treatment effects over time seem absent or small.

1. Introduction

Depression is a very common disorder, as is marked by its lifetime prevalence of 16.6% (Kessler et al., 2012). The unfavorable long-term course of depression is characterized by very high relapse rates (Mueller, 1999; Solomon, 2000; Simon, 2000), substantial residual symptomatology (ESEMED/MHEDEA 2000 consortium, 2004a), serious impairment, and high health care costs (Kessler, 2005). Moreover, about 10–20% of all cases run a chronic course (Eaton et al., 2008).

Most naturalistic long-term studies concern either community (e.g. Spijker et al., 2001) or psychiatric samples (e.g. Keller et al., 1992). Examining the long-term naturalistic course of depression in primary care, however, is of particular interest because most depressed patients are treated in this setting (ESEMED/MHEDEA 2000, 2004b). Nevertheless, such studies are rare and have limitations, including relative short follow-up periods, i.e. 18 months (Vuorilehto et al., 2009)

to three years (Stegenga et al., 2012). The studies that covered long-term follow-up, e.g. five to 23 years, were comprised because of the method applied, i.e. historical case record examination that did not allow for the assessment of continuous depression outcomes and yielded uncertain diagnosis rates (Van Weel-Baumgarten et al., 1998; Wilson et al., 2003), or, although applying DSM diagnostic criteria using a life chart interview, examined a small sample (Yiend et al., 2009).

The knowledge that is available about the unfavorable short- and medium-term course of depression in primary care (Vuorilehto et al., 2009; Stegenga et al., 2012) underscores the need for effective treatment. Antidepressant medication, the most widely applied treatment strategy, has proven effective in both the acute phase (Cipriani et al., 2009), and at long-term follow-up when applied as maintenance treatment (Geddes et al., 2003). Discontinuation of antidepressants, however, is associated with a return of the risk of relapse (Dobson

* Corresponding author.

E-mail address: h.j.conradi@uva.nl (H.J. Conradi).

et al., 2008; Huijbers et al., 2016), which is problematic, since long-term compliance may not be realistic. A low-intensity alternative to pharmacological interventions is disease management consisting of psychoeducation and motivational techniques. A meta-analysis (Cuijpers et al., 2009) revealed small effects of such low-intensity treatments on depression severity. Cognitive Behavioral Therapy (CBT) shows more favorable outcomes than psychoeducation, with moderate to large effect sizes, which is comparable to other psychotherapies or pharmacotherapy (Cuijpers et al., 2013).

Problematic, however, is that most of the mentioned treatment effectiveness studies are confined to acute phase treatment or at best medium-term follow-up effects of up to two years. A meta-analysis by Steinert et al. (2014) on longer-term treatment effects identified 11 randomized clinical trials (RCTs) with an average follow-up duration of 4.4 years, including the original RCT on which the current long-term follow-up study is based (Conradi et al., 2007). It was found that psychotherapy (mainly CBT), resulted in significantly less relapse than non-psychotherapeutic treatments (mainly care as usual, medication and psychoeducation), i.e. 53.1% vs. 71.3% respectively. Primary care-based long-term treatment studies, however, are absent.

The fact that only 11 longer-term treatment studies were identified is due to high costs associated with conducting such trials, but also with methodological problems like increasing attrition rates and the mounting effect of potential confounders as additional care seeking and medication use. This makes it harder to unravel the effect of the treatment to which patients were originally randomized and that of additional care. To complicate matters further, the treatment to which the patients were originally randomized may also affect the degree to which they consume additional care thereafter.

Taken together, more insight into the long-term course of depression and treatment effects in primary care is needed. In the current study we covered a follow-up of ten years after a randomized treatment phase of several months, and studied the course of depression during this ten-year period in terms of medication use, health care utilization, relapse/recurrence rates, duration of depression diagnosis-free time and symptom-free time, and severity of depression. We examined:

(1) as the main objective of the study the naturalistic long-term course of primary care depression by assessing the course of the outcomes in all available participants;

(2) and as a secondary more explorative study aim the potential differential long-term treatment effectiveness by comparing the outcomes across the four treatments to which patients were originally randomized, i.e. Care as Usual (CAU) by the general practitioner (GP), the Psychoeducation Prevention program (PEP), Psychiatric Consultation followed by PEP (PC+PEP), and CBT followed by PEP (CBT+PEP). The original study revealed no differences in the medium-term on most outcome measures except for PC+PEP and CBT+PEP showing lower severity of depression over the three-years follow-up (Conradi et al., 2007). Based on this finding and previous research showing favorable outcomes of CBT (Cuijpers et al., 2013) we anticipated CBT+PEP to have the most favorable long-term course. Although PC+PEP also showed lower depression severity at the three-year follow-up, prior research shows that favorable outcomes of antidepressants only hold when applied as maintenance treatment (Geddes et al., 2003). Because compliance during the ten-years period we studied, however, may not be realistic, this inevitably will result in an increase of the risk of relapse (Dobson et al., 2008; Huijbers et al., 2016). Therefore we anticipated a less favorable outcome with PC+PEP.

2. Methods

2.1. Patients and procedure

We sought to contact all patients who were included in the original RCT (INSTEL), conducted in primary care between January 1998 and

June 2003 (for details see Conradi et al., 2007). Inclusion criterion for INSTEL was meeting criteria for a current or recent Major Depressive Episode (MDE) treated by the GP. Exclusion criteria were suffering from a life-threatening somatic disease, meeting criteria for bipolar disorder, psychosis, substance abuse or dependency, dementia, being pregnant, or being already in psychotherapy for depression. Originally, 267 patients were randomized to one of four treatments: CAU ($n=72$), PEP ($n=112$), PC+PEP ($n=39$) or CBT+PEP ($n=44$). Because CBT+PEP and PC+PEP were expected to have greater positive effects than PEP-only in comparison to CAU, fewer patients were randomized to these two treatments. CAU consisted of brief supportive counseling, possible antidepressant prescription, and/or referral according to clinical guidelines. PEP was a low-intensity psychoeducation-based program consisting of three face-to-face sessions and short quarterly telephone contacts in the three years thereafter. In the PC+PEP condition one session with a psychiatrist, mainly focusing on antidepressant medication, preceded PEP, and in the CBT+PEP arm on average 10 sessions of CBT were provided prior to PEP. Patients in the INSTEL study were followed-up for up to three years (average 2.75 years; $SD=0.48$). The INSTEL study was approved by the Medical Ethics Committee of the University Medical Center Groningen (MEC96/02/028c).

The present Long-Term INSTEL (LTI) follow-up study took place between October 2010 and June 2012. After consent from their GP, patients were contacted by mail and subsequently by telephone. After reading the information brochure 166 patients signed the informed consent; CAU ($n=51$), PEP ($n=68$), PC+PEP ($n=21$) and CBT+PEP ($n=26$). Next, they were face-to-face interviewed by an experienced research assistant for about two hours. The procedure was approved by the Medical Ethics Committee of the University Medical Center Groningen (METc2009.319). Patients received a 15 euro coupon for participation.

2.2. Instruments

Outcomes during the three-year follow-up of the INSTEL study were assessed with the Composite International Diagnostic Interview (CIDI Auto 2.1; WHO, 1997; Ter Smitten et al., 1998) a valid and reliable structured interview (Wittchen, 1994). The lifetime CIDI was administered face-to-face at baseline and the end of follow-up concerning the previous three years. In-between a slightly adapted version of the CIDI was administered three-monthly by telephone. The adapted version contained additional questions probing onset and remission of each of the DSM-IV symptoms, subsequently allowing determination of diagnosis, duration of depressive episodes, depression-free time and symptom-free time. Questions with respect to medication and health care utilization were added.

Outcomes concerning the seven years after the end of the INSTEL study, i.e. the LTI follow-up, were covered by two face-to-face interviews at the patient's home in a single two hour session. First, as in INSTEL, the lifetime CIDI was administered. The CIDI contained extra questions with which month and year of onset and remission of the identified MDEs were established and subsequently duration of depression-free time. Second, a version of the Longitudinal Interval Follow-up Evaluation (LIFE) life-chart based interview as used by Yiend and colleagues (2009) was administered to measure month-by-month severity of depression and proportion of symptom-free time during the follow-up. The LIFE has shown good to excellent ICCs (Keller et al., 1987). Research has shown that retrospective long-term recall is a valid method when accompanied with proper anchoring of major events (Wells and Horwood, 2004). Therefore we provided patients with three types of anchor points. First, interviewers and patients spent approximately one hour to identify key personal and historical events that were used as aids for retrieval of severity of depressive complaints. These events were: relationships (start, crises and breakup), education and work (exams and change of jobs of self, partner and children), housing (moves), birth, diseases and death (self

and others), finance, birthdays and anniversaries, holidays and journeys, other life events (trauma etc.), and historical events. Second, year and month of onset and remission of the MDEs identified with the adapted lifetime CIDI were added to the LIFE chart. Third, two extra anchor points were used to support memory, namely the Beck Depression Inventory (BDI; Beck et al., 1961) scores at the end of the INSTEL study (the point from where the LIFE interview started), and at the day the LIFE interview was administered. For this purpose both BDI scores were converted to the 5-point response scale that was used in the LIFE interview to estimate the severity of depressive complaints. These scores were: 1 'not at all' (BDI 0 – < 5), 2 'little' (BDI 5 – < 10), 3 'pretty much' (BDI 10 – < 19), 4 'much' (BDI 19 – < 30), and 5 'very much' (BDI ≥ 30). For the statistical analyses, monthly scores were averaged into three-monthly scores in order to improve reliability. The research assistants who administered the CIDI and the LIFE chart interview were experienced interviewers who were extensively trained in both methods and were supervised about every 3 months by the project leader (HJC).

2.3. Outcome measures

Because the current study was developed as a follow-up of INSTEL, we selected outcomes that mirrored those used in the original RCT. All outcomes were computed to cover: (1) the duration of the LTI, i.e. 7.23 years ($SD=0.50$), and (2) the duration of the entire follow-up of INSTEL+LTI, i.e. 9.97 years ($SD=0.20$).

Duration of antidepressant medication usage, and healthcare utilization, i.e. number of contacts with the GP and number of sessions of psychotherapy (secondary care, psychiatric polyclinics and private practices ran by psychologists/psychiatrists), were assessed during LTI with the same questions as used during INSTEL (Conradi et al., 2007).

Relapse/recurrence rate, i.e. the percentage of patients who experienced at least one new MDE according to DSM-IV criteria after remission of the index-episode, was derived from the adapted versions of the CIDI during INSTEL and LTI.

Number of MDEs, was determined by the adapted versions of the CIDI used during INSTEL and LTI.

Proportion depression diagnosis-free time, i.e. the time during follow-up that patients did not meet DSM-IV criteria for MDE, was derived from the adapted versions of the CIDI during INSTEL and LTI.

Proportion symptom-free time, i.e. the time during follow-up that patients did not suffer from depressive symptoms. This outcome was measured during INSTEL with the adapted version of the CIDI and during LTI with the life chart-based interview.

Severity of depression in INSTEL was based on the number of CIDI symptoms and during LTI on the life chart-based interview (5-point severity scale). INSTEL and LTI scores were harmonized by converting the INSTEL CIDI number of symptoms (range 0–9) to the LTI response scale (range 1–5) by dividing the CIDI score by 2.25 and adding 1.

2.4. Power analysis for pairwise comparisons

We calculated the post-hoc achieved power for the between-treatment group comparisons for the severity of depression measurements during the ten-year follow-up with G*Power 3.1.9.2. The calculation was based on a Repeated Measurement ANOVA with 40 measurements (120 months aggregated per 3 months) with 0.8 intercorrelations. Because G*power requires the total sample size of the two compared groups in order to compute the power, and because groups sizes were unequal in the current RCT, we calculated the harmonic mean of the two sample sizes and multiplied this by 2 to compute the total sample size. G*Power showed that the power to detect a small to moderate effect size of Cohen's $d=0.30$ in the comparison of CAU vs. PEP was 0.95, CAU vs. PC+PEP 0.71, CAU vs. CBT+PEP 0.78, PEP vs. PC+PEP 0.75, and PEP vs. CBT+PEP 0.82. This means that, except for the comparisons of CAU vs. PEP, and PEP

vs. CBT+PEP, comparisons were underpowered. The comparisons concerning the continuous outcomes that were based on categorical variables (e.g. proportion depression diagnosis-free time) and the categorical outcomes (e.g. relapse/recurrence rates) were likely even more underpowered.

2.5. Statistical analyses

Because we were not able to trace and include all the patients from the original INSTEL sample, we analyzed, following a two-pronged strategy, whether drop-out status was random. We statistically compared: (a) LTI participants versus patients who did not participate in the LTI study (2 groups) in order to test for representativeness of the LTI sample for the whole original sample, and (b) the four treatment groups within the LTI sample in order to test for equality with respect to baseline characteristics (4 groups). We compared these groups on socio-demographic and clinical variables at the baseline measurement from the original INSTEL study. To that end, independent samples t -tests, one-way ANOVAs and Chi -square tests were applied.

The naturalistic course of the outcomes during follow-up in all individuals participating in the LTI was described by computing percentages for categorical variables, and medians and IQRs for all continuous variables. For severity of depression we report the estimated marginal means and SEs as obtained by the linear mixed model described below.

Comparisons between the treatment groups were guided by the hypotheses from the original study. In order to examine whether CAU could be improved, we compared CAU with PEP, PC+PEP and CBT+PEP. We further examined whether PEP could be enhanced by comparing PEP with PC+PEP and CBT+PEP. Differences during follow-up between groups on categorical outcomes were examined with Chi -square tests, and concerning continuous outcomes not normally distributed with non-parametric Mann-Whitney U tests. To test for differences between treatments on severity of depression we applied linear mixed models on the repeated measurements during follow-up. This analysis allows for evaluation of effects over time, while making optimal use of the available data at the repeated assessments and taking into account the clustering of assessments within subjects (Bryk and Raudenbush, 1987). Treatment group and severity of depression at baseline, to control for initial differences between groups, were used as predictors. AR1 was applied as covariance structure. A random intercept was added to the model. We were interested in the post-hoc pairwise comparisons between the treatment groups. Significance levels were set at $p < 0.05$ (two-tailed). The sample size in the analyses of the entire follow-up was 166 for most outcome measures. For relapse/recurrence, and proportions of depression-free time and symptom-free time, the sample sizes were 152, 155 and 155, respectively, due to missing data during the first three years. Finally, effect size (Cohen's d) for severity of depression was computed, using estimated marginal means and SD s of the raw scores.

3. Results

3.1. Patient characteristics and missingness at random

Sociodemographic and clinical characteristics at baseline of patients who were followed-up ten years later (LTI sample, $n=166$) and those who were not (non-LTI sample, $n=101$) are displayed in Table 1. The LTI patients were on average 42.6 years ($SD=11.3$) old (i.e. about 52 years old at the ten-year follow-up), 72.3% were female, mean BDI at baseline was 20.0 ($SD=8.7$), almost 68% of the patients met diagnostic criteria for recurrent major depression at baseline, and 37% had experienced more than three previous MDEs.

Of the 17 comparisons made between LTI participants and non-participants (Table 1), two significant differences emerged; the LTI sample comprised more women (72.3% vs. 50.5%) and fewer patients

Table 1
Demographic and clinical characteristics at baseline.

	Original INSTEL sample n=267	LTI sample n=166	Non-LTI sample n=101	Test difference LTI vs. non-LTI	Test difference between treatments within LTI
Mean age (SD)	42.8 (11.3)	42.6 (11.3)	43.1 (11.4)	$t=0.54; p=0.59$	$F=0.70; p=0.56$
Female	65.0%	72.3%	50.5%	$X^2=12.95; p < 0.001$	$X^2=0.80; p=0.85$
Social status					
Married/Cohabiting	64.8%	68.1%	59.4%	$X^2=2.70; p=0.44$	$X^2=11.65; p=0.23$
Not Married	19.1%	17.5%	21.8%		
Divorced	12.7%	12.0%	13.9%		
Widowed	3.4%	2.4%	5.0%		
Educational attainment				$X^2=3.00; p=0.22$	$X^2=6.66; p=0.35$
Low	43.8%	39.8%	50.5%		
Medium	36.3%	38.6%	32.7%		
High	19.9%	21.7%	16.8%		
Primary occupation				$X^2=4.29; p=0.12$	$X^2=6.04; p=0.42$
Employed	60.3%	65.1%	52.5%		
Homemaker	19.1%	16.3%	23.8%		
Other	20.6%	18.7%	23.8%		
Severity BDI	20.1 (9.4)	20.0 (8.7)	20.1 (10.5)	$t=-0.59; p=0.95$	$F=0.92; p=0.43$
Severity SCL depression	43.1 (13.1)	43.0 (12.5)	43.5 (14.0)	$t=0.33; p=0.74$	$F=1.32; p=0.27$
Number of DSM-IV criteria for MDE	6.0 (2.1)	6.1 (2.0)	5.9 (2.4)	$t=-1.01; p=0.32$	$F=0.92; p=0.43$
Severity DSM-IV MDE					
Mild	30.3%	30.1%	30.6%	$X^2=0.85; p=0.65$	$X^2=8.84; p=0.18$
Moderate	31.8%	30.1%	34.7%		
Severe	37.9%	39.9%	34.7%		
Recurrent DSM-IV depression	67.2%	67.9%	66.0%	$X^2=0.31; p=0.86$	$X^2=2.46; p=0.87$
> 3 previous DSM-IV MDEs	36.8%	37.0%	36.6%	$X^2=0.001; p=0.96$	$X^2=2.34; p=0.51$
Mean age first onset (SD)	31.1 (13.2)	31.8 (13.0)	30.1 (13.3)	$t=-0.65; p=0.51$	$F=0.60; p=0.62$
Suicide attempt ever	10.1%	10.2%	9.9%	$X^2=0.01; p=0.93$	$X^2=2.66; p=0.45$
Antidepressant medication	74.2%	74.1%	74.3%	$X^2=0.001; p=0.98$	$X^2=0.80; p=0.85$
Last-month comorbidity					
Dysthymia	8.2%	6.6%	10.9%	$X^2=3.62; p=0.31$	$X^2=9.47; p=0.15$
Panic Disorder	12.7%	9.0%	18.8%	$X^2=5.40; p=0.02$	$X^2=1.36; p=0.72$
Agoraphobia	8.2%	10.2%	5.0%	$X^2=2.33; p=0.13$	$X^2=0.74; p=0.86$
Social Phobia	15.4%	13.3%	18.8%	$X^2=1.49; p=0.22$	$X^2=2.35; p=0.50$
SCL total score	195.7 (55.5)	194.7 (50.5)	197.3 (63.3)	$t=0.56; p=0.57$	$F=1.36; p=0.26$
Neuroticism	43.4 (6.7)	42.4 (6.3)	42.3 (7.4)	$t=0.12; p=0.99$	$F=0.30; p=0.83$

with panic disorder the month before baseline (9.0% vs. 18.8%). With respect to the other variables no significant differences were found.

Participation versus non-participation in LTI did not significantly differ between the treatment groups: 70.8% CAU ($n=51$), 60.7% PEP ($n=68$), 53.8% PC+PEP ($n=21$), and 59.1% CBT+PEP ($n=26$) ($X^2=3.73; df=3; p=0.29$). Of the 17 comparisons made between the four treatment groups participating in LTI no significant differences emerged on baseline variables (Table 1). In the group that participated in the LTI the percentage of missing LIFE chart data was 2.3%.

3.2. Antidepressant and health care utilization

The proportion of patients using antidepressants during LTI was 50.6% and 84.3% over the entire ten-year follow-up (Table 2). Within the groups of antidepressant users, median proportion of time of utilization was 0.90 (6.51 years) during LTI and 0.31 (3.10 years) during the entire ten-year follow-up. In total 95.3% of all patients had contacted their GP during LTI; the median number of contacts was 3.00 per year. Over the ten-year follow-up, 100% of the patients had contacted their GP, with a median of 3.23 visits per year. More than half of the patients (53.3%) received some kind of psychotherapy during LTI, and they reported a median number of contacts of 15.00. During the ten-year follow-up 63.9% received psychotherapy outside the randomized treatments, with a median number of sessions of 18.21. No significant differences between treatment groups were revealed in antidepressant and health care utilization.

3.3. Depression outcomes

During LTI 57.1% of the patients suffered a new MDE (median

number of relapses/recurrences was 1.00), and 76.5% did so during the ten-year follow-up (median number of relapses/recurrences was 2.00). No significant differences emerged between treatment groups (Table 3). Median proportion of depression diagnosis-free time during LTI was 0.95 (6.87 years), and 0.90 (9.00 years) during the entire ten-year follow-up, not significantly differing between treatments. The median proportion of symptom-free time during LTI was 0.45 (3.25 years) and during the ten-year follow-up 0.38 (3.80 years). Although no significant differences were found between treatments, a non-significant trend was detected during LTI for PEP and CBT+PEP patients towards a lower proportion of symptom-free time compared to CAU patients. Estimated mean severity of depression during LTI was 1.94 on the 5-point scale used (comparable to a BDI score of 4.7), whereas during the entire ten-year follow-up this was 2.13 (BDI score of 5.7) (Table 3 and Fig. 1). Although no significant differences were found between treatment groups, a non-significant trend appeared during LTI between CAU and CBT+PEP in favor of the last group with an effect size of $d=0.43$.

4. Discussion

Long-term follow-up studies in primary care, the setting in which most depressed patients are treated, are scarce. We examined the ten-year course of depression in primary care patients who had participated in a randomized clinical trial aiming at enhancement of depression outcomes. Over the ten-year follow-up more than three quarters of the patients suffered a new depressive episode, 83.4% used antidepressants with a median usage duration of 3.1 years, median proportion of time patients met MDE criteria was one year, and median duration they suffered from (residual) symptoms was more than 6

Table 2
Antidepressants and health care utilization during LTI and INSTEL+LTI.

n during LTI	All n=165/166 ^a	CAU n=51	PEP n=67/68 ^a	PC+PEP n=21	CBT+PEP n=26	CAU vs. PEP	CAU vs. PC+PEP	CAU vs. CBT+PEP	PEP vs. PC+PEP	PEP vs. CBT+PEP
Antidepressants										
During LTI: percentage of users										
During LTI: median (IQR) proportion of time usage of users only										
During INSTEL+LTI: percentage of users										
During INSTEL+LTI: median (IQR) proportion of time usage of users only										
General Practitioner										
During LTI: percentage of users										
During LTI: median (IQR) number of visits per year of users only										
During INSTEL+LTI: percentage of users										
During INSTEL+LTI: median (IQR) number of visits per year of users only										
Psychotherapy										
During LTI: percentage of users										
During LTI: median (IQR) number of visits of users only ^c										
During INSTEL+LTI: percentage of users										
During INSTEL+LTI: median (IQR) number of visits of users only										

^a First and second numbers represent the range of numbers of patients included in analyses.

^b Not possible to calculate due to 100% users.

^c Based on number of visits during the last year of LTI.

Table 3
Depression outcomes during LTI and INSTEL+LTI.

	All n=163/165 ^a n=152/166 ^a	CAU n=50/51 ^a n=45/51 ^a	PEP n=67 n=63/68 ^b	PC+PEP n=20/21 ^b n=20/21 ^b	CBT+PEP n=26 n=24/26 ^b	CAU vs. PEP	CAU vs. PC+PEP	CAU vs. CBT+PEP	PEP vs. PC+PEP	PEP vs. CBT+PEP
n during LTI										
n during INSTEL+LTI										
Relapse/recurrence										
During LTI: percentage of patients with relapse/recurrence	57.1%	60.0%	58.2%	40.0%	61.5%	0.04; 0.85	2.30; 0.85	0.02; 0.90	2.06; 0.15	0.09; 0.77
During LTI: median (IQR) number of relapses/recurrences	1.00 (0.00–2.00)	1.00 (0.00–2.00)	1.00 (0.00–2.00)	0.50 (0.00–1.00)	1.00 (0.00–2.00)	-0.57; 0.57	-1.08; 0.28	-0.06; 0.99	-0.80; 0.43	-0.48; 0.63
During INSTEL+LTI: percentage of patients with relapse/recurrence	76.5%	77.8%	76.2%	61.9%	87.5%	0.04; 0.85	1.82; 0.18	0.97; 0.33	1.62; 0.20	1.36; 0.24
During INSTEL+LTI: median (IQR) number of relapses/recurrences	2.00 (1.00–4.00)	2.00 (1.00–4.00)	2.00 (1.00–4.00)	1.50 (0.00–3.75)	2.00 (1.00–4.75)	-0.25; 0.81	-1.16; 0.25	-0.33; 0.75	-1.02; 0.31	-0.53; 0.60
Depression-free time										
During LTI: median (IQR) proportion	0.95 (0.81–1.00)	0.92 (0.74–1.00)	0.96 (0.76–1.00)	1.00 (0.87–1.00)	0.95 (0.86–1.00)	-0.50; 0.62	-1.59; 0.11	-0.90; 0.37	-1.27; 0.21	-0.29; 0.77
During INSTEL+LTI: median (IQR) proportion	0.90 (0.77–0.98)	0.87 (0.76–0.97)	0.91 (0.70–0.97)	0.93 (0.82–0.99)	0.90 (0.78–0.97)	-0.08; 0.98	-1.36; 0.17	-0.78; 0.43	-1.32; 0.19	-0.74; 0.46
Symptom-free time										
During LTI: median (IQR) proportion	0.45 (0.00–0.84)	0.05 (0.00–0.75)	0.48 (0.01–0.85)	0.45 (0.01–0.69)	0.75 (0.00–0.92)	-1.83; 0.02	-0.51; 0.61	-1.70; 0.09	-0.90; 0.37	-0.67; 0.50
During INSTEL+LTI: median (IQR) proportion	0.38 (0.06–0.71)	0.30 (0.05–0.65)	0.41 (0.06–0.74)	0.34 (0.06–0.57)	0.58 (0.04–0.80)	-0.95; 0.34	-0.09; 0.93	-1.31; 0.19	-0.64; 0.52	-0.77; 0.44
Severity of depression										
During LTI: estimated mean (SE)	1.94 (0.07)	2.10 (0.11)	1.90 (0.09)	1.97 (0.17)	1.77 (0.15)	0.20; 0.15	0.13; 0.20	<u>0.33; 0.08</u>	-0.07; 0.70	0.13; 0.48
During INSTEL+LTI: estimated mean (SE)	2.13 (0.06)	2.26 (0.09)	2.11 (0.08)	2.12 (0.14)	2.02 (0.13)	0.15; 0.21	0.14; 0.40	0.25; 0.12	-0.01; 0.96	0.09; 0.54

^a First and second numbers represent the range of numbers of patients included in analyses.
^b Tests of the estimated mean differences of severity of depression are based on post-hoc pairwise comparisons from the linear mixed model (n=163 for LTI; n=166 for INSTEL+LTI).

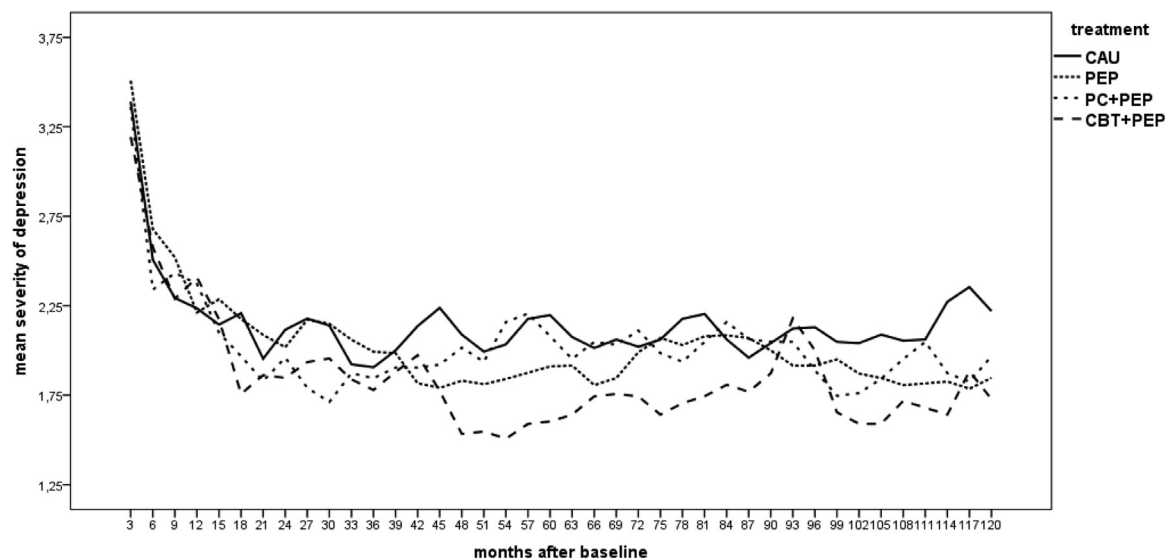


Fig. 1. Depression severity (range 1–5) during 10-year follow-up Mean severity of depression scores (measured with the CIDI from 0 to 36 months and with the LIFE chart from 37 to 120 months) were converted to a 1–5 scale with 1 indicating ‘not at all’ (BDI 0 - < 5), 2 ‘little’ (BDI 5 - < 10), 3 ‘pretty much’ (BDI 10 - < 19), 4 ‘much’ (BDI 19 - < 30), and 5 ‘very much’ (BDI \geq 30). CAU = Care as Usual; PEP = Psychoeducation Prevention program; PC+PEP = Psychiatric Consultation followed by PEP; CBT+PEP = Cognitive Behavioral Therapy followed by PEP.

years. No significant differences between treatments emerged, only trends were noted for lower depression severity in CBT+PEP, and, along with PEP, a higher proportion of symptom-free time.

4.1. The ten-year course of depression in primary care

Since only a few trends for between treatment effects were observed, the ten-year course of depression described in this study can largely be seen as naturalistic for primary care. The emerging image is rather bleak. After remission of the MDE present at baseline, more than three quarters of the primary care patients developed a new MDE during the ten-year course. They developed a median of 2.00 MDEs, with a median total duration of one year in these ten years. Mean severity of depression during these ten years was 2.13 on the 1–5 response scale we used (a BDI score of 5.7), whereas they suffered six year long to some degree from (residual) depressive symptoms. Consequently, health care utilization was high. All patients had contact with their GP during the ten-year follow-up for depression-related and other reasons. A median of more than three visits per year were reported. Almost two thirds of the patients received psychotherapy outside the randomized treatments, with a median of more than 18 sessions during the decade we followed them. Finally, four out of five patients were prescribed antidepressants during the ten-year course and they used them with a median duration of more than 3 years. This suggests that treatment with antidepressants in a developed country like the Netherlands is more substantial than sometimes assumed. Although antidepressant usage was substantial, we cannot determine whether the unfavorable course of depression in this study is not due to undertreatment with antidepressants. On the other hand, longer duration of antidepressant usage was apparently not preferred by patients and/or GPs, as patients were free to decide to (dis-)continue their treatment in the present effectiveness trial. Apart from this, it is uncertain whether treatment with antidepressants on such a substantial scale is indicated given the high relapse risk after discontinuation (Dobson et al., 2008; Huijbers et al., 2016).

Overall, depression outcomes are more unfavorable than previously presented in naturalistic studies of primary care patients. We found 76% of the patients to experience another MDE during the ten-year follow-up while Van Weel-Baumgarten et al. (1998) found 40% during ten year, and Yiend et al. (2009) 50% during ten year and 64% during 23 year. Yiend et al.’s sample was presumably less vulnerable than ours

with 10% recurrent depressed at inclusion, versus 67% in our study. Wilson et al. (2003) reported a comparable recurrence rate of 77.5%, however, during half of our follow-up time. However, the diagnostic reliability of this study was uncertain since diagnoses were based on medical records. Mean duration of depression diagnosis, only reported by the Yiend et al. (2009) study, was 15% of the follow-up period, which was comparable to the median of 10% we found. As far as we know proportion of symptom-free time was not reported in other studies, while severity of depression was measured differently, making comparisons across studies difficult.

The more unfavorable depression course we found presumably was reflected in a higher health care utilization compared to other primary care studies. Yiend reported 23% of the patients using antidepressants with a mean duration of 14.5% of the follow-up time. This compared favorably to the 93.5% using antidepressants in the Wilson et al. study (2003) and the 83.4% in our study during a median of 31% of follow-up period. Finally, Van Weel-Baumgarten et al. (1998) reported 9% of the patients being hospitalized during follow-up, Wilson et al. (2003) 7.3%, which equals our finding of 7.3%. However, Van Weel-Baumgarten et al. (1998) reported a modest 15% referral rate of patients to secondary care, while we found that 63% of the patients sought additional psychotherapeutic treatment. It may be that the aim of INSTEL, enhancement of depression care, stimulated especially GPs and PEP workers to advise patients to seek help timely when warning signs of a pending relapse appeared.

4.2. Treatment effects during the ten-year course

Clearly, it is difficult to examine treatment effects almost a decade after delivery. However, treatment groups did not differ in antidepressant and health care utilization, leaving confounding less probable, and attribution of possible differences to the treatment patients were randomized to more likely. Nevertheless, no significant differences between treatments were found on depression outcomes. Two of the three trends detected favored CBT+PEP over CAU and concerned severity of depression and proportion of symptom-free time. The trend regarding severity during LTI is in line with the significant difference on the BDI we found during the original INSTEL study (Conradi et al., 2007). Converted to BDI scores this meant 5.5 for CAU vs. 3.85 for CBT+PEP. In INSTEL PC+PEP performed better than CAU too, but in the LTI follow-up this effect had disappeared. This may be explained by the

fact that CBT is directed at enhancing patients' behavioral and cognitive coping repertoire which may have an effect long after CBT was ended, while PC+PEP was primarily directed at enhancing antidepressant usage which works as long as the patient is compliant.

Interestingly, both PEP and CBT+PEP patients performed better than CAU patients concerning symptom-free time. For CBT+PEP this may be again explained by the building of a coping repertoire. The trend found for PEP may be a chance finding, because in the original study we did not find favorable outcomes for PEP (Conradi et al., 2007), but if it refers to a real effect it may be understood as a consequence of training patients in recognizing residual symptoms and immediately applying behavioral activation. Of note, we should be cautious with these interpretations since the effects we found were relatively small, and comparisons between CAU and CBT+PEP were underpowered.

4.3. Limitations and strengths

There are some limitations that should be kept in mind when interpreting the findings from this study. First, in contrast with the original INSTEL study, which was characterized by three-monthly assessments of depression and care utilization, in LTI we administered one interview with which we looked approximately seven years back in time. Although we did put a lot of effort in obtaining reliable data by using the previously successfully applied life chart interview (Keller et al., 1987, Yiend et al., 2009), data are probably less precise than the INSTEL data. Second, with such a long follow-up, i.e. a decade, differences between treatment groups cannot automatically be attributed to the original treatments patients were randomized to. Treatment effects may be washed out by additional treatment, although additional health care utilization did not seem to differ between groups. Third, the comparisons with PC+PEP and CBT+PEP were underpowered making it difficult to reach firm conclusions about the absence of effects (type II error). However, because the absolute differences between groups were relatively small possible false negative findings are less of a concern.

Strengths of this study are the long-term follow-up of ten year in a substantial primary care sample. Moreover, participation rate to the follow-up interview was in line with a previous study (Yiend et al., 2009) who reported a loss to their 23-year follow-up of 77.6%. Assumed that loss to follow-up displays a linear trend over time, this would mean that the loss to our ten-year follow-up would be approximately 10 (follow-up years)/23 (follow-up years of Yiend et al.) *77.6% (loss to follow-up of Yiend et al.)=33.7%, which is comparable with our 37.8% loss to follow-up. Finally, missingness was random, meaning that the findings are representative for the cohort that was originally included in the INSTEL RCT and randomization remained successful within the LTI sample.

4.4. Recommendations for research and clinical practice

Important in long-term studies is the selection of depression outcomes. In medium-term studies, some categorical outcomes seem to have higher information value than when used in long-term studies. Because relapse/recurrence rates are highly correlated with follow-up duration (Solomon, 2000), ceiling effects tend to appear with longer follow-ups. Also less useful in long-term studies is persistence or chronicity of MDE (e.g. Vuorilehto et al., 2009). Only a few months of (partial) remission are sufficient to label patients as non-chronic, or intermittent, but when the level of residual symptomatology continues to be high, this categorical label is less informative than severity of (residual) symptomatology. Residual symptomatology is an informative outcome and also very common when patients are remitted (Conradi et al., 2011, 2012). Proportion of depression-free time is also an informative outcome in the long-term, given the fact that interludes between relapses/recurrences vary from months to years (De Jonge et al., 2010). In sum, we think continuous measures are especially

useful assessing long-term outcome, because these measures provide more precise and accurate insight in the state of well-being of the patients. Moreover, continuous measures enable unbiased comparisons between studies independent of length of follow-up.

We further recommend future research to stratify patients based on number of prior episodes. Four previous studies (Conradi et al., 2008; Bockting et al., 2005; Ma and Teasdale, 2004; Teasdale et al., 2000) found that patients who suffered from multiple MDEs profit more from CBT than patients who experienced fewer prior episodes. Vulnerability for rumination may be a key factor (Ma and Teasdale, 2004; Conradi et al., 2008). Unfortunately, the group of patients receiving CBT in our study was too small to perform subgroup analyses with sufficient statistical power.

The possible treatment effects we found were small, which makes sense when following patients over a decade. In order to outperform CAU more clearly, we therefore recommend CBT, the most promising treatment, to be prolonged with booster sessions (i.e. 15–30 sessions), and to be delivered by experienced therapists.

In conclusion, to our knowledge, this is the first study to report on the ten-year course of depression and treatment effectiveness in a substantial sample (N=166) from primary care, i.e. the setting in which most patients are treated. Findings suggest that the long-term course of depression in primary care is unfavorable, whereas treatment effects seem absent or small.

Acknowledgements

The LTI study was financially supported by the Netherlands Foundation for Mental Health (NFMH), and the original INSTEL trial by grants from the Dutch Organization for Scientific Research (NWO), the Medical Sciences Program and Chronic Diseases Program, Research Foundations of the Health Insurance Company 'Het Groene Land', the Regional Health Insurance Company (RZG), the Netherlands Foundation for Mental Health (NFMH), and the University Hospital Groningen to J. Ormel and H.J. Conradi.

References

- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., 1961. An inventory for measuring depression. *Arch. Gen. Psychiatry* 4, 561–571.
- Bockting, C.L.H., Schene, A.H., Spinhoven, Ph, Koeter, M.W.J., Wouters, L.F., Huysler, J., Kamphuis, J.H., 2005. Preventing relapse/recurrence in recurrent depression with cognitive therapy: a randomized controlled trial. *J. Consult. Clin. Psychol.* 73, 647–657.
- Bryk, A.S., Raudenbush, S.W., 1987. Application of hierarchical linear models to assessing change. *Psychol. Bull.* 101, 147–158.
- Cipriani, A., Furukawa, T.A., Salanti, G., Geddes, J.R., Higgins, J.P.T., Churchill, R., Watanabe, N., Nakagawa, A., Omori, I.M., McGuire, H., Tansella, M., Barbui, C., 2009. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 373, 746–758.
- Conradi, H.J., De Jonge, P., Ormel, J., 2008. Cognitive-behavioural therapy v. usual care in recurrent depression. *Br. J. Psychiatry* 193, 505–506.
- Conradi, H.J., De Jonge, P., Ormel, J., 2011. Presence of individual (residual) symptoms during depressive episodes and periods of remission: a three-year prospective study. *Psychol. Med.* 41, 1165–1175.
- Conradi, H.J., Ormel, J., De Jonge, P., 2012. Symptom profiles of DSM-IV defined remission, recovery, relapse and recurrence of depression: the role of the core symptoms. *Depress. Anxiety* 29, 638–645.
- Conradi, H.J., de Jonge, P., Kluiters, H., Smit, A., van der Meer, K., Jenner, J.A., van Os, T.W.D.P., Emmelkamp, P.M.G., Ormel, J., 2007. Enhanced treatment for depression in primary care: long-term outcomes of a psychoeducational program alone and enriched with psychiatric consultation or cognitive behavioral therapy. *Psychol. Med.* 37 (6), 849–862.
- Cuijpers, P., Muñoz, R.F., Clarke, G.N., Lewinsohn, P.M., 2009. Psychoeducational treatment and prevention of depression: the “coping with depression” course thirty years later. *Clin. Psychol. Rev.* 29, 449–458.
- Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., Dobson, K.S., 2013. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Can. J. Psychiatry* 58, 376–385.
- De Jonge, P., Conradi, H.J., Kaptein, K.L., Bockting, C., Korf, J., Ormel, J., 2010. Duration of subsequent episodes and periods of recovery in recurrent major depression. *J. Affect. Disord.* 125, 141–145.
- Dobson, K.S., Hollon, S.D., Dimidjian, S., Schmalzing, K.B., Kohlenberg, R.J., Gallop, R.J., Rizvi, S.L., Gollan, J.K., Dunner, D.L., Jacobson, N.S., 2008. Randomized trial of

- behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. *J. Consult. Clin. Psychol.* 76, 468–477.
- Eaton, W.W., Shao, H., Nestadt, G., Lee, H.B., Bienvenu, O.J., Zandi, P., 2008. Population-based study of first onset and chronicity in major depressive disorder. *Arch. Gen. Psychiatry* 65, 513–520.
- ESEMeD/MHEDEA 2000 Investigators, 2004a. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr. Scand. Suppl.* 109, 21–27.
- ESEMeD/MHEDEA 2000 Investigators, 2004b. Use of mental health services in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr. Scand. Suppl.* 109, 47–54.
- Geddes, J.R., Carney, S.M., Davies, C., Furukawa, T.A., Kupfer, D.J., Frank, E., Goodwin, G.M., 2003. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. *Lancet* 361, 653–661.
- Huijbers, M.J., Spinhoven, P., Spijker, J., Ruhé, H.G., van Schaik, D.J., van Oppen, P., Nolen, W.A., Ormel, J., Kuyken, W., van der Wilt, G.J., Blom, M.B., Schene, A.H., Rogier, A., Donders, T., Speckens, A.E., 2016. Discontinuation of antidepressant medication after mindfulness-based cognitive therapy for recurrent depression: randomised controlled non-inferiority trial. *Br. J. Psychiatry* 208 (4), 366–373.
- Keller, M.B., Lavori, P.W., Mueller, T.I., Endicott, J., Coryell, W., Hirschfeld, R.M., Shea, T., 1992. Time to recovery, chronicity, and levels of psychopathology in major depression. A 5-year prospective follow-up of 431 subjects. *Arch. Gen. Psychiatry* 49, 809–816.
- Kessler, R.C., Chiu, W.T., Demler, O., Merikangas, K.R., Walters, E.E., 2005. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* 62, 617–627.
- Kessler, R.C., Petukhova, M., Sampson, N.A., Zaslavsky, A.M., Wittchen, H.U., 2012. Twelve-month and life time prevalence and life time morbid risk of anxiety and mood disorders in the United States. *Int. J. Methods Psychiatr. Res.* 21, 169–184.
- Ma, S.H., Teasdale, J.D., 2004. Mindfulness-based cognitive therapy for depression: replication and exploration of differential relapse prevention effects. *J. Consult. Clin. Psychol.* 72, 31–40.
- Mueller, T.I., Leon, A.C., Keller, M.B., Solomon, D.A., Endicott, J., Coryell, W., Warshaw, M., Maser, J.D., 1999. Recurrence after recovery from major depressive disorder during 15 years of observational follow up. *Am. J. Psychiatry* 156, 1000–1006.
- Simon, G.E., 2000. Long-term prognosis of depression in primary care. *Bull. World Health Organ.* 78, 439–445.
- Solomon, D.A., Keller, M.B., Leon, A.C., Mueller, T.I., Lavori, P.W., Shea, M.T., Coryell, W., Warshaw, M., Turvey, C., Maser, J.D., Endicott, J., 2000. Multiple recurrences of major depressive disorder. *Am. J. Psychiatry* 157, 229–233.
- Spijker, J., Bijl, R.V., de Graaf, R., Nolen, W.A., 2001. Care utilization and outcome of DSM-III-R major depression in the general population. Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta Psychiatr. Scand.* 104, 19–24.
- Stegenga, B.T., Kamphuis, M.H., King, M., Nazareth, I., Geerlings, M.I., 2012. The natural course and outcome of major depressive disorder in primary care: the PREDICT-NL study. *Soc. Psychiatry Psychiatr. Epidemiol.* 47, 87–95.
- Steinert, Ch, Hofmann, M., Kruse, J., Leichsenring, F., 2014. Relapse rates after psychotherapy for depression – stable long-term effects? A meta-analysis. *J. Affect. Disord.* 168, 107–118.
- Teasdale, J.D., Segal, Z.V., Williams, J.M.G., Ridgeway, V.A., Soulsby, J.M., Lau, M.A., 2000. Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *J. Consult. Clin. Psychol.* 68, 615–623.
- Ter Smitten, M.H., Smeets, R.M.W., van Den Brink, W., 1998. Composite International Diagnostic Interview (CIDI). Basis version 2.1. Lifetime Manual. WHO, Amsterdam.
- Van Weel-Baumgarten, E.M., van den Bosch, W.J., van den Hoogen, H.J., Zitman, F.G., 1998. Ten year follow-up of depression after diagnosis in general practice. *Br. J. Gen. Pract.* 48, 1643–1646.
- Vuorilehto, M.S., Melartin, T.K., Isometsa, E.T., 2009. Course and outcome of depressive disorders in primary care: a prospective 18-month study. *Psychol. Med.* 39, 1697–1707.
- Wells, J.E., Horwood, L.J., 2004. How accurate is recall of key symptoms of depression? A comparison of recall and longitudinal reports. *Psychol. Med.* 34, 1001–1011.
- Wilson, I., Duszynski, K., Manta, A., 2003. A 5-year follow-up of general practice patients experiencing depression. *Fam. Pract.* 20 (6), 685–689.
- World Health Organization, 1997. The Composite International Diagnostic Interview (CIDI). World Health Organization, Geneva.
- Wittchen, H.U., 1994. Reliability and validity studies of the WHO-Composite International Diagnostic Interview: a critical review. *J. Psychiatr. Res.* 28, 57–84.
- Yiend, J., Paykel, E., Merritt, R., Lester, K., Doll, H., Burns, T., 2009. Long term outcome of primary care depression. *J. Affect. Disord.* 118, 79–86.