View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Open Research Exeter

Psychological Medicine

Stepped care treatment delivery for depression: a systematic review and meta-analysis

--Manuscript Draft--

Manuscript Number:	PSM-D-13-00741R1
Full Title:	Stepped care treatment delivery for depression: a systematic review and meta-analysis
Article Type:	Review Article
Corresponding Author:	Annemieke van Straten, Ph.D. VU university Amsterdam, NETHERLANDS
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	VU university
Corresponding Author's Secondary Institution:	
First Author:	Annemieke van Straten, Ph.D.
First Author Secondary Information:	
Order of Authors:	Annemieke van Straten, Ph.D.
	Jacqueline J Hill, MSc
	Dave A Richards, Prof. PhD
	Pim Cuijpers, Prof. PhD
Order of Authors Secondary Information:	
Manuscript Region of Origin:	UNITED KINGDOM
Abstract:	Background In stepped-care models patients typically start with a low intensity evidence-based treatment. Progress is monitored systematically and those patients who do not respond adequately step up to a subsequent treatment of higher intensity. Despite the fact that many guidelines have endorsed this stepped care principle it is not clear if stepped care really delivers similar or better patient outcomes against lower costs compared to other systems. We performed a systematic review and meta-analysis of all randomised trials on stepped care for depression. Methods We carried out a comprehensive literature search. Selection of studies, evaluation of study quality, and extraction of data, was performed independently by two authors. Results Fourteen studies were included and ten were used in the meta-analyses (4,580 patients). All studies used screening to identify possible patients and care-as-usual as a comparator. Study quality was relatively high. Stepped care had a moderate effect on depression (pooled six month between group effect size Cohen's d was 0.34; 95% confidence interval 0.20 to 0.48). The stepped care interventions varied a lot in number and duration of treatment steps, treatments offered, professionals involved, and criteria to step up. Conclusions There is currently only limited evidence to suggest that stepped care should be the dominant model of treatment organisation. Evidence on (cost-) effectiveness compared with high intensity psychological therapy alone, as well as with matched care, is required.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Manuscript
Click here to download Manuscript: Stepped care manuscript_resubmission.doc

Stepped care treatment delivery for depression: a systematic review and meta-analysis

Authors

A. van Straten[1,2]

J. Hill[1]

D. A. Richards[3]

P. Cuijpers[2]

Affiliations

[1] Mood Disorders Centre, University of Exeter, Exeter, UK

[2] dept. of Clinical Psychology, VU University Amsterdam & EMGO Institute for Health and Care Research, The Netherlands

[3]University of Exeter Medical School, University of Exeter, Exeter, UK

Corresponding author

A. van Straten

VU university

FPP, Dept. Clinical Psychology

Van der Boechorststraat 1

1081 BT Amsterdam, The Netherlands

Tel: +31205988970

e-mail: a.van.straten@vu.nl

Abstract

Background In stepped-care models patients typically start with a low intensity evidence-based treatment. Progress is monitored systematically and those patients who do not respond adequately step up to a subsequent treatment of higher intensity. Despite the fact that many guidelines have endorsed this stepped care principle it is not clear if stepped care really delivers similar or better patient outcomes against lower costs compared to other systems. We performed a systematic review and meta-analysis of all randomised trials on stepped care for depression.

Methods We carried out a comprehensive literature search. Selection of studies, evaluation of study quality, and extraction of data, was performed independently by two authors.

Results Fourteen studies were included and ten were used in the meta-analyses (4,580 patients). All studies used screening to identify possible patients and care-as-usual as a comparator. Study quality was relatively high. Stepped care had a moderate effect on depression (pooled six month between group effect size Cohen's *d* was 0.34; 95% confidence interval 0.20 to 0.48). The stepped care interventions varied a lot in number and duration of treatment steps, treatments offered, professionals involved, and criteria to step up.

Conclusions There is currently only limited evidence to suggest that stepped care should be the dominant model of treatment organisation. Evidence on (cost-) effectiveness compared with high intensity psychological therapy alone, as well as with matched care, is required.

Keywords: meta-analysis; review; depression; stepped care; collaborative care; self-help; psychological treatment

Introduction

It is generally acknowledged that care for depression could be improved because the delivery and uptake of antidepressant medication and evidence-based psychotherapies is often suboptimal (Simon, 2002; Bijl *et al*. 2003; National Institute for Health and Clinical Excellence, 2011; Piek *et al*. 2011; Piek *et al*. 2012). Improvement of care is more likely to come from changes in the way care is provided than from adding new treatment options (Katon & Unutzer, 2006).

Currently, the standard approach in which mental health care is delivered to patients is called matched care. In this approach the patient is referred to a certain therapist or therapy. The therapy choice is based (matched) on patients' characteristics and preferences. As a result, the treatment may vary (e.g. antidepressant medication or different types of psychotherapy) as well as the setting (primary care, mental health care, online therapy, group therapy, individual therapy) and the provider (e.g. GP, nurse, PWP, psychologist, psychiatrist). A major problem with this model at present is our lack of clear prognostic determinants with which to match patients to the available treatments. It has been argued that some patients receive too much treatment (Lovell & Richards, 2000), whilst others too little, as those lucky enough to be given treatment utilise highly scarce resources to the detriment of many others who receive little or nothing.

An alternative approach is called 'stepped care'. Within the last ten years and in the context of international concern regarding the cost and prevalence of common mental health problems, stepped care has been recommended as a means to increase access and efficiency of mental health care (Andrews *et al*. 2006; NICE, 2009). In stepped-care models, the default position is that patients start with an evidence-based treatment of low intensity as a first step. Progress is monitored systematically and those patients who do not respond adequately will step up to a subsequent treatment of higher intensity (Bower & Gilbody, 2005). Low-intensity treatments are usually defined as those treatments that require less time from a professional than a conventional treatment (Bennett-Levy *et al*. 2010). However, intensity may also mean the time required of patients, cost, and therapists' level of expertise and it is possible for treatments to differ in one but not all of these dimensions. Patients, for example, may themselves spend similar amounts of time undertaking high- or low-intensity treatments which require a different amount of time from a professional.

Whilst the concept of intensity readily applies to psychological therapies, it is difficult to characterise pharmacological and, perhaps, physical treatments as intensive or otherwise. Given the widespread use of pharmacotherapy alongside psychological treatment for depression, it is perhaps

unsurprising that the term 'stepped care' is also used to refer to treatment that is not organised in order of increased intensity; at each 'step' patients switch or add treatments of different modalities (pharmacological, psychological) - patients may start with intensive psychological therapy (Araya *et al*. 2003; Katon *et al.* 2004; Ell *et al.* 2008).

In practice, self-help treatments (through books or the internet) are often used as a first step in stepped care. The effectiveness of self-help for depression, guided by a mental health worker but still of less intensity than traditional psychological therapy, has been demonstrated convincingly (Gellatly *et al*. 2007; Andrews *et al.* 2010; Cuijpers *et al*. 2010; Richards & Richardson 2012). Therefore, the assumption of stepped care is that for most patients the low intensity treatment will be sufficient and only few will need a higher intensity treatment, thereby making better use of scarce and expensive resources such as therapist time. Many depression treatment guidelines have endorsed this stepped care principle e.g. the English NICE guideline (NHS National Institute for Health and Clinical Excellence, 2009; National Collaborating Centre for Mental Health, 2010), and the Dutch multidisciplinary guideline (Spijker *et al*. 2010). This has also led to implementing stepped care in routine practice. The most notable initiative in this respect is the implementation of the Improving Access to Psychological Therapies (IAPT) programme (www.iapt.nhs.uk), for which stepped care underpins the organisational structure.

The question remains how much evidence there is for the effectiveness of stepped care. Does stepped care really deliver similar or better patient outcomes compared to other systems? Although, observational data from the first year of English IAPT services show that recovery rates were higher in services making use of the full range of low and high-intensity treatments in stepped care systems (Clark, 2011) no systematic review of randomised trials has been published yet. Therefore, our aim in this study was to conduct a systematic review and meta-analysis of studies investigating the effectiveness of stepped care for depression.

Methods

Search strategy

We carried out a comprehensive literature search in PubMed, PsycINFO, EMBASE, and the Cochrane Central Register of Controlled Trials. We combined terms indicative of depression with those of stepped care, e.g. for Medline we used (depression [Mesh] OR depressive disorder [Mesh] or mood disorders [mesh]) AND (stepped [all fields] AND care [all fields]). We searched all literature up to April 2012

without any language restrictions and followed up identified protocol papers published before April 2012 to determine if the researchers had subsequently published their findings before May 2013. Two independent researchers (AvS and JH) reviewed all abstracts and titles of retrieved references for eligibility. We retrieved the full papers for all references that had been judged as potentially eligible and the full papers were examined independently by two of the research team (AvS, JH, DR). In case of disagreement the paper was discussed with the third reviewer until a consensus was achieved. We also checked the reference lists of the included papers and a recent meta-analysis on collaborative care (Archer *et al.* 2012).

Inclusion criteria

We used the following inclusion criteria: (1) the study had to be a randomized controlled trial (2) aimed at adults (3) with a DSM-IV depressive disorder identified through a diagnostic interview, or with depressive symptoms established by scoring above a cut-off on a depression questionnaire, (4) investigating 'stepped care' as one of the randomised trial groups. Stepped care had to include psychological therapy and was defined as the availability of more than one psychological treatment of different intensities and/or the availability of more than one treatment modality (pharmacological and psychological). We defined the intensity of psychological treatments with respect to the time to deliver; non-psychological (pharmacological) treatments were not characterised in this respect. We did not require treatments to be organised in a hierarchy of low- to high-intensity. Decisions about stepping up had to be based on a systematic clinical evaluation undertaken by a clinician or through questionnaire assessment, done at a pre-specified time interval and with an explicit aim to determine the next treatment step. We included studies in which only a proportion of patients were depressed, for example studies including patients with a common mental health disorder and a sub-group of patients specifically diagnosed with depression. We allowed both physical and psychiatric comorbidity. Studies were included regardless of their setting or control group.

Data extraction

We coded the following general characteristics of the studies: year of publication, country, randomisation level (patient or cluster), the way depression or depressive symptoms were established (e.g. diagnostic interview or scoring above a cut-off on a questionnaire), possible comorbidity as an inclusion criterion (e.g. cancer patients, diabetes), age, and total number of patients included in the

study. The stepped care interventions were coded as follows: number of steps, the content of the interventions in the different steps, criteria to step up, and total duration of the program. Two independent assessors coded each study and differences were discussed among the review team until consensus was reached.

Quality assessment

We assessed the validity of the studies using the criteria as suggested by the Cochrane Handbook (The Cochrane Collaboration, 2011): adequate sequence generation, concealment of allocation, blinding of outcome assessors, adequate handling of incomplete outcome data, selective reporting of data and other potential threats to validity. Two reviewers conducted the quality assessment independently of each other.

Meta-analyses

We calculated between group effect sizes (Cohen's *d*) for all individual studies. The effect size represents the difference between two groups in number of standard deviations (Hedges & Olkin, 1985; Cooper & Hedges, 1994). To calculate between group effect sizes we used the available statistics as published in the papers (means and standard deviations, mean difference score and 95% confidence interval, or proportions of patients improved or recovered). When more than one outcome was reported (e.g. more than one depression questionnaire or more than one cut-off score) we performed a sensitivity analysis. We pooled the effects using (a) the highest reported effect sizes for all studies and (b) the lowest reported effect sizes for all studies and (c) the average or combined effect size for all studies.

To calculate the individual effect sizes as well as the pooled mean effect size we used the computer program Comprehensive Meta-analysis version 2.2.046 for Windows, developed for support in meta-analysis (www.metaanalysis.com). As we expected considerable heterogeneity, we calculated pooled effect sizes with the random effects model. However, we first tested heterogeneity under the fixed effects model using the statistics I^2 and Q. I^2 describes the variance between studies as a proportion of the total variance. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity, with 25% as low, 50% as moderate, and 75% as high heterogeneity. The statistical significance of the heterogeneity is tested with the Q statistic. A significant Q value rejects the null hypothesis of homogeneity. We mark all results in which the p-value is lower than 0.05.

In addition, we performed subgroup analyses. In these analyses we tested whether there were significant differences between the effect sizes in different categories of studies. We used the mixed effects model, which pools studies within subgroups with the random effects model, but tested for significant differences between subgroups with the fixed effects model. Lastly, publication bias was tested by inspecting the funnel plot, and by Duval and Tweedie's trim and fill procedure, which yields an estimate of the effect size after publication bias has been taken into account (as implemented in Comprehensive Meta-analysis; Duval & Tweedie, 2000).

Results

Inclusion of studies

We retrieved 61 manuscripts for eligibility after screening 438 references (Figure 1). We excluded 47 of the 61 that did not fulfill our inclusion criteria. In total, we included 14 studies on stepped care for depression (Unutzer *et al*. 2002 [13]; Araya *et al*. 2003 [2]; Katon *et al*. 2004 [10]; Ell *et al*. 2008 [7]; Van't Veer-Tazelaar *et al*. 2009 [14]; Bot *et al*. 2010 [3]; Davidson *et al*. 2010 [4]; Ell *et al*. 2010 [8]; Patel *et al*. 2010 [11]; Seekles *et al*. 2011 [12]; Apil *et al*. 2012 [1]; Dozeman *et al*. 2012 [6]; Davidson *et al*. 2013 [5]; Huijbregts *et al*. 2013 [9]). In one trial [3], only part of the results were published and we contacted the authors to obtain the (unpublished) research protocol and additional data.

We included ten of the 14 studies in our quantitative meta-analyses on the treatment of depression in which outcomes were expressed as the reduction of depressive symptoms. One treatment trial was excluded from this analysis because the authors did not report post–treatment data but only long-term follow-up. The three remaining trials were aimed at prevention of depression, either as indicated prevention [6,14] or as relapse prevention [1] with the incidence of depressive disorders as the main outcomes. Given that it is not useful to pool results from treatment and prevention we excluded the prevention trials from our quantitative meta-analyses.

Characteristics of the 14 included treatment and prevention studies

The 14 studies included a total of 5,194 patients of whom 2,560 were randomized to stepped care and 2,634 to a control condition. For the ten studies included in the quantitative meta-analyses the total number of included patients is 4,580 with 2,243 in the stepped care arms and 2,337 in the control conditions (Table 1).

Twelve trials were patient-randomised [1-8,10, 12-14], and two were cluster-randomised [9, 11]. Six trials were conducted in the US [4-5,7-8,10,13], six in The Netherlands [1,3,6,9,12,14], one in Chile [2] and one in India [11]. Participants were recruited mainly from primary care [2,9-11,12-14], or secondary care [3-5,7]. All studies compared stepped care to usual care, either standard [1-6,9-10,12-14] or 'enhanced' [7-8,11].

Five of the treatment trials [3-5,8,10] included patients scoring above a cut-off on a self-rated depression questionnaire only (two also used the core symptoms of MDD) while five others [6,9,11-13] performed diagnostic interviews to include patients with MDD (one also included minor depression, and two also included dysthymia). The three prevention trials [1,6,14] used a diagnostic interview to exclude patients with MDD. Six of the studies were aimed at depressive symptoms among patients with either co-morbid acute coronary syndrome [4-5], cancer [7] or diabetes mellitus [3,8,10] and five trials, including the three prevention studies, were specifically aimed at older adults [1,3,6,13,14].

Characteristics of the stepped care interventions

We found considerable between study heterogeneity in numbers of steps (two, three or four), types of treatments offered at each step, and duration of the total intervention (between three and 12 months; table 2).

Seven studies [4, 5,7-10,13], six of which were US trials, were based on the 'IMPACT' model and used Problem Solving Treatment (PST) and antidepressant medication (ADM) as the core of the intervention. The IMPACT intervention is primarily a collaborative intervention in which a dedicated team works together to provide optimal depression care, meeting our inclusion criteria as a stepped care approach because patients were evaluated at predetermined time intervals according to defined improvement criteria and care was adjusted or augmented if the patient did not improve sufficiently. Treatments were provided according to patients' needs and preferences. In all seven 'IMPACT' studies and one other involving both psychological treatment (psycho-education) and ADM [2], there was no progression of increasing therapeutic intensity.

In contrast, care was delivered in the other six trials [1,3, 6,11-12,14] through steps of increasing intensity. Five of the six studies started with watchful waiting although two studies [12,14] only included patients after the watchful waiting period while the other three [1,3,6] included watchful waiting as part of their stepped care model. The first therapeutic component included psycho-education or bibliotherapy alone or combined, offered either as self-help (with online, telephone or face-to-face

support), in a group, or as individual sessions. The next step in these six studies varied widely and included psychological therapy (CBT, life review, IPT, PST, Coping with Depression Course) [1,3,6,12,14] or a psychological therapy (IPT) combined with ADM[11]. The last step typically consisted of referral to specialists, a GP or mental health services. Only two of those six studies which used steps of increasing intensity are included in the quantitative meta-analysis [11,12]. As mentioned above, one study was excluded because of unavailability of post-test data [3], and the three other trials were aimed at (relapse) prevention [1,6,14].

In twelve studies more than one healthcare professional was involved in stepped care [1-2,4-13] including nurses [1-2,4-6,10,12-13], psychiatrists [4-5,7-11,13], General Practitioners [2,5,8,9,11,13], social workers [2,4,7-8], psychologists [4-5,12-13] and relatively less qualified staff (residential home staff [6], an assistant patient navigator [8], lay health counselor [11] and study researcher [1]). In two studies, treatment was provided by one healthcare professional: a nurse or psychologist [3] or a nurse only [14]. No details are available for external professionals providing treatment after referral outside the core stepped care team.

Patient progress was assessed using one [1-7,9-11,13-14], two [8] or three [12] self-rated instruments. In five studies the decision to 'step up' was contingent on patients' score relative to a specific cut-off on the HDRS [2], CES-D [1,14], PHQ-9 [7] or HADS, IDS and WSAS [12]. In five studies the decision to 'step up' was dependent on improvement (relative to baseline or the last assessment) on the PHQ-9 [4-5,10,13] or CES-D [6]. Three studies used a combination of improvement and a specific cut-off on the CES-D [3], PHQ-9 [9] or PHQ and SCL [8]. In one study [11] improvement was assessed by health counselors following application of the GHQ with no further detail specified.

Quality of the included studies

In one study [3] we rated all quality criteria as either unclear or at high risk of bias and in a second [1] we rated five of the six criteria as unclear or at high risk of bias. For the remaining twelve studies quality on most criteria was high. The description of randomisation sequence generation was adequate but four of these twelve studies did not clearly report methods of allocation concealment [4,10,11,14]. No studies were able to blind patients or clinicians but all studies used assessors to measure outcomes who were unaware of the randomisation status of the patients or used self-report. Post-intervention study drop-out ranged between 8.0% [5] and 49.6% [3] and one study [9] was rated at high risk of bias with respect to handling incomplete outcome data. All studies used intention-to-treat analyses. Three of the twelve

studies were at high risk of other biases because of the potential for contamination between trial arms [6,8,13] or because patients were recruited in different ways in the intervention and control groups [9].

Effects of stepped care

Most of the studies used more than one depression outcome measure so we averaged the between group differences from the various measures as a single combined measures effect size for each study (Table 3). We found an overall post-intervention effect size of *d* = 0.38 (95% CI 0.18 to 0.57). We also examined the post-test effect sizes from the measure with the highest effect size for each study (*d* = 0.42; 95% CI 0.22 to 0.62) and repeated this with the measure producing the lowest effect sizes (*d* = 0.33; 95% CI 0.13 to 0.52). All effect sizes were significantly in favour of stepped care.

The stepped care interventions varied in duration between three and 12 months. We used the combined measures effect size to examine outcomes at different time points. The effects were *d* = 0.57 at two to four months (95% CU 0.21 to 0.94), *d*=0.34 at six months (95% CI 0.20 to 0.48), *d*=0.43 at nine to 12 months (95% CI 0.20 to 0.65), and *d* = 0.26 at 18 months (one study only). All effects were significantly in favour of the stepped care intervention with the exception of the 18 month result. Heterogeneity, as indicated by I^2, was high for the post-intervention effect sizes as well as for the effect sizes at the different time points. From Figure 2 it can be observed how the six month effect sizes varied between the different studies. To examine this heterogeneity we performed subgroup analyses.

Subgroup analysis and publication bias

We analysed the association of the six months outcomes (overall *d* = 0.34; Table 3) with the following variables: country in which the study was performed (USA, Netherlands, or other), treatment based on IMPACT protocol (yes or no), stepped care treatment using progressive intensity (yes or no), physical health comorbidity (present or absent), and diagnostic status at inclusion (diagnosis assessed or not). The effect of the eight studies on stepped care models without progressive intensity was significantly higher (*d*=0.41) than those of the two studies examining stepped care models with progressive intensity (*d*=0.07; $p < 0.01$). None of the remaining variables were significantly related to the effect size. Even though not statistically significant (p=0.63) the effect size for the two Dutch studies was lower (*d*=0.18) than for those conducted in the USA (*d*=0.38) or other countries (*d*=0.44).

We found no indication of publication bias in our funnel plot on the six month outcomes or in Duval and Tweedie's trim and fill procedure. No studies needed to be imputed.

Effects of stepped care intervention for depression: four studies excluded from the quantitative analyses

The treatment study of Bot [3] only provided two year follow-up data for the complete cases (49.6%) and reported no difference between the groups (*d*=-0.12; 95% CI -0.62 to 0.39). Both of the trials on indicated prevention showed results in favour of stepped care [6, 14]. One [6] demonstrated 12 month MDD rates of 6.5% in the intervention group and 14.1% in the control group (Incidence Rate Ratio = 0.46; 95% CI 0.17 to 1.21). The other [14] demonstrated 12 month prevalence rates of combined MDD and anxiety disorders of 11.6% in the intervention group and 23.8% in the control group (Incidence Rate Ratio = 0.49; 95% CI 0.24 to 0.98). The pooled rate ratio of the two studies was 0.48 (95% CI 0.27 to 0.83; $I^2 = 0$). The study on relapse prevention [1] reported no difference in the 12 month MDD incidence rate between stepped care and care-as-usual.

Discussion

We identified 14 trials on stepped care for depression, ten of which could be used in a meta-analysis of treatment outcomes. Stepped care has a moderate effect on depression (*d*=0.34 at six months and *d*=0.38 post-intervention). Stepped care interventions based on progressive treatment intensity performed worse (n=2; *d*=0.07) than those without a clear intensity order (n=8; *d*=0.41; $p < 0.01$). Most trials were of good quality. The stepped care interventions were extremely heterogeneous with different numbers of steps, different treatment components, different duration of the steps, different rules about stepping up and different professionals involved.

Even though we demonstrated that stepped care is effective, the effect sizes were modest. Meta-analyses have demonstrated higher effect sizes (Cohen's *d* between 0.42 and 0.88) for self-help interventions which are usually considered as a first step in stepped care (Gellatly *et al*. 2007; Andrews *et al.* 2010; Richards & Richardson 2012, Bower *et al.* 2013). However, the majority of the trials on self-help have been performed in population samples rather than in clinical samples. Even though baseline severity of symptoms do not seem to be associated with the effect of self-help interventions (Bower *et al*. 2013) there might be other differences between clinical and population samples which might account for differences in effects.

The stepped care six month effect size (*d*=0.34) was similar to the one found in the Cochrane review on collaborative care (Archer *et al.* 2012). (Collaborative care may include a broad range of

interventions, settings and providers; defining characteristics are that a team of health care professionals are responsible for providing the 'right' care at the 'right' time and that there is a structured management plan which includes scheduled patient follow-ups (Bower et al. 2006; Gunn et al. 2006).) This finding may not be suprising given that six out of the ten studies [2,7 8,10,11,13] included in our meta-analysis were also included in the meta-analysis of collaborative care.

In stepped care the primary focus is on psychological interventions of different intensity. However, as we noted in our introduction, it is unclear how medication management, which might be offered with significant support from case managers, fits into stepped care programs. Since medication management is an important treatment option in depression care, we decided to include it in our definition of stepped care (the availability of more than one treatment modality, medication and psychotherapy). This choice led to the inclusion of several of the collaborative care trials, albeit the majority of which were also described as stepped care [2, 7, 8, 10, 11], and three other studies [4, 5, 9] in which stepped care was not defined by a progressive increase in treatment intensity. Our definition is debatable: others may choose to review or conduct future research on stepped care in line with how it was originally conceived; findings based on one definition of stepped care may not generalize to the other; future research may be required to compare stepped care defined by a progressive increase in treatment intensity and stepped care that is not.

We compared the results of the eight studies without a hierarchy in treatment intensity with the two studies which did provide 'true' stepped care with increasing treatment intensity. This comparison demonstrated that the 'true' stepped care studies performed significantly worse. This indicates that it might be better to match the first treatment to the patient's need than to offer a low intensity treatment regardless of the patient's clinical profile. However, we think that this conclusion would be premature. First, because the results of 'true' stepped care are based on two studies only. Second, because seven of the eight studies without increasing intensity were based on the IMPACT protocol. Those seven IMPACT studies did not show better results than the three non-IMPACT studies. In other words, the difference in results between the two subgroup analyses (IMPACT vs. non-IMPACT, and increasing intensity vs. no increasing intensity) was actually based on one study with a very high effect size [2]. Third, because the two studies aiming at prevention of (indicated) depression both offered 'true' stepped care and they demonstrated very large effects (almost halving the incidence of depression). In conclusion we think that more 'true' stepped care studies need to be performed before we can reach a definite conclusion. Moreover, it is important not only to look at treatment studies but also prevention studies especially as

it has been argued that prevention contributes most in reducing the global burden of depression (Cuijpers *et al.* 2012). This and other key areas for future research are summarised in Box 1.

The central tenet of stepped care is that for many patients the first (low intensity) treatments are sufficient and relatively few patients need to step up. This means that similar (or better) patient outcomes could be achieved against lower costs. In the current meta-analyses only a limited number of trials provided data on the proportion of patients recovered after the first treatment. The data that was available was hard to interpret since the definition of adequate recovery varied between the studies as well as the duration of the steps, the number of patients dropping out of treatment and the number of patients not reporting health status. We also do not know how many patients needed to step up or the actual percentage of patients who took up this second step. This is important information because within stepped care there is a risk that patients do not start a second higher intensity treatment after failure of the first. To improve reporting on clinical trials of stepped care for depression, we identify data that are important to include (Box 2); including this would maximize subsequent systematic reviews.

We did demonstrate that better outcomes were reached in stepped care compared to care-as-usual. However, the question is whether or not care-as-usual is the best comparator. One could argue that care-as-usual is similar to matched care since this is the current dominant treatment approach. However, all the trials used an active approach to find and select patients. In four trials it was reported that the GP was informed about the diagnostic status of the patients in the control group, while the other studies refrained from informing the GP or did not report how they handled this. This indicates that care-as-usual probably more closely resembled 'no care'. In other words we demonstrated that stepped care is better than doing nothing. The ideal test, against true matched care or against high intensity care for all patients, has not been performed yet. We identified five (Dutch) protocol manuscripts on stepped care (Braamse *et al*. 2010; Krebber *et al*. 2012; Pommer *et al*. 2012; Van Dijk *et al*. 2012; Van der Weele *et al*. 2012); none compare stepped with matched care or with intensive psychological treatment for all.

The remaining assumption of stepped care is that it reduces health care costs. Six out of the ten studies included in the meta-analyses published a separate paper on the cost-effectiveness of their (collaborative) stepped care program (Katon et al. 2005; Araya et al. 2006; Simon et al. 2007; Van 't Veer-Tazelaar et al. 2010; Butorff et al. 2012; Hay et al. 2012; Ladapo et al. 2012). The results of the studies performed in Chile and India are hard to generalize to the Western world. The remaining four (US) papers either report savings or incremental costs which are offset by the health gains. This means that

there is an indication that stepped care interventions might indeed be more cost-effective. However, because stepped care has not been compared to either matched care or high intensity care, final conclusions about cost-effectiveness cannot be made.

Our study has several limitations. First is the limited number of studies. This made it especially hard to perform subgroup analyses. In this respect, the five protocol manuscripts on stepped care are relevant, indicating that there is considerable clinical trials work in progress. Second, the stepped care interventions varied a lot as well as the samples included in the studies (countries, with or without comorbidity, age, definitions of depression etc). This may limit the generalizability of our findings. A strength of this study is that it is the first to systematically describe all the available evidence with respect to stepped care which is regarded in many countries as the preferred way to offer depression care. Furthermore, most of the studies were of good quality.

Although many guidelines recommend stepped care, there is currently only limited evidence to suggest it should be the dominant model of treatment organization compared to alternative systems. Consistent with a previous observational study (Richards *et al*. 2012), we found considerable variety in the implementation of stepped care (with respect to the number and duration of treatment steps, treatments offered, professionals involved and criteria to step up) and only one significant difference between subgroups of studies (progressive intensity, yes/no) which requires further research. Hence, it was not possible to identify any optimal component of stepped care or to suggest a preferred model for delivery which may be associated with increased effectiveness. It was also not possible to determine with any certainty the relative effectiveness of stepped care models defined by combined treatment modalities (psychological and pharmacological) compared to those defined by progressive intensity of psychological treatment. The balance of costs, effectiveness and acceptability has not been investigated and further research is needed to determine if stepped care really should have such prominence in treatment guidelines. The first stage of such a research programme should be a fully powered clinical trial of stepped psychological versus high-intensity treatment to test both the non-inferiority hypothesis and the potential cost advantages of stepped versus more intensive treatment.

Declaration of interests

None

Author contributions

AvS and DR were jointly responsible for conception, AvS, JJH and DR for design, study identification and selection, and data extraction. AvS was mainly responsible for meta-analysis, JJH for the narrative synthesis and description of included studies. All authors' reviewed and reflected on results and contributed to writing; writing was led by AVS and JJH and JJH was mainly responsible for response to peer review. All authors approved the manuscript for publication.

References

Andrews G, Tolkein II Team (2006). A needs-based, costed stepped-care model for mental health services. Sydney: CRUFAD, University of New South Wales.

Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N (2010). Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. P*LoS One*, 13, 5(10), e13196.

Apil SRA, Hoencamp E, Haffmans PMJ, Spinhoven P (2012). A stepped care relapse prevention program for depression in older people: a randomized controlled trial. *International Journal of Geriatratric Psychiatry,* 27**,** 583-91.

Araya R, Rojas G, Fritsch R, Gaete J, Rojas M, Simon G, Peters TJ (2003). Treating depression in primary care in low-income women in Santiago, Chile: a randomised controlled trial. *The Lancet,* 361**,** 995-1000.

Araya R, Flynn T, Rojas G, Fritsch R, Simon G (2006). Cost-Effectiveness of a Primary Care Treatment Program for Depression in Low-Income Women in Santiago, Chile. *American Journal of Psychiatry,* 163, 1379–1387.

Archer J, Bower P, Gilbody S, Lovell K, Richards D, Gask L, Dickens C, Coventry P (2012). Collaborative care for depression and anxiety problems. *Cochrane Database of Systematic Reviews*.

Bennett-Levy J, Richards DA, Farrand P (2010). Low intensity CBT interventions: a revolution in mental health care. In *Low Intensity CBT Interventions* (ed. J. Bennett-Levy, D.A. Richards, P. Farrand, H. Christensen, K.M. Griffiths, D.J. Kavanagh, B. Klein, M.A. Lau, J. Proudfoot, L. Ritterband, J. White and C. Williams), pp. 3-18. Oxfor University Press: Oxford.

Bijl RV, de Graaf R, Hiripi E, Kessler RC, Kohn R, Offord D R, Ustun TB, Vicente B, Vollebergh WAM, Walters EE, Wittchen HU(2003)The prevalence of treated and untreated mental disorders in five countries. *Health Affairs,* 22**,** 122-133.

Bot M, Pouwer F, Ormel J, Slaets JP, de Jonge P (2010). Predictors of incident major depression in diabetic outpatients with subthreshold depression. *Diabetic Medicine,* 27**,** 1295-301.

Bower P, Gilbody S (2005). Stepped care in psychological therapies: access, effectiveness and efficiency - Narrative literature review. *British Journal of Psychiatry,* 186**,** 11-17.

Bower P, Gilbody S, Richards D, Fletcher J, Sutton A (2006). Collaborative care for depression in primary care - Making sense of a complex intervention: systematic review and meta-regression. *British Journal of Psychiatry,* 189**,** 484-493.

Bower P, Kontopantelis E, Sutton A, Kendrick T, Richards DA, Gilbody S, Knowles S, Cuijpers P, Andersson G, Christensen H, Meyer B, Huibers M, Smit F, van Straten A, Warmerdam L, Barkham M, Bilich L, Lovell K, Liu ET (2013). Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. *British Medical Journal*,26, 346, f540.

Braamse AMJ, van Meijel B, Visser O, van Oppen P, Boenink AD, EeltinkC, Cuijpers P, Huijgens PC, Beekman ATF, Dekker J (2010). Distress and quality of life after autologous stem cell transplantation: a randomized clinical trial to evaluate the outcome of a web-based stepped care intervention. *BMC Cancer*, 10, 361.

Buttorff C, Hock RS, Weiss HA, Naik S, Araya R, Kirkwood BR, Chisholm D, Patel V (2012). Economic evaluation of a task-shifting intervention for common mental disorders in India. *Bulletin World Health Organization,* 90, 813–821.

Clark DM (2011). Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: The IAPT experience. *International Review of Psychiatry,* 23**,** 318-327.

Cooper H, Hedges LV (1994). *The handbook of research synthesis.,* New York, Russell Sage Foundation Press.

Cuijpers P, Donker T, van Straten A, Li J, Andersson G (2010) Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychological Medicine*, 40, 1943–1957.

Cuijpers P, Beekman ATF, Reynolds CF III (2012). Preventing depression. A global priority. *Journal of the American Medical Association,* 307, 1033-1034.

Davidson KW, Rieckmann N, Clemow L, Schwartz JE, Shimbdo D, Medina V, Albanese G, Kronish I, Hegel M, Burg MM (2010). Enhanced depression care for patients with Acute Coronary Syndrome and persistent depressive symptoms. *Archives of Internal Medicine,* 170**,** 600-608.

Davidson K W, Bigger JT, Burg MM, Carney RM, Chaplin WF, Czajkowski S, Dornelas E, Duer-Hefele J, Frasure-Smith N, Freedland KE, Haas DC, Jaffe AS, Ladapo JA, Lesperance F, Medina V, Newman JD, Osorio GA, Parsons F, Schwartz JE, Shaffer JA, Shapiro PA, Shep DS, Vaccarino V, Whang GW, Ye S (2013). Centralized, stepped, patient preference-based treatment for patients with post- Acute Coronary Syndrome depression: CODIACS Vanguard randomized controlled trial. *JAMA Internal Medicine,* 173**,** 997-1004.

Dozeman E, van Marwijk HW, van Schaik DJ, Smit F, Stek ML, van der Horst HE, Bohlmeijer ET, Beekman AT (2012). Contradictory effects for prevention of depression and anxiety in residents in homes for the elderly: a pragmatic randomized controlled trial. *International Psychogeriatrics,* 24**,** 1242-51.

Duval S, Tweedie R (2000). Trimm and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics,* 56, 455–463.

Ell K, Xie B, Quon B, Quinn DI, Dwight-Johnson M, Lee PJ (2008). Randomized controlled trial of collaborative care management of depression among low-income patients with cancer. *Journal of Clinical Oncology,* 26**,** 4488-96.

Ell K, Katon W, Xie B, Lee PJ, Kaoetanovic S, Guterman J, Chou CP (2010). Collaborative care management of major depression among low-income, predominantly Hispanic subjects with diabetes: a randomized controlled trial. *Diabetes Care,* 33**,** 706-13.

Gellatly J, Bower P, Hennessy S, Richards D, GilbILBdy S, Lovell K (2007). What makes self-help interventions effective in the management of depressive symptoms? Meta-analysis and meta-regression. *Psychological Medicine,* 37**,** 1217-28.

Gilbody S (2006). Is collaborative care cost effective? *Journal of Affective Disorders.,* 91**,** S4-S5.

Gunn J, Diggens J, Hegarty K, Blashki G (2006). A systematic review of complex system interventions designed to increase recovery from depression in primary care. *BMC Health Services Research.,* 6.

Hay JW, Katon WJ, Ell K, Lee P-J, Guterman JJ (2012). Cost effectiveness analysis of collaborative care management of major depression among low-income, predominantly Hispanics with diabetes. *Value Health*, 15(2), 249–254.

Hedges LV. Olkin I (1985). *Statistical methods for meta-analysis.* Academic Press: San Diego.

Huijbregts KM, de Jong FJ, van Marwijk HW, Beekman AT, Ader HJ, Hakkaart-van Roijen L, Unutzer J, van der Feltz-Cornelis CM (2013). A target-driven collaborative care model for Major Depressive Disorder is effective in primary care in the Netherlands. A randomized clinical trial from the depression initiative. *Journal of Affective Disorders,* 146**,** 328-37.

Katon W, von Korff M, Lin E, Walker E, Simon GE, ,Bush T, Robinson P, Russo J (1995). Collaborative management to achieve treatment guidelines. Impact on depression in primary care. *Journal of the American Medical Association,* 273**,** 1026-31.

Katon WJ, von Korff M, Lin EHB, Simon G, Ludman E, Russo J, Ciechanowski P, Walker E, Bush T (2004). The Pathways Study. A randomized trial of collaborative care in patients with diabetes and depression. *Archives of General Psychiatry,* 61**,** 1042-1049.

Katon WJ, Schoenbaum M, Fan M-Y, Callahan CM, Williams J, Hunkeler E, Harpole L, Zhou X-HA, Langston C, Unutzer J (2005). Cost-effectiveness of Improving Primary Care Treatment of Late-Life Depression. *Archives of General Psychiatry*, 62, 1313-1320.

Katon W, Unutzer J (2006). Collaborative care models for depression - Time to move from evidence to practice. *Archives of Internal Medicine,* 166**,** 2304-2306.

Krebber A-MH, Leemans CR, de Bree R, van Straten A, Smit F, Smit EF, Becker A, Eeckhout GM, Beekman ATF, Cuijpers P, Verdonck-de Leeuw IM (2012). Stepped care targeting psychological distress in head and neck and lung cancer patients: a randomized clinical trial. *BMC Cancer,* 12, 173.

Ladapo JA, Shaffer JA, Fang Y, Ye S, Davidson KW (2012). Cost-effectiveness of Enhanced Depression Care After Acute Coronary Syndrome: Results From the Coronary Psychosocial Evaluation Studies Randomized Controlled Trial. *Archives of Internal Medicine*, 172, 1682-1683.

Lipsey MW (1990). *Design sensitivity. Statistical power for experimental research.*SAGE publications*:* Newbury Park, California.

Lipsey MW, Wilson DB (1993). The efficacy of psychological, educational and behavioural treatment. Confirmation from meta-analysis. *American Psychologist,* 48**,** 1181-1209.

Lovell K, Richards D (2000). Multiple Access Points and Levels of Entry (MAPLE): ensuring choice, accessibility and equity for CBT services. *Behavioural and Cognitive Psychotherapy,* 28**,** 379-391.

National Collaborating Centre for Mental Health (2010). *Depression: The NICE guideline on the treatment and management of depression in adults. Updated edition.* The British Psychological Society and The Royal College of Psychiatrists: London.

National Institute for Health and Clinical Excellence (2009). Depression in adults. The treatment and management of depression in adults. NICE clinical guideline 90.

National Institute for Health and Clinical Excellence (2011). Common Mental Health disorders: Identification and Pathways to Care - NICE guidance (Clinical Guidelines CG123).

Patel V, Weiss HA, Neerja C, Smita N, Sulochana P, Chatterjee S, de Silva MJ, Bhat B, Araya R, King M, Simon G, Verdeli H, Kirkwood BR (2010). Effectiveness of an intervention led by lay health counsellors for depressive and anxiety disorders in primary care in Goa, India (MANAS): a cluster randomised controlled trial. *The Lancet,* 376**,** 2086-20895.

Piek E, van der Meer K, Penninx BWJH, Verhaak PFM, Nolen WA (2011). Referral of patients with depression to mental health care by Dutch general practitioners: an observational study. *BMC Family Practice,* 12.

Piek E, Nolen WA, van der Meer K, Joling KJ, Kollen BJ, Penninx BWJH, van Marwijk HWJ, van Hout HPJ (2012). Determinants of (non-)recognition of depression by general practitioners: Results of the Netherlands study of depression and anxiety. *Journal of Affective Disorders,* 138**,** 397-404.

Pommer AM, Pouwer F, Denollet J, Pop VJM (2012). Managing co-morbid depression and anxiety in primary care patients with asthma and/or chronic obstructive pulmonary disease: study protocol for a randomized controlled trial. *Trials*, 13, 6.

Richards D, Richardson T (2012). Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clinical Psychology Review*, 32, 329-342.

Richards, D.A., Bower P., Pagel, C., Weaver, A., Utley, M., Cape, J., Pilling, S., Lovell, K., Gilbody, S., Leibowitz, J., Owens, L., Paxton, R., Hennessy, S., Simpson, S., Gallivan, S., Tomson, D. and Vasilakis, C. (2012). Delivering stepped care: an analysis of implementation in routine practice, *Implementation Science,* 7:3.

Seekles W, van Straten A, Beekman A, van Marwijk H, Cuijpers P (2011). Stepped care treatment for depression and anxiety in primary care. a randomized controlled trial. *Trials,* 12**,** 171.

Simon GW (2002). Evidence review: efficacy and effectiveness of antidepressant treatment in primary care. *General Hospital Psychiatry,* 24**,** 213-224.

Simon GE, Katon WJ, Lin EHB, Rutter C, Manning WG, Von Korff M, Ciechanoswki P, Ludman EJ, Young BA (2007). Cost-effectiveness of Systematic Depression Treatment Among People With Diabetes Mellitus. *Archives of General Psychiatry*, 64, 65-72.

Spijker J, van Vliet IM,Meeuwissen JAC, van Balkom AJLM (2010). Update of the multidisciplinary guidelines for anxiety and depression. *Tijdschrift voor psychiatrie,* 52**,** 715-8.

The Cochrane Collaboration (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011)*, Wiley-Blackwell.

Unutzer J, Katon W, Callahan CM, Williams JW, Hunkeler E, Harpole L, Hoffing M, Della Penna RD, Noel PH, Lin EHB, Arean PA, Hegel M, Tang LQ, Belin TR, Oishi S, Langston C, et al (2002). Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *Journal of the American Medical Association,* 288**,** 2836-2845.

Van der Weele GM, dAJM, de Waal MWM, van den Hout WB, de Craen AJM, Spinhoven P, Stijnen T, Assendelft WJJ, van der Mast RC, Gussekloo J (2012). Effects of a stepped-care intervention programme among older subjects who screened positive for depressive symptoms in general practice: the PROMODE randomized controlled trial. *Age and Ageing*, 41, 482–488.

Van Dijk SEM, Pols AD, Adriaanse MC, Bosmans JE, Elders PJM, van Marwijk HWJ, van Tulder MW (2013). Cost-effectiveness of a stepped-care intervention to prevent major depression in patients with type 2 diabetes mellitus and/or coronary heart disease and subthreshold depression: design of a cluster-randomized controlled trial. *BMC Psychiatry*, 13, 128.

Van 't Veer-Tazelaar PJ, van Marwijk HWJ, van Oppen P, van Hout HPJ, van der Horst HE, Cuijpers P, Smit F, Beekman ATF (2009). Stepped-care prevention of anxiety and depression in late life: a randomized controlled trial. *Archives of General Psychiatry,* 66**,** 297-304.

Van 't Veer-Tazelaar PJ, Smit F, van Hout HPJ, van Oppen P, van der Horst HE, Beekman ATF, van Marwijk HWJ (2010). Cost-effectiveness of a stepped care intervention to prevent depression and anxiety in late life: randomised trial*. British Journal of Psychiatry*, 196:319-325.

Van Steenbergen-Weijenberg KM, van der Feltz-Cornelis CM, Horn EK, Van Marwijk HWJ, Beekman ATF, Rutten FFH, Hakkaart-van Roijen L (2010). Cost-effectiveness of collaborative care for the treatment of major depressive disorder in primary care. A systematic review. *BMC Health Services Research,* 10.

Figure 1
Click here to download Figure(s): Figure 1 flow chart.doc

Figure 1. Flow-chart of studies included in the meta-analysis on stepped care for depression

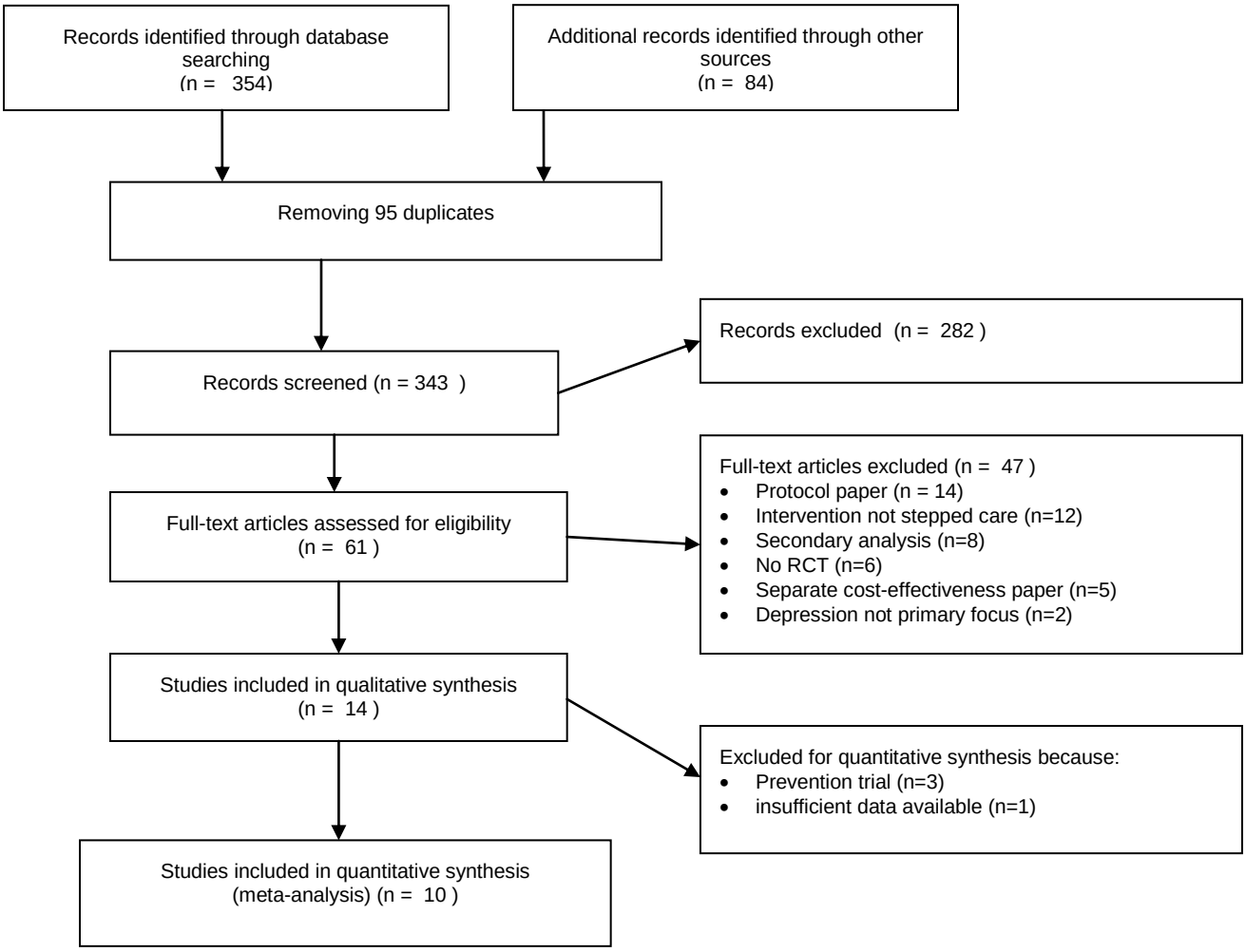

Figure 2
Click here to download Figure(s): Figure 2 plot ES 6 months.docx

Figure 2. Effects of stepped care versus care-as-usual (6 months outcomes)

Study name	Outcome	Time point	Statistics for each study		
			Std diff in means	Lower limit	Upper limit
Araya, 2003	Combined	6 months	0,839	0,530	1,148
Davidson, 2010	BDI	6 months	0,485	0,168	0,803
Davidson, 2013	Combined	6 months	0,443	0,057	0,828
Ell, 2008	Combined	6 months	0,196	-0,041	0,433
Ell, 2010	Combined	6 months	0,369	0,108	0,630
Huijbregts, 2013	Combined	6 months	0,455	-0,120	1,031
Katon, 2004	Combined	6 months	0,238	-0,030	0,506
Patel, 2010	Combined	6 months	0,073	-0,091	0,237
Seekles, 2011	IDS	6 months	0,031	-0,326	0,389
Unutzer, 2002	Combined	6 months	0,419	0,309	0,528
			0,339	0,197	0,481

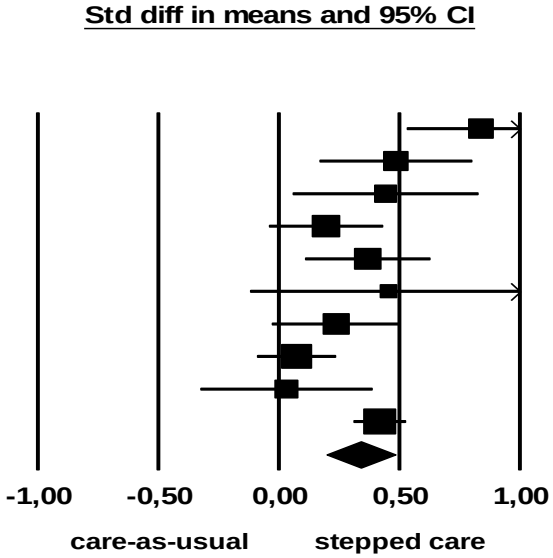

Table 1
Click here to download Table(s): Table 1 characteristics studies.docx

Table 1. Characteristics of randomized controlled trials comparing stepped care for depression with usual care

ID	Author	Year	Country	Random level	Target of the trial	Control condition	Depression criteria	Comorbid disorder	Age (years)	IMPACT based	Total N (EXP / CTRL)
1	Apil	2012	Nether lands	Patient	Prevention	Usual care: depressive symptoms monitored.	Not depressed (MINI)	-	55+	No	136[1] (74/62)
2	Araya	2003	Chile	Patient	Treatment	Usual care: GPs given guidelines on depression treatment.	MDD (MINI)	-	18-70	No	240 (120/120)
3	Bot	2010	Nether lands	Patient	Treatment	Usual care: ADs or psychotherapy were available.	Depressive symptoms (CES-D ≥ 16)	Diabetes	55+	No	123[1] (64 / 59)
4	Davidson	2010	USA	Patient	Treatment	Usual care: physicians informed of patients' depressive symptoms/ MDD criteria.	Persistent depressive symptoms (BDI ≥10 and < 45 at week 1 and 13)	Acute Coronary Syndrome	NS[3]	Yes	157 (80/77)
5	Davidson	2013	USA	Patient	Treatment	Usual care: PCPs and/or cardiologists informed of patients' depressive symptoms.	Depressive symptoms (BDI ≥ 10 on 2 occasions or ≥ 15 on 1 occasion, 2 to 6 months after hospitalization for ACS)	Acute Coronary Syndrome	35+	Yes	150 (73/77)
6	Dozeman	2012	Nether lands	Patient	Prevention	Usual care [4]	Depressive symptoms (CES-D ≥ 8), no MDD (MINI)	-	Elderly in residential homes	No	185[1] (93/92)
7	Ell	2008	USA	Patient	Treatment	Enhanced usual care: patient/family depression and cancer educational pamphlets + resource list. [5]	1 or 2 core depressive symptoms, and PHQ ≥ 10, and/or 2 questions from the SCID indicating dysthymia	Cancer	18+	Yes	472 (242/230)
8	Ell	2010	USA	Patient	Treatment	Enhanced usual care: depression educational pamphlets + resource list; PCPs informed of patient depression diagnoses.	Depressive symptoms (PHQ ≥ 10 and 1 or 2 core symptoms)	Diabetes	18+	Yes	387 (193/194)

9	Huijbregts	2013	Nether lands	Cluster	Treatment	Usual care: patients informed of diagnosis and advised to consult GP.	MDD (MINI) and PHQ ≥ 10	-	18+	Yes	150 (101/49)
10	Katon	2004	USA	Patient	Treatment	Usual care: patients advised to consult PCP.	Persistent depressive symptoms (PHQ ≥ 10 and mean SCL ≥ 1.1 at 2 weeks)	Diabetes	NS[3]	Yes	329 (164/165)
11	Patel	2010	India	Cluster	Treatment	Enhanced usual care: physicians & patients given screening results and a treatment manual.	MDD (CIS-R) and GHQ > 5	-	18+	No	774[2] (304/470)
12	Seekles	2011	Nether lands	Patient	Treatment	Usual care: patients advised to consult GP	Persistent depressive symptoms (K10 ≥ 21 at week 1 and 4), MDD, dysthymia, minor depression (CIDI)	-	18-65	No	120 (60/60)
13	Unutzer	2002	USA	Patient	Treatment	Usual care [4]	MDD or dysthymia (SCID)	-	60+	Yes	1801 (906/895)
14	Van 't Veer	2009	Nether lands	Patient	Prevention	Usual care [4]	Persistent depressive symptoms (CES-D ≥ 16 at week 1 and 13), no MDD or anxiety disorder (MINI)	-	75+	No	170[1] (86/84)

Notes: [1] not included in quantitative meta-analysis; [2] total N in this trial is 2796 but we only used the depressed subsample in our meta-analysis; [3] age in- and exclusion criteria 'not specified'; [4] no particular feature of usual care described; [5] oncologists may have attended a depression treatment didactic session by the study psychiatrist at the start of the study and yearly after and may have been informed of patients' depression status although it is unclear whether these features applied to patients in the Enhanced Usual Care group.

Abbreviations: GP – General Practitioner; PCP Primary Care Physician; MDD – Major Depressive Disorder; ACS – Acute Coronary Syndrome; GHQ – General Health Questionnaire; other abbreviations refer to depressive symptom checklists (CES-D, BDI, PHQ, SCL) and diagnostic interviews for depression (MINI, CIS-R, SCID, CIDI).

Table 2

Click here to download Table(s): Table 2 characteristics interventions.docx

Table 2. Characteristics of the stepped care interventions for depression

ID	Author	N step	Step 1	Step 2	Step 3	Step 4	Providers[1]	Stepping up rules	Total duration
1	Apil	4	Watchful waiting (1 phone call)	Bibliotherapy based on CWD (3 phone calls)	Individual CWD course (12 sessions)	Referral to a GP or psychotherapist	Nurse, Researcher	CES-D >=16 at 6 weeks, 3 months and 6 months	6 months
2	Araya	2	PE group (9 sessions) + self-help book. If HDRS > 19 also structured ADs	Initiating or adjusting ADs	-		Social workers, nurses, GP	HDRS > 12 at 6 weeks	3 months
3	Bot	4	Watchful waiting + 3 phone	Bibliotherapy based on CWD (3 phone calls)	CBT: 4 modules of CWD course (5 sessions)	Referral to psychiatrist	Prevention worker (nurse or psychologist)	CES-D improvement < 5 or CES-D ≥ 16 at 6, 12 & 24 weeks	36 weeks ≈ 8 months
4	Davidson 2010	3	PST (no predetermined number of sessions) or ADs (patient preference)	Switching treatments, adding treatments, intensifying original treatment (patient preference)	Referral to usual care provider	-	nurse, psychologist, social worker, psychiatrist	Initial PHQ9 5-10 and improvement < 30%; initial PHQ9 11-20 and improvement < 50%; initial PHQ9 > 20 and improvement < 60%. Assessed every 8 weeks.	6 months
5	Davidson 2013	4	PST (number of sessions not specified) and/or Ads, or neither	Switching treatments, adding treatments (patient preference)	Switching treatments, adding treatments (patient preference)	Switching treatments, adding treatments (patient preference)	PST therapist, psychiatrist, clinical psychologist, GP or advanced practice nurse	See Davidson 2010. Assessed every 6-8 weeks.	6 months
6	Dozeman	4	Watchful waiting	Bibliotherapy based on CWD (face-to-face guidance; no predetermined number of sessions)	Individual face-to-face Life Review (no predetermined number of sessions) + advise to consult GP	If CESD ≥ 16: advised to consult GP or referral to mental health specialist	Residential home staff, mental health nurses	CES-D improvement < 5 at 1 & then every 3 months.	10 months
7	Ell 2008	3	1 visit CDCS then PST (8 to 12 sessions) and/ or ADs (patient preference)	ADs and additional psychotropic medications	Referral to usual care provider / public safety net clinic	-	Social workers (Cancer Depression Clinical Specialist), psychiatrist	PHQ9 ≥ 10. Timing unclear.	12 months

8	Ell 2010	3	PST (number of sessions in this step not specified) or ADs (patient preference)	PST in step 1: addition of pharmacotherapy; ADs in step 1: change of ADs or adding PST (patient preference)	Additional PST, adding insomnia medication, referral to specialty mental health care.	-	Social work diabetes depression clinical specialists, GP, psychiatrist, assistant patient navigator	Partial or non-response: clinical improvement = SCL or PHQ 50% reduction of symptoms; remission = PHQ < 5 or SCL < 0.5. Assessed at 8 & 12 weeks.	12 months
9	Hujbregts	3	Self-help book (all patients) plus PST (6 or 12 sessions) or PST + ADM (patient preference)	Self-help book, also switching treatments (PST / ADs, patient preference)	Referral to specialty mental health care	-	Depression Care Manager, GP, consultant psychiatrist	PHQ-9 reduction <5 and/or PHQ-9 score ≥ 5 at 6 & 12 weeks.	18 weeks ≈ 4.5 months
10	Katon	3	1 vist + PST (6 sessions) or ADs (patient preference)	Switching treatments, adding treatments, changing ADs and psychiatric consultation	Referral to specialty mental health care	-	Nurses, psychiatrist	PHQ-9 reduction < 50% at 10-12 weeks then + 8-12 weeks.	6 months
11	Patel	4	Face-to-face PE	ADs or IPT (6 to 12 sessions) + adherence management	ADs + IPT (6 to 12 sessions) + adherence management	Continue all treatments + referral to clinical specialist	Lay health counselor (non-medical graduate), GP, psychiatrist	Routine clinical assessment by the health counsellor. Time point not reported.	6 months
12	Seekles	3	PE (1 face-to-face session) + bibliotherapy (content depending on diagnosis, online/telephone support on request)	PST (5 sessions)	Contact with Care Manager (1 session): referral to GP or specialist mental health setting	-	Mental health nurse, junior psychologist	IDS ≥ 14 or HADS-A ≥ 8 or WSAS ≥ 6 every 8 weeks.	18-24 weeks ≈ 6 months
13	Unutzer	3	Videotape + booklet + 1 DCM visit then PST (6 to 8 sessions) or ADs (patient preference)	Switching treatments, adding treatments, changing ADs (patient preference	Team considered alternative treatment for each patient individually (e.g. hospitalisation)	-	Depression Care Manager (nurses, psychologist), psychiatrist, GP	PHQ9 reduction <50% and more than 2 out of the 9 symptoms of MDD. Assessed end step 1 (precise timing not reported) & after 10 weeks step 2 treatment.	12 months
14	Van 't Veer	3	Bibliotherapy (based on CWD; support by telephone calls or face-to-face visits, no predetermined number)	PST (7 sessions)	Referral to GP	-	Home care / community mental health nurse	CES-D ≥16 every 3 months.	12 months

[1] Providers' includes the role of all health care professionals involved in the stepped care intervention except for professionals who cared for patients 'on referral'.
Abbreviations: ADs=antidepressants; CBT = cognitive behavioral therapy; CDCS = cancer depression clinical specialist; CES-D = center for epidemiological studies depression scale; CGI-S= Clinical Global Impression Severity Scale; CWD = Coping with depression; HADS-A = Hospital Anxiety and depression scale-Anxiety; HDRS = Hamilton depression rating scale; IPT = interpersonal psychotherapy; MDD = major depressive disorder; PST = problem solving treatment; PE = psycho-education; WSAS = work and social adjustment scale

Table 3
Click here to download Table(s): Table 3 results meta analysis.docx

Table 3 . Meta-analysis, and subgroup analysis, of 10 studies examining the effects of stepped care for depression compared to care-as-usual: effect sizes (Cohen's *d*)

	N_{comp}	*d*	95% CI	I^2	P value
Post intervention effect sizes					
• Outcomes combined	10	0.38	0.18 to 0.57	81.53*	Na
• Outcomes with highest ES	10	0.42	0.22 to 0.62	81.33*	
• Outcomes with lowest ES	10	0.33	0.13 to 0.52	84.81*	
Effect sizes for different time points (outcomes combined)					
• 2-4 months	4	0.57	0.21 to 0.94	83.61*	Na
• 6 months	10	0.34	0.20 to 0.48	68.11*	
• 9-12 months	5	0.43	0.20 to 0.65	74.81*	
• 18 months	1	0.26	< - 0.01 to 0.53	-	
Subgroup analysis on six months outcomes (*d* = 0.34; 95% CI 0.20 to 0.48)					
Country					
• USA	6	0.38	0.29 to 0.46	0.00	0.63
• Netherlands	2	0.18	-0.22 to 0.58	33.54	
• other	2	0.44	-0.31 to 1.19	94.57*	
IMPACT based					0.79
• yes	7	0.38	0.30 to 0.46	0.00	
• no	3	0.31	-0.18 to 0.80	89.78*	
Progressive treatment intensity?					
• yes	2	0.07	-0.08 to 0.22	0.00	< 0.01
• no	8	0.41	0.33 to 0.49	44.03	
Physical co-morbidity					
• Present	5	0.32	0.19 to 0.44	0.00	0.82
• Absent	5	0.35	0.09 to 0.62	84.11*	
Inclusion based on diagnosis					
• yes	5	0.35	0.09 to 0.62	84.11*	0.82
• no	5	0.32	0.19 to 0.44	0.00	

N_{comp} = number of comparisons; * = P < 0.01; CI = 95% confidence intervals; na = not applicable

Click here to download Other Supplementary Material: Box 1 future research.docx

Box 1. Key areas of future research on stepped care.

Appropriately powered, non-inferiority randomised controlled trial of stepped care for depression and/or other disorders defined by a progressive increase in treatment intensity compared with a single-step high intensity psychological treatment; cost-effectiveness and process analysis of above to be included.

Pilot research into defining a) stepping criteria (algorithm) for stepped care and b) stratification criteria for matched care, leading to an appropriately powered, non-inferiority randomised controlled trial of stepped care for depression and/or other disorders defined by a progressive increase in treatment intensity compared with a matched care control.

Appropriately powered, non-inferiority randomised controlled trial of stepped care for depression defined by progressive intensity of psychological vs. stepped care defined by combined treatment modalities (psychological and pharmacological).

Following more published trials, an updated systematic review of stepped care to help identify (via subgroup analysis) optimal components of stepped care.

Additional randomised controlled trials to compare stepped care with other treatment for the prevention of depression.

Click here to download Other Supplementary Material: Box 2 reporting standards.docx

Box 2. Recommended reporting standards on stepped care

Data to include in the report of a clinical trial on stepped care for depression

Number of patients in stepped care and control group(s)
Drop out prior to step one and between steps (n, %)
Number, % of people discharged from treatment at each step
Number, % of people stepping up to subsequent steps

For each step:
N treated
Health care professionals involved
Training and education provided to deliver clinical protocols
Treatment received

- n patients in receipt
- dose e.g. n sessions of psychological therapy (mean, SD)
- duration e.g. n weeks (mean, SD)

Drop out of treatment during specific step (n, %)
Patient outcomes on end of each treatment step

- n patients' health status assessed
- depressive symptoms (mean, SD, n in analysis)
- n, % recovered or improved with definition of recovery / improvement specified

Stepping criteria:
Measure
Frequency and timeframe of assessment
Definition of improvement / recovery required to end treatment or to step

For the control group:
Number treated
Treatment received (detail as above)
Treatment drop out (n, %)