Illness burden and physical outcomes associated with collaborative care in patients with comorbid depressive disorder in chronic medical conditions

DOI:

[10.1016/j.genhosppsych.2017.08.003](https://doi.org/10.1016/j.genhosppsych.2017.08.003)

Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

van Eck van der Sluijs, J. F., Castelijns, H., Eijsbroek, V., Rijnders, C. A. T., van Marwijk, H. W. J., & van der Feltz-Cornelis, C. M. (2018). Illness burden and physical outcomes associated with collaborative care in patients with comorbid depressive disorder in chronic medical conditions: A systematic review and meta-analysis. *General Hospital Psychiatry*, 50, 1-14. <https://doi.org/10.1016/j.genhosppsych.2017.08.003>

Published in:

General Hospital Psychiatry

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.

Manuscript Details

Manuscript number	GHP_2017_110
Title	Illness burden and physical outcomes associated with collaborative care in patients A systematic review and meta-analysis
Article type	Research Paper

Abstract

Abstract Objective Collaborative care (CC) improves depressive symptoms in people with comorbid depressive disorder in chronic medical conditions, but its effect on physical symptoms has not yet systematically been reviewed. This study aims to do so. **Methods** Systematic review and meta-analysis was conducted using Pubmed, the Cochrane Library, and the European and US Clinical Trial Registers. Eligible studies included randomized controlled trials (RCTs) of CC compared to care as usual (CAU), in primary care and general hospital setting, reporting on physical and depressive symptoms as outcomes. Overall treatment effects were estimated for illness burden, physical outcomes and depression, respectively. **Results** Twenty RCTs were included, with N = 4774 patients. The overall effect size of CC versus CAU for illness burden was OR 1.64 (95%CI 1.47;1.83), d = 0.27 (95%CI 0.21;0.33). Overall, depression outcomes were better for CC than for CAU. Moderator analyses did not yield statistically significant differences. **Conclusions** CC is more effective than CAU in terms of illness burden, physical outcomes and depression, in patients with comorbid depression in chronic medical conditions. More research covering multiple physical disorders is needed.

Corresponding Author	Christina Van der Feltz-Cornelis
Corresponding Author's Institution	Clinical Centre of excellence for Body, Mind and Health, GGz Breburg, Tilburg, The Netherlands
Order of Authors	Jonna van Eck van der Sluijs, Hilde Castelijns, Vera Eijsbroek, Cees Rijnders, Harm van Marwijk, Christina Van der Feltz-Cornelis

Submission Files Included in this PDF

File Name [File Type]

Cover letter.docx [Cover Letter]

Rebuttal letter.docx [Response to Reviewers]

Manuscript.docx [Revised Manuscript with Changes Marked]

Title page.docx [Title Page (with Author Details)]

Manuscript clean.docx [Manuscript (without Author Details)]

Appendix A_searchterm.docx [Figure]

Appendix B_PRISMA 2009 checklist.docx [Figure]

Appendix C_risk of bias assessment.docx [Figure]

Forest plot 1.docx [Figure]

Forest plot 2.docx [Figure]

Forest plot 3.docx [Figure]

Forest plot 4.docx [Figure]

Forest plot 5.docx [Figure]

Funnel plot 1.docx [Figure]

PRISMA flow diagram.docx [Figure]

Table 1.docx [Table]

Table 2 Data extraction table.docx [Table]

Table 3 Data extraction table.docx [Table]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Dear Dr. Huffman,

We have received your comments with regard to our manuscript 'Physical outcomes of collaborative care and illness burden in patients with comorbid depressive disorder in chronic medical conditions: a systematic review and meta-analysis' (GHP_2017_110_R1). We would like to thank you for the fast and careful review of our manuscript, and are pleased that your Journal is interested in considering a revised version. We have carefully attended to the remarks, which were most helpful for improving the manuscript. Enclosed you will also find the revised version of the manuscript, in which we have marked the revisions in yellow. During the revision process we have also changed the title to: 'Illness burden and physical outcomes associated with collaborative care in patients with comorbid depressive disorder in chronic medical conditions: A systematic review and meta-analysis'. We hope that our revised manuscript is suitable for publication in General Hospital Psychiatry.

Editor comments and our response:

Thank you very much for your comprehensive revision. It is clear that a great deal of effort was made to address reviewers' comments, and the inclusion of numerous additional studies, completion of moderator analyses, and multiple other changes all have strengthened the manuscript.

The manuscript still has a number of issues to be addressed prior to re-consideration for publication. These are as follows:

1. The issue of what constitutes a physical outcome is still somewhat unclear. The definition in one location discusses diagnostic and treatment evaluations (this apparently includes readmissions), physical health status, and pain, and elsewhere it is listed as items such as A1C, seizures, HIV-associated symptoms, and so forth. This continues to remain quite confusing and it would be useful to explicitly and in some detail list exactly what constitutes a physical outcome and to keep this consistent throughout the manuscript.

We have explicitly described the outcome measures investigated in our meta-analysis in a new Table (Table 1: Outcomes in meta-analysis). In this Table we list the exact outcome measures.

2. Overall, phrasing and writing remain somewhat unclear in many places. It would be useful to have a native English speaker re-review the text just to ensure the writing is as clear as possible. While the English is not grossly off, the manuscript is complex enough that this modestly impedes review of the manuscript, and at the very least a careful review of the manuscript and good editing will help immensely.

Thank you for this suggestion. We have sent our manuscript to a professional proof reading service to fine tune the writing. Changes are highlighted in the manuscript.

3. Statements about the importance of the hypertension finding should be substantially tempered given that there is only a single study involving hypertension.

We have tempered the importance of this finding in the Discussion section: 'However, given that there is only one study involving hypertension in the meta-analysis, more studies are needed to draw firm conclusions.' Further, we have removed the statement about the hypertension finding from the Abstract.

The issue of the Rollman study (which seems that it should be classified as post-CABG and not multimorbidity) and Huffman (which includes admissions for HF and arrhythmia, which are NOT acute coronary syndrome) persist.

We agree that Huffman 2014 should not be mentioned under acute coronary syndrome, and have relabelled this to cardiac events (ACS, myocardial infarction, unstable angina, heart failure, arrhythmia). Given the variety of physical outcome measures in the Rollman study, we classified this under multi-morbidity; we hope the new Table 1 also clarifies this.

4. The forest plots (especially 1 and 2) should be better labelled in terms of which outcomes they are examining (e.g., illness burden, 'physical outcomes') and the specific constructs (most of them say "combined," and this is somewhat confusing).

We have adjusted the titles of the Forest plots to clarify which outcomes they examine. A more extensive description of the outcome measures is provided in the data extraction table.

5. The title may be better as "Illness burden and physical outcomes associated with collaborative care in patients..."

Thank you for this suggestion. We agree that this is an improvement, and have changed the title.

6. Further detail about how exactly this analysis is different from other analyses performed about collaborative care in patients with medical conditions would be very useful.

We have further explained this in the Methods section: 'An important difference from earlier reviews is that we focus explicitly on illness burden and physical outcomes in CC' and in the Discussion section: 'This is the first systematic review that focusses explicitly on illness burden and physical outcomes of CC for patients with comorbid depression in chronic medical conditions.'

Thank you again for your vigorous revision. If you are able to better and more consistently clarify the physical outcome, clarify how this analysis differs from prior work, refine the overall clarity of the writing, and address these smaller issues, this has the potential to be a very useful contribution to the literature.

Thank you for your consideration of our work. Because of my maternity leave, Prof. Dr. Christina van der Feltz-Cornelis will be available as the corresponding author. Please address all correspondence concerning this manuscript to her at GGz Breburg, and feel free to correspond with her by e-mail (C.M.vdrFeltz@uvt.nl).

Kind regards, and on behalf of my co-authors,

Jonna van Eck van der Sluijs, M.D.

Clinical Centre of Excellence for Body, Mind and Health, GGz Breburg, Lage Witsiebaan 4, 5042 AD Tilburg, The Netherlands.

Contact information for Christina van der Feltz-Cornelis:

Tel +31(0)88-0161555; Fax +31(0)88-0161699; Email C.M.vdrFeltz@uvt.nl

Appendix: rebuttal letter revision GHP_2017_110

Dear Dr. Huffman,

We have received the assessors comments with regard to our manuscript 'Physical outcomes of collaborative care in patients with combined chronic medical conditions and depression in primary care and in general hospital care: a systematic review and meta-analysis' (GHP_2017_110). Hereby we would like to thank you and both assessors for the careful review of our manuscript and are pleased that your Journal is interested in considering a revised version. We have carefully attended to the remarks, which were most helpful to improve the manuscript. Enclosed you will also find the revised version of the manuscript, in which we marked the adaptations in yellow. After the revision we changed the title to: 'Physical outcomes of collaborative care and illness burden in patients with comorbid depressive disorder in chronic medical conditions: a systematic review and meta-analysis'. We hope it is suitable for publication in General Hospital Psychiatry.

Reviewers comments and our response:

Reviewer 1

GHP_2017_110

Strengths of this paper include; Language and publication date limits are specified. The protocol is registered on PROSPERO. An example search string is provided for replication. There is a flow chart of article selection.

The main limitation is that there is no clear definition of what constitutes a physical outcome. This is very unclear from the protocol, and in the paper the explanation is insufficient. The presented results and data extraction table lists a number of outcomes that are reported in the previous systematic reviews for specific diseases. This includes depression (references 23, 24, 25, also see Atlantis et al., BMJ open 2014; Panagioti et al JAMA Psychiatry 2016), glycemic control (reference 23, also see Atlantis et al., BMJ open 2014), quality of life (reference 25), and adverse events (references 24 and 25).

We clarified this on page 10:

'Our definition of a physical outcome is physical symptoms in terms of:

- somatic diagnostic or treatment interventions
- physical health status (including disability)
- pain'

In fact the secondary outcome in the protocol (e.g. depression) substantially overlaps the previous reviews in this area.

Page 4: We clarified the main focus of our review in the Rationale:

'Although research into comorbid physical and depression is being done, in general physical outcomes are not the main focus of this research. We focus on this and define physical outcomes as physical burden of disease, i.e. physical symptoms including pain, somatic events such as myocardial infarction (MI), and physical functioning.'

Curiously there is no exclusion criteria listed.

Page 8: we added the exclusion criteria:

'Exclusion criteria were:

- Articles that focus on only physical symptoms or only depressive-anxiety symptoms, instead of on both.
- Interventions that did not meet our criteria for CC
- High risk of bias, indicated by a score of <4 out of 7 in the risk of bias assessment performed according to the Cochrane quality criteria [41].'

No PRISMA checklist was provided.

We added the PRISMA checklist as Appendix B.

There are too few studies for disease-specific analyses as presented here. Each chronic disease is comprised by 3 or fewer studies. The authors attempt to collate all diseases together which despite the I² statistic introduces substantial clinical and methodological heterogeneity. For example, how is a pooled-analysis of HbA1C, number of epileptic seizures, HIV-symptom severity, pain, or physical functioning meaningful? The overall pooled effect is omitted from Figure 1 (i.e. .20, 95% CI .13 - .27).

We have added several studies based on the suggestions of the reviewers (see further) and now present a total of 20 RCTs in the meta-analysis. We now focus more on 'disease burden' instead of on the disease-specific analyses. See the description on page 9/10:

'Outcomes should define impact of the intervention on the clinical condition of the patients in terms of physical symptomatology and depressive symptom severity. Outcomes should report objective measures of such dimensions, as assessed by validated questionnaires or methods. These questionnaires could either be based on self-report or clinician rated.

Assessment of physical outcomes was done by:

- somatic diagnostic or intervention procedures
- physical health status (including disability)
- Pain

Assessment of depressive outcomes was done in terms of severity of symptoms, reduction of symptoms, treatment response or remission.'

Studies are excluded based on risk of bias but this was not stated in the protocol or review exclusion criteria. 6 studies are excluded due to risk of bias that is only apparent after reading the results section – para 1. The authors elected to not present all studies in analysis stratified by risk of bias. Justification is provided from the Cochrane handbook in the results section rather than the methods section. This seems ad-hoc. The authors should probably present all data and then stratify by risk of bias. As the Cochrane handbook states, "The intended strategy should be described in the protocol for the review."

We agree that we should have explained this approach in the methods section and clarified this in the methods section.

The authors label Huffman 2014 as acute coronary syndrome when they included arrhythmia and heart failure. Despite this, the authors class Rollman et al's work in coronary heart disease as multimorbidity.

We agree there is an overlap between physical conditions in several articles. We search for the effect of a generic CC model for which we make a distinction between one chronic medical condition and multimorbidity. We label Huffman 2014 as acute cardiac syndrome (including myocardial infarction or unstable angina). We label Rollman 2009 as multimorbidity, given that they focus on hypertension, diabetes, hyperlipidemia, stroke, COPD, chronic renal insufficiency, myocardial infarction, and congestive heart failure.

Can questions about setting, CC model and care manager be answered statistically by meta-regression?

We looked at these three potentially interesting effect moderators by calculating and pooling effect sizes. After inclusion of the studies as described further, no statistically significant differences were found between CC and CAU. Hence we did not perform a further analysis.

Other weaknesses include that the strength of evidence is not graded according to GRADE criteria. Indeed we did not use GRADE criteria, instead we used the Cochrane criteria. We revised our rating of the risk of bias (ROB) and ran a new analysis including all articles.

Is Coventry et al BMJ 2015 eligible, the COINCIDE trial?

We contacted Dr. Coventry about the results of the COINCIDE trial, with a special interest in the (secondary) outcomes related to diabetes and CHD. He responded that no physical health outcomes were measured. Therefore we cannot include the COINCIDE trial.

Suggested minor changes

Abstract

Remove “depressive outcomes” from the conclusion as this was not a focus of the paper... “CC is more effective than care as usual for physical outcomes and depressive outcomes....”

We removed ‘depressive outcomes’.

Introduction

Suggest rephrasing the sentence... “So far, this has been explored....” To
Previous systematic reviews have examined mental health outcomes in comorbid diabetes and depression [23], cancer and depression [24], and coronary heart disease and depression [25].
However, no systematic review has assessed physical outcomes in comorbid depression in general chronic medical conditions.”

We rephrased the sentence as suggested.

Suggest rephrasing the sentence... “Besides in general health...” to “Besides general healthcare settings.

We rephrased the sentence as suggested.

Change “A score of <4 meant...” and replace with “too many”

We rephrased the sentence as suggested.

Reviewer 2

This manuscript is an updated systematic review of collaborative care for depression that essentially confirms prior reviews on its effectiveness for depression and is reassuring, but the twist is on the impact of collaborative care for depression on physical outcomes. Also, the manuscript is submitted by a European group who liberally reference the European literature that U.S. readers of GHP are likely to be less familiar with.

We clarified in the Rationale that the main focus of this review is on the physical outcomes. In order to further stay in line with the literature readers of GHP are familiar with, we added a number of relevant references from GHP (including US setting) in the Introduction section.

Questions/Suggestions.

1) The meta-analysis only includes 14 papers so it is hard to make sweeping generalizations on the impact of CC within the disorders described. Indeed, many times in the Discussion section they write about a difference in effect sizes, but hedge that the 95% CIs overlap.

Indeed, the number of papers is low. We did an extra search by contacting experts with one extra resulting study (Carney et al 2016). And we added two articles that were suggested by the reviewers. We excluded Bogner's trial (2007), because there were no results reported on the subgroup of depressed patients with diabetes. We excluded Eli 2011, because there were no physical outcomes reported; and excluded Pols 2008, because the focus was on somatization instead of on a specific physical disorder. We also reran the ROB assessment and now included all selected trials. Hence we could rerun the meta-analysis on 20 trials. We rewrote our results and discussion sections accordingly, with clear descriptions about which results were statistically significant and which not.

2) Several important trials may have been overlooked. For a systematic review, I would have expected Kurt Kroenke's SCAMP (2009) and INCPAD Trials (2010) for depression and pain, Karina Davidson's CODIACS Trial (2013) to be included (I looked at the abstracts and tables myself and am unclear why these didn't make their cut) and there may be other important trials missed. Please review these studies and consider including.

We would like to thank the reviewer for these suggestions. In the previous version of our review these trials were excluded based on the information in the abstract. We now checked the full texts of these trials and included Kroenke's SCAMP trial (2009) and Davidson's CODIACS trial (2013) in the selection of articles. Kroenke's INCPAD trial did not report outcomes about the subgroup of patients with both pain and depression and we could therefore not include it.

3) Overlooked Trials. The authors present their database search strategy, but did they also contact PIs of key studies about other trials that might be eligible for inclusion to complete their search (crowdsharing")?

Thank you for this suggestion. We did contact PIs about other possible trials, which yielded several suggestions but no more articles that included physical outcome data.

We also did a search in two clinical trial registers to find ongoing trials: the 'European Union Clinical Trials Register' and 'ClinicalTrials.gov'; a service of the U.S. National Institutes of Health'. This yielded 52 suggestions but no articles that we could include.

We also checked the articles in the updated search of Archer et al's review. This yielded one new article we could include.

4) Overlooked Trials. See Panagioti's 2016 meta-analysis on chronic physical conditions and the effectiveness of collaborative care for depression published in JAMA Psychiatry. It included 31 RCTs that may be relevant for inclusion in this report.

In this meta-analysis nine studies are described that 'explicitly recruited patients with chronic physical conditions'. We checked these studies:

- Katon WJ, Von Korff M, Lin EH, et al. The Pathways Study: a randomized trial of collaborative care in patients with diabetes and depression. *Arch Gen Psychiatry*. 2004;61(10):1042-1049. → this study was already included in our review
- Katon WJ, Lin EH, Von Korff M, et al. Collaborative care for patients with depression and chronic illnesses. *N Engl J Med*. 2010;363(27): 2611-2620. → this study was already included in our review
- Davidson KW, Bigger JT, Burg MM, et al. Centralized, stepped, patient preference-based treatment for patients with post-acute coronary syndrome depression: CODIACS vanguard randomized controlled trial. *JAMA Intern Med*. 2013; 173(11):997-1004. → we included this study, also based on the suggestion of the reviewer (see point 2).
- Eli K, Xie B, Quon B, Quinn DI, Dwight-Johnson M, Lee P-J. Randomized controlled trial of collaborative care management of depression among low-income patients with cancer. *J Clin Oncol*. 2008;26(27):4488-4496. → this study was already included in our review.
- Eli K, Katon W, Xie B, et al. Collaborative care management of major depression among low-income, predominantly Hispanic subjects with diabetes: a randomized controlled trial. *Diabetes Care*. 2010;33(4):706-713. → this study was already included in our review.

- Huffman JC, Mastromauro CA, Sowden GL, Wittmann C, Rodman R, Januzzi JL. A collaborative care depression management program for cardiac inpatients: depression characteristics and in-hospital outcomes. *Psychosomatics*. 2011;52(1): 26–33. → we excluded this study in our selection process, because cardiac symptoms were only mentioned as baseline characteristics.
- Huffman JC, Mastromauro CA, Beach SR, et al. Collaborative care for depression and anxiety disorders in patients with recent cardiac events: the Management of Sadness and Anxiety in Cardiology (MOSAIC) randomized clinical trial. *JAMA Intern Med*. 2014;174(6):927–935. → this study was already included in our review.
- Rollman BL, Belnap BH, LeMenager MS, et al. Telephone-delivered collaborative care for treating post-CABG depression: a randomized controlled trial. *JAMA*. 2009;302(19):2095–2103. → this study was already included in our review.
- Coventry PA, Hudson JL, Kontopantelis E, et al. Characteristics of effective collaborative care for treatment of depression: a systematic review and meta-regression of 74 randomised controlled trials. *PLoS One*. 2014;9(9):e108114. → this study is not a RCT about CC for depression or anxiety disorder with comorbid chronic medical conditions and therefore does not meet our inclusion criteria

5) Excluded Trials. The investigators identified 20 trials, but excluded 6 because of bias (Table 1). I can't speak with expertise on all of the papers excluded, but it raises a red flag that Katon's TEAMcare paper published in NEJM was rated a "3" and excluded. On pp 10-11 they write that "there were no statistically significant results....", but this is false for TEAMcare as their blended collaborative care intervention demonstrated significant improvements in blood pressure, HgbA1c, and cholesterol. Moreover, the paper suggests that blended models of CC may produce particularly strong ES improvements in mood (ES>0.6 vs. Katon's 2004 Pathway Trial cited with ES 0.22). Please re-review excluded papers, or at the very least given how few studies are included, please conduct a 'sensitivity analysis' that include the 6 excluded trials to see if the results change. This is listed as a limitation, but it shouldn't be.

We checked our ROB rating assessment and adapted this for incomplete ratings. Hence, all selected trials had better ROB ratings and were no longer excluded from the meta-analysis. We added the complete risk of bias tables for all studies (Appendix C, we suggest to add this as an online appendix or on request available). We also added a full description of all studies in Table 2 (data extraction table) and conducted the meta-analysis again with all articles included.

6) New Figure. Both ES for mood and physical health are listed in the summary Table 2. It would be interesting to plot the ES in a graph to visually see if there is a relationship between improvements in mood and physical health observed in each trial and generate a numerical correlation coefficient if there may be. i.e., did greater improvements in ES for mood produce larger ES improvements in physical health?

In the current version of the data extraction table (Table 2) we present: the illness burden combined for the physical and depression outcome, the physical outcome, and the depression outcome (odds ratio's and effect sizes). We also present the forest plots of these outcomes. We chose to focus on these outcomes, and (given the different physical outcome measures and limited number of RCTs) chose not to add the suggested new figure.

7) Discussion. Given how few studies were included within each disease in the present analyses, it is difficult to make sweeping generalizations. Indeed, the authors frequently cite that the 95% CIs often overlap several times in the Discussion. Therefore, suggest that the Discussion section be shortened and less sweeping.

We agree that the 95%CIs should be taken into account and reconsidered which results to present and rewrote the discussion.

8) Discussion. There is mention of cost-effectiveness (p. 16). Several of the CC trials have published related cost-effectiveness papers. I am familiar with C-E papers referencing Davidson's COPES and CODIACS trials, Rollman, and Katon's TeamCare and there may be others worth citing.

Our review focusses on effectiveness of CC for depression or anxiety disorder with comorbid chronic medical conditions, not on cost-effectiveness. However, we agree that we could mention the subject of cost effectiveness and did add a paragraph on that subject in the Discussion section: 'More somatic diagnostic or treatment procedures are performed in case of CC when compared to CAU (OR 2.25). Although more procedures cost more, several studies show that CC improves quality of life for acceptable additional costs [38, 83]. This was the case for several chronic medical conditions with comorbid depression [84-89].'

9) Discussion/Conclusions. I'd like to see a call for more "blended" CC trials like TEAMcare be conducted to examine the impact of one CM who treats both mental and physical health conditions to see if this model is more effective at improving mood and physical health as found by Dr. Katon (I know of at least two U.S. trials in progress or underway and there may be others).

We think that the use of the term 'blended' is confusing here. Blended is mostly used in case of combined real life and e-health treatments, which was not the subject here in this article. The idea that it is important to address both mental and physical health conditions is relevant but it is part of most collaborative care interventions already, not only recently. However, we agree that the importance of addressing both physical and mental health in monitoring, counseling and consultation in CC should receive more attention in the discussion and hence we adapted the Discussion section: '...Apparently, it is more important for improvement in both physical and depressive symptoms that the CC model is applied, than by who it is applied or whether it is provided in primary care or general hospital setting. Therefore, addressing both physical and mental health in monitoring, counselling and consultation can be seen as important factors in the success of CC. This is supported by the finding that reduction in physical symptoms is associated with reduction in depression [90].'

10) PRISMA Checklist. Did the authors submit the PRISMA checklist in preparing their meta-analysis? I see it mentioned and referenced in the text, but do not see the completed checklist itself included with the submitted manuscript. <http://www.prisma-statement.org>
We added the PRISMA checklist as Appendix B.

We thank both reviewers for their useful comments. We do hope that our revision is sufficient to the reviewers and that our article is suitable for publication in General Hospital Psychiatry. We await your response.

Thank you for your consideration of our work. Because of my upcoming maternity leave, prof.dr. Christina van der Feltz-Cornelis will be available as the corresponding author. Please address all correspondence concerning this manuscript to her at GGz Breburg and feel free to correspond with her by e-mail (C.M.vdrFeltz@uvt.nl).

Kind regards, also on behalf of my co-authors,
Jonna van Eck van der Sluijs, M.D.
Clinical Centre of excellence for Body, Mind and Health, GGz Breburg, Lage Witsiebaan 4, 5042 AD Tilburg, The Netherlands

Contact information Christina van der Feltz-Cornelis:
Tel +31(0)88-0161555; Fax +31(0)88-0161699; Email C.M.vdrFeltz@uvt.nl

Dear Dr. Huffman,

We have received your comments with regard to our manuscript 'Physical outcomes of collaborative care and illness burden in patients with comorbid depressive disorder in chronic medical conditions: a systematic review and meta-analysis' (GHP_2017_110_R1). We would like to thank you for the fast and careful review of our manuscript, and are pleased that your Journal is interested in considering a revised version. We have carefully attended to the remarks, which were most helpful for improving the manuscript. Enclosed you will also find the revised version of the manuscript, in which we have marked the revisions in yellow. During the revision process we have also changed the title to: 'Illness burden and physical outcomes associated with collaborative care in patients with comorbid depressive disorder in chronic medical conditions: A systematic review and meta-analysis'. We hope that our revised manuscript is suitable for publication in General Hospital Psychiatry.

Editor comments and our response:

Thank you very much for your comprehensive revision. It is clear that a great deal of effort was made to address reviewers' comments, and the inclusion of numerous additional studies, completion of moderator analyses, and multiple other changes all have strengthened the manuscript.

The manuscript still has a number of issues to be addressed prior to re-consideration for publication. These are as follows:

1. The issue of what constitutes a physical outcome is still somewhat unclear. The definition in one location discusses diagnostic and treatment evaluations (this apparently includes readmissions), physical health status, and pain, and elsewhere it is listed as items such as A1C, seizures, HIV-associated symptoms, and so forth. This continues to remain quite confusing and it would be useful to explicitly and in some detail list exactly what constitutes a physical outcome and to keep this consistent throughout the manuscript.

[We have explicitly described the outcome measures investigated in our meta-analysis in a new Table \(Table 1: Outcomes in meta-analysis\). In this Table we list the exact outcome measures.](#)

2. Overall, phrasing and writing remain somewhat unclear in many places. It would be useful to have a native English speaker re-review the text just to ensure the writing is as clear as possible. While the English is not grossly off, the manuscript is complex enough that this modestly impedes review of the manuscript, and at the very least a careful review of the manuscript and good editing will help immensely.

Thank you for this suggestion. We have sent our manuscript to a professional proof reading service to fine tune the writing. Changes are highlighted in the manuscript.

3. Statements about the importance of the hypertension finding should be substantially tempered given that there is only a single study involving hypertension.

We have tempered the importance of this finding in the Discussion section: 'However, given that there is only one study involving hypertension in the meta-analysis, more studies are needed to draw firm conclusions.' Further, we have removed the statement about the hypertension finding from the Abstract.

The issue of the Rollman study (which seems that it should be classified as post-CABG and not multimorbidity) and Huffman (which includes admissions for HF and arrhythmia, which are NOT acute coronary syndrome) persist.

We agree that Huffman 2014 should not be mentioned under acute coronary syndrome, and have relabelled this to cardiac events (ACS, myocardial infarction, unstable angina, heart failure, arrhythmia). Given the variety of physical outcome measures in the Rollman study, we classified this under multi-morbidity; we hope the new Table 1 also clarifies this.

4. The forest plots (especially 1 and 2) should be better labelled in terms of which outcomes they are examining (e.g., illness burden, 'physical outcomes') and the specific constructs (most of them say "combined," and this is somewhat confusing).

We have adjusted the titles of the Forest plots to clarify which outcomes they examine. A more extensive description of the outcome measures is provided in the data extraction table.

5. The title may be better as "Illness burden and physical outcomes associated with collaborative care in patients..."

Thank you for this suggestion. We agree that this is an improvement, and have changed the title.

6. Further detail about how exactly this analysis is different from other analyses performed about collaborative care in patients with medical conditions would be very useful.

We have further explained this in the Methods section: 'An important difference from earlier reviews is that we focus explicitly on illness burden and physical outcomes in CC' and in the Discussion section: 'This is the first systematic review that focusses explicitly on illness burden and physical outcomes of CC for patients with comorbid depression in chronic medical conditions.'

Thank you again for your vigorous revision. If you are able to better and more consistently clarify the physical outcome, clarify how this analysis differs from prior work, refine the overall clarity of the writing, and address these smaller issues, this has the potential to be a very useful contribution to the literature.

Thank you for your consideration of our work. Because of my maternity leave, Prof. Dr. Christina van der Feltz-Cornelis will be available as the corresponding author. Please address all correspondence concerning this manuscript to her at GGz Breburg, and feel free to correspond with her by e-mail (C.M.vdrFeltz@uvt.nl).

Kind regards, and on behalf of my co-authors,

Jonna van Eck van der Sluijs, M.D.

Clinical Centre of Excellence for Body, Mind and Health, GGz Breburg, Lage Witsiebaan 4, 5042 AD Tilburg, The Netherlands.

Contact information for Christina van der Feltz-Cornelis:

Tel +31(0)88-0161555; Fax +31(0)88-0161699; Email C.M.vdrFeltz@uvt.nl

Appendix: rebuttal letter revision GHP_2017_110

Dear Dr. Huffman,

We have received the assessors comments with regard to our manuscript 'Physical outcomes of collaborative care in patients with combined chronic medical conditions and depression in primary care and in general hospital care: a systematic review and meta-analysis' (GHP_2017_110). Hereby we would like to thank you and both assessors for the careful review of our manuscript and are pleased that your Journal is interested in considering a revised version. We have carefully attended to the remarks, which were most helpful to improve the manuscript. Enclosed you will also find the revised version of the manuscript, in which we marked the adaptations in yellow. After the revision we changed the title to: 'Physical outcomes of collaborative care and illness burden in patients with comorbid depressive disorder in chronic medical conditions: a systematic review and meta-analysis'. We hope it is suitable for publication in General Hospital Psychiatry.

Reviewers comments and our response:

Reviewer 1

GHP_2017_110

Strengths of this paper include; Language and publication date limits are specified. The protocol is registered on PROSPERO. An example search string is provided for replication. There is a flow chart of article selection.

The main limitation is that there is no clear definition of what constitutes a physical outcome. This is very unclear from the protocol, and in the paper the explanation is insufficient. The presented results and data extraction table lists a number of outcomes that are reported in the previous systematic reviews for specific diseases. This includes depression (references 23, 24, 25, also see Atlantis et al., BMJ open 2014; Panagioti et al JAMA Psychiatry 2016), glycemic control (reference 23, also see Atlantis et al., BMJ open 2014), quality of life (reference 25), and adverse events (references 24 and 25).

We clarified this on page 10:

'Our definition of a physical outcome is physical symptoms in terms of:

- somatic diagnostic or treatment interventions
- physical health status (including disability)
- pain'

In fact the secondary outcome in the protocol (e.g. depression) substantially overlaps the previous reviews in this area.

Page 4: We clarified the main focus of our review in the Rationale:

'Although research into comorbid physical and depression is being done, in general physical outcomes are not the main focus of this research. We focus on this and define physical outcomes as physical burden of disease, i.e. physical symptoms including pain, somatic events such as myocardial infarction (MI), and physical functioning.'

Curiously there is no exclusion criteria listed.

Page 8: we added the exclusion criteria:

'Exclusion criteria were:

- Articles that focus on only physical symptoms or only depressive-anxiety symptoms, instead of on both.
- Interventions that did not meet our criteria for CC
- High risk of bias, indicated by a score of <4 out of 7 in the risk of bias assessment performed according to the Cochrane quality criteria [41].'

No PRISMA checklist was provided.

We added the PRISMA checklist as Appendix B.

There are too few studies for disease-specific analyses as presented here. Each chronic disease is comprised by 3 or fewer studies. The authors attempt to collate all diseases together which despite the I² statistic introduces substantial clinical and methodological heterogeneity. For example, how is a pooled-analysis of HbA1C, number of epileptic seizures, HIV-symptom severity, pain, or physical functioning meaningful? The overall pooled effect is omitted from Figure 1 (i.e. .20, 95% CI .13 - .27).

We have added several studies based on the suggestions of the reviewers (see further) and now present a total of 20 RCTs in the meta-analysis. We now focus more on 'disease burden' instead of on the disease-specific analyses. See the description on page 9/10:

'Outcomes should define impact of the intervention on the clinical condition of the patients in terms of physical symptomatology and depressive symptom severity. Outcomes should report objective measures of such dimensions, as assessed by validated questionnaires or methods. These questionnaires could either be based on self-report or clinician rated.

Assessment of physical outcomes was done by:

- somatic diagnostic or intervention procedures
- physical health status (including disability)
- Pain

Assessment of depressive outcomes was done in terms of severity of symptoms, reduction of symptoms, treatment response or remission.'

Studies are excluded based on risk of bias but this was not stated in the protocol or review exclusion criteria. 6 studies are excluded due to risk of bias that is only apparent after reading the results section – para 1. The authors elected to not present all studies in analysis stratified by risk of bias. Justification is provided from the Cochrane handbook in the results section rather than the methods section. This seems ad-hoc. The authors should probably present all data and then stratify by risk of bias. As the Cochrane handbook states, "The intended strategy should be described in the protocol for the review."

We agree that we should have explained this approach in the methods section and clarified this in the methods section.

The authors label Huffman 2014 as acute coronary syndrome when they included arrhythmia and heart failure. Despite this, the authors class Rollman et al's work in coronary heart disease as multimorbidity.

We agree there is an overlap between physical conditions in several articles. We search for the effect of a generic CC model for which we make a distinction between one chronic medical condition and multimorbidity. We label Huffman 2014 as acute cardiac syndrome (including myocardial infarction or unstable angina). We label Rollman 2009 as multimorbidity, given that they focus on hypertension, diabetes, hyperlipidemia, stroke, COPD, chronic renal insufficiency, myocardial infarction, and congestive heart failure.

Can questions about setting, CC model and care manager be answered statistically by meta-regression?

We looked at these three potentially interesting effect moderators by calculating and pooling effect sizes. After inclusion of the studies as described further, no statistically significant differences were found between CC and CAU. Hence we did not perform a further analysis.

Other weaknesses include that the strength of evidence is not graded according to GRADE criteria. Indeed we did not use GRADE criteria, instead we used the Cochrane criteria. We revised our rating of the risk of bias (ROB) and ran a new analysis including all articles.

Is Coventry et al BMJ 2015 eligible, the COINCIDE trial?

We contacted Dr. Coventry about the results of the COINCIDE trial, with a special interest in the (secondary) outcomes related to diabetes and CHD. He responded that no physical health outcomes were measured. Therefore we cannot include the COINCIDE trial.

Suggested minor changes

Abstract

Remove “depressive outcomes” from the conclusion as this was not a focus of the paper... “CC is more effective than care as usual for physical outcomes and depressive outcomes....”

We removed ‘depressive outcomes’.

Introduction

Suggest rephrasing the sentence... “So far, this has been explored....” To
Previous systematic reviews have examined mental health outcomes in comorbid diabetes and depression [23], cancer and depression [24], and coronary heart disease and depression [25].
However, no systematic review has assessed physical outcomes in comorbid depression in general chronic medical conditions.”

We rephrased the sentence as suggested.

Suggest rephrasing the sentence... “Besides in general health...” to “Besides general healthcare settings.

We rephrased the sentence as suggested.

Change “A score of <4 meant...” and replace with “too many”

We rephrased the sentence as suggested.

Reviewer 2

This manuscript is an updated systematic review of collaborative care for depression that essentially confirms prior reviews on its effectiveness for depression and is reassuring, but the twist is on the impact of collaborative care for depression on physical outcomes. Also, the manuscript is submitted by a European group who liberally reference the European literature that U.S. readers of GHP are likely to be less familiar with.

We clarified in the Rationale that the main focus of this review is on the physical outcomes. In order to further stay in line with the literature readers of GHP are familiar with, we added a number of relevant references from GHP (including US setting) in the Introduction section.

Questions/Suggestions.

1) The meta-analysis only includes 14 papers so it is hard to make sweeping generalizations on the impact of CC within the disorders described. Indeed, many times in the Discussion section they write about a difference in effect sizes, but hedge that the 95% CIs overlap.

Indeed, the number of papers is low. We did an extra search by contacting experts with one extra resulting study (Carney et al 2016). And we added two articles that were suggested by the reviewers. We excluded Bogner's trial (2007), because there were no results reported on the subgroup of depressed patients with diabetes. We excluded Eli 2011, because there were no physical outcomes reported; and excluded Pols 2008, because the focus was on somatization instead of on a specific physical disorder. We also reran the ROB assessment and now included all selected trials. Hence we could rerun the meta-analysis on 20 trials. We rewrote our results and discussion sections accordingly, with clear descriptions about which results were statistically significant and which not.

2) Several important trials may have been overlooked. For a systematic review, I would have expected Kurt Kroenke's SCAMP (2009) and INCPAD Trials (2010) for depression and pain, Karina Davidson's CODIACS Trial (2013) to be included (I looked at the abstracts and tables myself and am unclear why these didn't make their cut) and there may be other important trials missed. Please review these studies and consider including.

We would like to thank the reviewer for these suggestions. In the previous version of our review these trials were excluded based on the information in the abstract. We now checked the full texts of these trials and included Kroenke's SCAMP trial (2009) and Davidson's CODIACS trial (2013) in the selection of articles. Kroenke's INCPAD trial did not report outcomes about the subgroup of patients with both pain and depression and we could therefore not include it.

3) Overlooked Trials. The authors present their database search strategy, but did they also contact PIs of key studies about other trials that might be eligible for inclusion to complete their search (crowdsharing")?

Thank you for this suggestion. We did contact PIs about other possible trials, which yielded several suggestions but no more articles that included physical outcome data.

We also did a search in two clinical trial registers to find ongoing trials: the 'European Union Clinical Trials Register' and 'ClinicalTrials.gov'; a service of the U.S. National Institutes of Health'. This yielded 52 suggestions but no articles that we could include.

We also checked the articles in the updated search of Archer et al's review. This yielded one new article we could include.

4) Overlooked Trials. See Panagioti's 2016 meta-analysis on chronic physical conditions and the effectiveness of collaborative care for depression published in JAMA Psychiatry. It included 31 RCTs that may be relevant for inclusion in this report.

In this meta-analysis nine studies are described that 'explicitly recruited patients with chronic physical conditions'. We checked these studies:

- Katon WJ, Von Korff M, Lin EH, et al. The Pathways Study: a randomized trial of collaborative care in patients with diabetes and depression. *Arch Gen Psychiatry*. 2004;61(10):1042-1049. → this study was already included in our review
- Katon WJ, Lin EH, Von Korff M, et al. Collaborative care for patients with depression and chronic illnesses. *N Engl J Med*. 2010;363(27): 2611-2620. → this study was already included in our review
- Davidson KW, Bigger JT, Burg MM, et al. Centralized, stepped, patient preference-based treatment for patients with post-acute coronary syndrome depression: CODIACS vanguard randomized controlled trial. *JAMA Intern Med*. 2013; 173(11):997-1004. → we included this study, also based on the suggestion of the reviewer (see point 2).
- Eli K, Xie B, Quon B, Quinn DI, Dwight-Johnson M, Lee P-J. Randomized controlled trial of collaborative care management of depression among low-income patients with cancer. *J Clin Oncol*. 2008;26(27):4488-4496. → this study was already included in our review.
- Eli K, Katon W, Xie B, et al. Collaborative care management of major depression among low-income, predominantly Hispanic subjects with diabetes: a randomized controlled trial. *Diabetes Care*. 2010;33(4):706-713. → this study was already included in our review.

- Huffman JC, Mastromauro CA, Sowden GL, Wittmann C, Rodman R, Januzzi JL. A collaborative care depression management program for cardiac inpatients: depression characteristics and in-hospital outcomes. *Psychosomatics*. 2011;52(1): 26–33. → we excluded this study in our selection process, because cardiac symptoms were only mentioned as baseline characteristics.
- Huffman JC, Mastromauro CA, Beach SR, et al. Collaborative care for depression and anxiety disorders in patients with recent cardiac events: the Management of Sadness and Anxiety in Cardiology (MOSAIC) randomized clinical trial. *JAMA Intern Med*. 2014;174(6):927–935. → this study was already included in our review.
- Rollman BL, Belnap BH, LeMenager MS, et al. Telephone-delivered collaborative care for treating post-CABG depression: a randomized controlled trial. *JAMA*. 2009;302(19):2095–2103. → this study was already included in our review.
- Coventry PA, Hudson JL, Kontopantelis E, et al. Characteristics of effective collaborative care for treatment of depression: a systematic review and meta-regression of 74 randomised controlled trials. *PLoS One*. 2014;9(9):e108114. → this study is not a RCT about CC for depression or anxiety disorder with comorbid chronic medical conditions and therefore does not meet our inclusion criteria

5) Excluded Trials. The investigators identified 20 trials, but excluded 6 because of bias (Table 1). I can't speak with expertise on all of the papers excluded, but it raises a red flag that Katon's TEAMcare paper published in NEJM was rated a "3" and excluded. On pp 10-11 they write that "there were no statistically significant results....", but this is false for TEAMcare as their blended collaborative care intervention demonstrated significant improvements in blood pressure, HgbA1c, and cholesterol. Moreover, the paper suggests that blended models of CC may produce particularly strong ES improvements in mood (ES>0.6 vs. Katon's 2004 Pathway Trial cited with ES 0.22). Please re-review excluded papers, or at the very least given how few studies are included, please conduct a 'sensitivity analysis' that include the 6 excluded trials to see if the results change. This is listed as a limitation, but it shouldn't be.

We checked our ROB rating assessment and adapted this for incomplete ratings. Hence, all selected trials had better ROB ratings and were no longer excluded from the meta-analysis. We added the complete risk of bias tables for all studies (Appendix C, we suggest to add this as an online appendix or on request available). We also added a full description of all studies in Table 2 (data extraction table) and conducted the meta-analysis again with all articles included.

6) New Figure. Both ES for mood and physical health are listed in the summary Table 2. It would be interesting to plot the ES in a graph to visually see if there is a relationship between improvements in mood and physical health observed in each trial and generate a numerical correlation coefficient if there may be. i.e., did greater improvements in ES for mood produce larger ES improvements in physical health?

In the current version of the data extraction table (Table 2) we present: the illness burden combined for the physical and depression outcome, the physical outcome, and the depression outcome (odds ratio's and effect sizes). We also present the forest plots of these outcomes. We chose to focus on these outcomes, and (given the different physical outcome measures and limited number of RCTs) chose not to add the suggested new figure.

7) Discussion. Given how few studies were included within each disease in the present analyses, it is difficult to make sweeping generalizations. Indeed, the authors frequently cite that the 95% CIs often overlap several times in the Discussion. Therefore, suggest that the Discussion section be shortened and less sweeping.

We agree that the 95%CIs should be taken into account and reconsidered which results to present and rewrote the discussion.

8) Discussion. There is mention of cost-effectiveness (p. 16). Several of the CC trials have published related cost-effectiveness papers. I am familiar with C-E papers referencing Davidson's COPES and CODIACS trials, Rollman, and Katon's TeamCare and there may be others worth citing.

Our review focusses on effectiveness of CC for depression or anxiety disorder with comorbid chronic medical conditions, not on cost-effectiveness. However, we agree that we could mention the subject of cost effectiveness and did add a paragraph on that subject in the Discussion section: 'More somatic diagnostic or treatment procedures are performed in case of CC when compared to CAU (OR 2.25). Although more procedures cost more, several studies show that CC improves quality of life for acceptable additional costs [38, 83]. This was the case for several chronic medical conditions with comorbid depression [84-89].'

9) Discussion/Conclusions. I'd like to see a call for more "blended" CC trials like TEAMcare be conducted to examine the impact of one CM who treats both mental and physical health conditions to see if this model is more effective at improving mood and physical health as found by Dr. Katon (I know of at least two U.S. trials in progress or underway and there may be others).

We think that the use of the term 'blended' is confusing here. Blended is mostly used in case of combined real life and e-health treatments, which was not the subject here in this article. The idea that it is important to address both mental and physical health conditions is relevant but it is part of most collaborative care interventions already, not only recently. However, we agree that the importance of addressing both physical and mental health in monitoring, counseling and consultation in CC should receive more attention in the discussion and hence we adapted the Discussion section: '...Apparently, it is more important for improvement in both physical and depressive symptoms that the CC model is applied, than by who it is applied or whether it is provided in primary care or general hospital setting. Therefore, addressing both physical and mental health in monitoring, counselling and consultation can be seen as important factors in the success of CC. This is supported by the finding that reduction in physical symptoms is associated with reduction in depression [90].'

10) PRISMA Checklist. Did the authors submit the PRISMA checklist in preparing their meta-analysis? I see it mentioned and referenced in the text, but do not see the completed checklist itself included with the submitted manuscript. <http://www.prisma-statement.org>
We added the PRISMA checklist as Appendix B.

We thank both reviewers for their useful comments. We do hope that our revision is sufficient to the reviewers and that our article is suitable for publication in General Hospital Psychiatry. We await your response.

Thank you for your consideration of our work. Because of my upcoming maternity leave, prof.dr. Christina van der Feltz-Cornelis will be available as the corresponding author. Please address all correspondence concerning this manuscript to her at GGz Breburg and feel free to correspond with her by e-mail (C.M.vdrFeltz@uvt.nl).

Kind regards, also on behalf of my co-authors,
Jonna van Eck van der Sluijs, M.D.
Clinical Centre of excellence for Body, Mind and Health, GGz Breburg, Lage Witsiebaan 4, 5042 AD Tilburg, The Netherlands

Contact information Christina van der Feltz-Cornelis:
Tel +31(0)88-0161555; Fax +31(0)88-0161699; Email C.M.vdrFeltz@uvt.nl

Illness burden and physical outcomes associated with collaborative care in patients
A systematic review and meta-analysis

Abbreviated running title: Collaborative care: Review and meta-analysis

Total number of words: 8111 (Introduction through to Discussion: 4908)

Number of Tables: 3

Number of Figures: 7

Abstract

Objective

Collaborative care (CC) improves depressive symptoms in people with comorbid depressive disorder in chronic medical conditions, but its effect on physical symptoms has not yet systematically been reviewed. This study aims to do so.

Methods

Systematic review and meta-analysis **was conducted using** Pubmed, the Cochrane Library, and the European and US Clinical Trial Registers. **Eligible studies** included randomized controlled trials (RCTs) of CC compared to care as usual (CAU), in primary care and general hospital setting, reporting on physical and depressive symptoms as outcomes. Overall treatment effects were estimated for illness burden, physical outcomes and depression, **respectively.**

Results

Twenty RCTs were included, with N = 4774 patients. The overall effect size of CC versus CAU for illness burden was OR 1.64 (95%CI 1.47;1.83), d = 0.27 (95%CI 0.21;0.33). Overall, depression outcomes were better for CC than for CAU. Moderator analyses did not yield statistically significant differences.

Conclusions

CC is more effective than CAU in terms of illness burden, physical outcomes and depression, in patients with comorbid depression in chronic medical conditions. More research covering multiple physical disorders **is needed.**

Protocol Registration Number

The protocol for this systematic review and meta-analysis has been registered at the International Prospective Register of Systematic Reviews (PROSPERO) on February 19th 2016: <http://www.crd.york.ac.uk/PROSPERO/DisplayPDF.php?ID=CRD42016035553>.

Keywords

Collaborative care; Depression; Chronic medical disorder; Systematic review; Meta-analysis; Physical outcomes; Illness burden; Comorbidity.

Acronyms used in text

95%CI 95% Confidence Interval

ACS Acute Coronary Syndrome

CAU Care As Usual
CC Collaborative Care
CM Care Manager
COPD Chronic Obstructive Pulmonary Disease
DM Diabetes Mellitus
GP General Practitioner
HIV Human Immunodeficiency Virus
MACE Major Adverse Cardiac Events
RCT Randomized Controlled Trial

1. Introduction

1.1 Rationale

Collaborative care (CC) is a protocolled multi-disciplinary chronic care treatment for depressive and anxiety disorders, originally developed in general health care in the 1970s [1]. It involves a number of health care professionals working **together** with a patient [2]. Besides primary care settings, CC is now also provided in other settings, such as in general hospitals and occupational health care [3-7]. It involves a care manager (CM) with training in the treatment of depression and anxiety disorders, a medical doctor (a general practitioner (GP), medical specialist or a pharmacist) and a psychiatrist [8-10]. The CM has regular contact with the patient and organises care in collaboration with the medical doctor and/or pharmacist, supervised by the psychiatrist the focus is on monitoring of symptoms and adherence to medication and other treatments. **In addition to patient** monitoring, the CM may also provide psychotherapy. The form of psychotherapy that is most often provided in CC is Problem Solving Treatment (PST) [11]. PST is a brief form of evidence-based psychotherapy that addresses problems that cause and maintain psychiatric symptoms.

Although CC **has been** found to have some, although small, effects on physical outcomes in comorbid depression and diabetes [12], **as compared to** CAU, it remains **unclear** what effect CC has on physical outcomes in chronic medical conditions. In general, physical outcomes are not the main focus of such research.

Individuals with chronic medical conditions frequently suffer from comorbid depression or anxiety disorders [13-17]. This comorbidity can easily lead to disability, **such as** in the workplace [18, 19], and high health care costs [20]. In **such cases of** comorbidity, the treatment response is **often** less favourable for both the chronic medical condition [21] and the mental disorder [22-24]. Research has shown that CC is effective, as compared to CAU, in improving depressive and anxiety disorders [25-33], and the National Institute for Health and Care Excellence (NICE) guidelines on depression recommend CC as the treatment of choice in the case of concomitant chronic medical disorders [34]. Previous systematic reviews have examined mental health outcomes in comorbid diabetes and depression [12], cancer and depression [35], and coronary heart disease and depression [36]. However, no systematic review has assessed physical outcomes in comorbid depression and chronic medical conditions. This study aims to do so, with a focus on illness burden and physical symptoms.

1.2 Objectives

Literature reviews to date have focused on the effectiveness of CC for psychological outcomes, rather than for physical outcomes, in patients with comorbid physical **conditions**

and depressive or anxiety disorders [30-32]. Therefore, the objective of this study was to perform a systematic review and meta-analysis that extends beyond previous reviews by reporting the effects of CC on physical disorders and co-morbid depressive or anxiety disorders, with a primary focus on physical outcomes and illness burden. We will explore results for each chronic medical condition separately, as well as multi-morbidity. Results as provided in the trials will be included.

1.3 Hypotheses

We hypothesize that the effect of CC on illness burden and physical outcomes **will be superior to** the effect of CAU. **Further**, we hypothesize that the effect might not be the same for all chronic medical conditions. The effect might be larger when self-management can improve the physical outcomes, such as in case of type II diabetes mellitus, chronic obstructive pulmonary disease (COPD) and multi-morbidity, where patients need to take their medication on a regular basis and make far-reaching lifestyle adaptations, for which CC may be of particular help. **The effect may be weaker** for cancer, **for example**, where most of the treatment is provided in hospital.

We hypothesize that CC will be associated with more somatic diagnostic and treatment interventions, and with lower illness burden, as compared to CAU. We hypothesize that this effect will be larger in patients with somatic multi-morbidity than in patients with one chronic medical condition, as the integrative effect of **CC is predicted to** be most effective. In view of ongoing discussions regarding **whether** CC is more effective in the case of two or three collaborating professionals [37], if CC **is more effective when provided** in the primary care setting or the general hospital setting [9, 38], and if CMs in CC should provide monitoring only or psychotherapeutic treatment **as well** [39, 40], we will explore if CC model, setting and the role of the CM differentiates the effect of CC in patients with depressive or anxiety disorder and comorbid chronic physical disorders.

2. Methods

2.1 Eligibility criteria, information sources and search

The protocol for this systematic review and meta-analysis has been registered at the International Prospective Register of Systematic Reviews (PROSPERO) [41]. The search terms and selection criteria for this review were defined by three **researchers**: J.E.S., H.C. and C.F.C. We identified potentially eligible randomized controlled trials (RCTs) and systematic reviews by searching the PubMed **database** and the Cochrane database of systematic reviews. MeSH terms and free text terms for “somatic”, “depression” and

“collaborative care” were used to search PubMed, with limitations of “RCT” or “systematic review”. We included only articles published in English or Dutch language. Appendix A presents the detailed and complete search strategy. The reference lists of three recent published systematic reviews on CC [30-32] were checked for studies missed by the PubMed search. Further, the search terms of the systematic review [18] were used to update the search up until August 1st 2017.

We also did a search in two clinical trial registers to find ongoing trials. We used the European Union Clinical Trials Register [42] and ‘ClinicalTrials.gov; a service of the U.S. National Institutes of Health’ [43] for this purpose. We searched for interventional studies of CC for depression or anxiety disorder with comorbid chronic medical conditions. Furthermore, we contacted principal investigators of key studies as well as the members of EURONET-SOMA [44] for suggestions about other trials that might be eligible for inclusion.

2.2 Study selection

In this systematic review, we included RCTs or cluster RCTs that evaluated the effect on physical outcomes of collaborative care for patients with physical disorders and co-morbid depression or anxiety disorders. We defined CC a priori as a treatment intervention performed by a multidisciplinary team with at least two of the following three care givers involved: a medical doctor (a GP or medical specialist), a consultant psychiatrist and/or a CM, who is most often a nurse. The trials were required to compare a treatment intervention, meeting the CC definition, with CAU, and assess the effect of CC on both physical symptoms as well as on depressive/anxiety symptoms. The outcomes could be either self-reported or clinician rated. Because CC is used in specialist care settings as well as in primary care, both settings were investigated. Exclusion criteria were:

- Articles that focused on only physical symptoms or only depressive-anxiety symptoms, instead of on both;
- Interventions that did not meet our criteria for CC;
- High risk of bias, indicated by a score of <4 out of 7 in the risk of bias assessment performed according to the Cochrane quality criteria [45].

Two researchers (J.E.S and H.C.) independently and in duplicate judged whether each study met our inclusion criteria based on the title and abstract of the article. After this, three researchers (J.E.S., H.C. and V.E.) independently and in triplicate ascertained whether each study met the inclusion criteria on the basis of the full text. Disagreement regarding inclusion status was discussed and we tried to come to a mutual agreement. If no consensus could be met, a fourth researcher (C.F.C.) was consulted. Consensus was reached in all cases. C.F.C. approached the EURONET-SOMA members for relevant publications. J.E.S. checked the trial registers.

2.3 Data collection process

Relevant outcome analyses were independently extracted from full text articles by three researchers (J.E.S., H.C. and V.E.). Extracted data included the following study characteristics and outcome analyses: setting, mean age, somatic diagnosis, psychiatric diagnosis, treatment model, treatment provided by CM, and main outcome (mean) compared to CAU. We pilot tested the standardized form to check if it was complete. In the case of multiple follow-up assessments, we chose the short-medium term main finding of the study (3-12 months). Authors were contacted to clarify data when there were uncertainties.

2.4 Data items

2.4.1 Type of study design

Randomized controlled trials.

2.4.2 Definition of the treatment interventions included in the review

The treatment intervention had to be carefully examined to assess whether the intervention was in accordance with our definition of CC. The characteristics of the treatment intervention had to be specified in terms of the roles of the GP/medical specialist, psychiatrist and CM.

2.4.3 Patient groups

Trials were included if the patients were aged 18 years or older, and were diagnosed with both a chronic medical condition and a depressive disorder and/or anxiety disorder. Trials were eligible for inclusion if they were conducted in primary care and/or specialist settings. Depressive disorders were classified as minor depressive disorder, major depressive disorder, dysthymic disorder or persistent depressive symptoms. Anxiety disorders were classified as any possible anxiety disorder according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision (DSM-IV-TR) [46]. Both depressive disorders and anxiety disorders had to be assessed by a validated questionnaire with a cut-off score, or by a diagnostic interview.

2.4.4 Outcome definitions

An important difference between our study and earlier reviews is that our study focuses explicitly on illness burden and physical outcomes of CC. Outcomes should define the impact of the intervention on the clinical condition of the patients in terms of physical symptomatology and depressive or anxiety symptom severity. Outcomes should report objective measures of such dimensions, as assessed by validated questionnaires or methods. These questionnaires could either be self-report or clinician rated. Table 1 describes all physical and depressive outcome measures included in the current study.

-Insert Table 1-

2.5 Risk of bias in individual studies

Risk of bias in each trial was assessed in triplicate, at study and outcome level, by three independent assessors (J.E.S., V.E. and H.C.). The assessment was performed according to the Cochrane quality criteria [45], of which randomisation was the most important [45].

2.6 Summary measures

We performed a random effects meta-analysis [47]. We calculated effect sizes for each study using the statistical program Comprehensive Meta-Analysis, version 2 [48]. If studies reported on multiple outcomes, we chose to take the following outcome measures into account for the analysis:

- (1) Physical symptom severity, such as cardiac symptoms, HbA1c, hypertension parameters and pain.
- (2) Somatic diagnostic and intervention procedures such as hospital admission, surgery, diagnostic procedures, etc.
- (3) Physical functioning.
- (4) Depressive symptom severity, i.e., treatment response or remission.

2.7 Synthesis of results and risk of bias across studies

2.7.1 Illness burden

We calculated and pooled effect sizes for both **physical and depressive** outcomes using a weighted mean that should indicate a combined assessment of illness burden. **Illness burden was defined as an** outcome that establishes the impact of the intervention at hand on the general clinical condition of the patients. The clinical meaning of this is envisioned to be somewhat similar to the Clinical Global Impression [49], an assessment that a clinician would make in order to indicate if the general clinical condition of a patient has improved in terms of symptoms and functioning. For this, we used the physical and depressive outcomes mentioned above, and translated these measures to a standardized pooled combined equally weighted effect size, using a statistical program that corrects for possible inflated weights, in the case of multiple outcomes, using equal weighting [48]. The program computes the weighted mean for each study, by taking the average of the outcomes, and the average variance. This computation assumes that the correlation between the outcomes of one study is 1.0. In fact, this correlation is almost certainly lower than 1.0, so this is a conservative approach, aimed at correcting for possible inflated weight caused by multiple outcomes. We gave the outcomes similar weights in the weighted means calculation. **This**

meant that illness burden was pooled consistent with previous research [12]. We calculated an Odds Ratio (OR) with a 95% confidence interval (95%CI).

2.7.2 Physical and depressive outcomes

In the next step, we pooled the effects for physical outcomes and for depressive outcomes separately. Then, we pooled across different chronic medical conditions, and multi-morbidity, as the effect of CC might be of a general nature and thus, comparison of the general effect of CC across several chronic medical conditions would be of high clinical relevance. Disease control parameters were included if they were reported in the study.

In the random-effects meta-analysis [47], between-study heterogeneity was assessed using the Q-statistic, which reflects the observed dispersion [50]. In order to quantify this dispersion, the I^2 statistic was used, which describes the percentage of total variation across the studies that is the result of heterogeneity rather than of chance. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity; I^2 values of 25%, 50% and 75% correspond to low, moderate and high heterogeneity, respectively [50]. Publication bias was examined by constructing a Begg [51] funnel plot and performing the fail-safe N [52]. High-resolution plots are presented. All statistical pooling was conducted using the program Comprehensive Meta-Analysis, version 2. Missing data were not imputed. The findings are presented according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [53]. The PRISMA checklist is included in Appendix B.

2.8 Additional pre-envisioned subgroup analyses

Additional pre-defined analyses were performed for three theoretically interesting potential moderating effects, based on the literature [32], and as described in our hypotheses. As such, we explored if CC model, setting and the role of the CM differentiates the effect of CC in chronic medical conditions with comorbid depressive or anxiety disorder. The effect sizes were calculated and pooled.

Results

3.1 Study selection

The initial search resulted in 1109 hits, from which 199 studies were selected based on the titles. Of those, 63 studies were selected based on the abstracts, and of those studies, 20 met the inclusion criteria based on the full text. The search of the clinical trial registers

yielded one trial that fitted our inclusion criteria, but this was a trial we had already included (see Flow diagram 1) [53].

-Insert Flow diagram-

The total N from these 20 studies was 4774. All studies focused on depression; only one study investigated the combination of depression and anxiety. Because we found only one study that reported physical outcomes of CC compared to CAU for the combination of a chronic medical condition with comorbid anxiety disorder, we will only report the effects of CC for the combination of a chronic medical condition and depressive disorder.

3.2 Study characteristics

In accordance with our inclusion criteria, all studies were RCTs. The sample sizes varied between 55 and 1001 patients. All 20 studies reported on a physical outcome separately. Follow-up periods were short to medium term, i.e., 3-12 months.

3.3 Risk of bias

The risk of bias was rated for all 20 studies; see Table 2. All 20 studies were rated as acceptable (low risk of bias) for inclusion in the meta-analysis. In this review, a cut-off score of at least four out of a possible seven scores was chosen to define a low risk of bias, according to the Cochrane quality criteria [45]. In all 20 studies, blinding of clinical personnel to treatment type provided was not possible due to the study design. However, this is a common feature of psychotherapeutic intervention studies [54].

-Insert Table 2-

3.4 Results of individual studies

The characteristics of the 20 included studies, as well as the effect sizes calculated in our meta-analyses, are shown in Table 3. A variety of chronic medical conditions were reported in the 20 studies: Acute coronary syndrome (ACS) (five studies; n = 866) [55-59], arthritis (two studies; n = 1251) [60, 61], cancer (three studies; n = 727) [62-64], chronic obstructive pulmonary disease (COPD) (one study; n = 138) [65], diabetes mellitus (DM) type II (three studies; n = 774) [66-68], epilepsy (one study; n = 80) [69], human immunodeficiency virus (HIV) (one study; n = 276) [70], hypertension (one study; n = 64) [71], and multi-morbidity (three studies; n = 598) [9, 72, 73]. Fifteen of the 20 studies diagnosed depression or dysthymia using self-report, 12 of the studies used the Patient Health Questionnaire (PHQ-9), two studies used the Beck Depression Inventory (BDI) and

one study used the Centre for Epidemiologic Studies Depression Scale (CES-D). Four studies used the Structured Clinical Interview for DSM-IV to diagnose major or minor depression, and one study use the Mini International Neuropsychiatric Interview (MINI). The mean age of the patients in the studies varied between 44 and 72 years.

-Insert Table 3-

3.5 Synthesis of results

3.5.1 Illness burden

In order to establish the illness burden based on the physical and depressive outcomes, a meta-analysis was performed across the 20 studies, N = 4774. The overall OR was 1.64 (95%CI 1.47;1.83) with an overall effect size of CC of $d = 0.27$, 95%CI 0.21;0.33 **for CC compared to CAU**. The results showed low statistical heterogeneity ($I^2 < 0.1$) and no significant dispersion (Q value of 14.95, df = 19, $p = 0.75$). The Forest plot is shown below in Figure 2.

-Insert Figure 2-

3.5.2 Combined physical outcomes

In order to establish the effect of CC on physical outcomes, a meta-analysis was performed for the combined physical outcomes: HbA1C, number of epileptic seizures, HIV-symptom severity, pain and physical functioning. This yielded 19 studies, total N = 4692. The overall OR for CC compared to CAU was 1.46 (95%CI 1.28;1.67) and the overall effect size of CC was $d = 0.21$, 95%CI 0.14;0.26. The results showed low statistical heterogeneity ($I^2 = 25.21$) and no significant dispersion (Q value of 24.07, df = 18, $p = 0.15$). The **Forest plot is shown in Figure 3**.

-Insert Figure 3-

The results varied depending on the chronic medical condition **investigated**. Results were significant for all chronic medical conditions, except for DM type II and epilepsy. The highest OR (4.18, 95%CI 2.18;8.04; $d = 0.79$, 95%CI 0.43;1.15) was found for hypertension. HIV, COPD, multi-morbidity, arthritis, cancer and ACS showed smaller ORs, with no significant dispersion (Q = 16,854, df = 8, $p = 0.032$). The results are shown in Figure 4.

-Insert Figure 4-

3.5.3 Somatic diagnostic or treatment procedures

Four studies reported on somatic diagnostic or treatment procedures; see the Forest plot in Figure 5. Low heterogeneity was found between the different treatment studies in terms of somatic diagnoses and intervention procedures ($I^2 < 0.01$; Q value = 0.25, df = 3, p = 0.97). Overall, more procedures were performed in CC compared to CAU, with an OR of 2.25 (95%CI 1.67;3.04), d = 0.45 (95%CI 0.28;0.61).

-Insert Figure 5-

3.5.4 Depression outcomes

All studies reported a depression outcome [9, 55-57, 60, 62-66, 69, 70, 72]. There was no statistical heterogeneity across the studies ($I^2 < 0.01$, Q value = 18.63, df = 19, p = 0.48). Results for CC were significantly better than for CAU, with an overall OR of 1.90 (95%CI 1.70;2.12). The overall effect size of CC for the depression outcome was small to moderate (d = 0.35, 95%CI 0.29;0.42). No statistically significant differences in depression outcomes were found across the different physical disorders. With the exception of COPD (based on only one study), all depressive outcomes were statistically significant. The results are shown in Figure 6.

-Insert Figure 6-

3.5.5 Additional analyses: Moderating effects

We selected three potential moderating effects, as indicated in Table 3. The first potential moderating effect was the CC model. CC treatment was either given by a CM in collaboration with a GP (four studies) or by a CM in collaboration with a physician and a psychiatrist (16 studies). The second hypothesised potential moderating effect was the setting in which the CC was provided; hospital care setting (11 studies) or primary care setting (nine studies). The third potential moderating effect was the role of the CM: monitoring only (10 studies) or providing psychotherapy as well as monitoring (10 studies). For all three moderating effects, no statistically significant differences in treatment effects of CC compared to CAU were found for illness burden, physical outcomes, somatic and diagnostic interventions and depression outcomes.

3.6 Risk of bias across studies: Publication bias

Based on the funnel plot (see Figure 6) is it unlikely that any studies that have been missed would change the results, given that we would have to include at least 431 extra studies to find diverging results.

4. Discussion

4.1 Summary of evidence

This is the first systematic review that focusses explicitly on the effects of CC on illness burden and physical outcomes for patients with comorbid depression in chronic medical conditions. We performed a systematic review and meta-analysis of 20 RCTs, and reported overall illness burden, physical outcomes and depressive outcomes for CC compared to CAU, among patients with comorbid depressive disorder and chronic medical conditions. Our results indicate that CC is more effective than CAU in terms of illness burden, physical outcomes and depression outcomes. Although the overall effect sizes for illness burden and combined physical outcomes were small ($d = 0.27$ and 0.21 , respectively), the effect sizes for the specific chronic medical conditions varied considerably. This confirms our first hypothesis and makes a strong case for providing CC in chronic medical conditions with comorbid depression.

Regarding our second hypothesis, the largest statistically significant effect size in terms of chronic medical conditions was found for hypertension, followed by HIV, COPD, multi-morbidity, arthritis, cancer and ACS. The strong effects of CC on hypertension can most likely be explained by the greater adherence of these patients to both antidepressant and antihypertensive medications [71]. However, given that there is only one study in our meta-analysis that investigated hypertension, more studies are needed to draw firm conclusions. We hypothesized that CC would be most effective in chronic medical conditions with high potential for self-management and case-management; our findings partly confirm this. However, our finding that physical outcomes are not significantly better for CC compared to CAU for patients with type II DM shows that self-management may not be enough to improve outcomes for patients with type II DM, especially in terms of HbA1c.

In patients with DM, the presence of multiple complications is significantly associated with comorbid depression [74], meaning that more complex patients are at higher risk for depression. Because of this complexity, both enhanced coordination of care and self-management are important. Both are an important part of CC, and a possible explanation for the success of CC in complex patient groups. For patients with DM, CC has previously been recommended as a method to provide all necessary aspects of depression care, such as psychotherapy and anti-depressant medication [12]. The results of our meta-analysis support this recommendation. We also found that more somatic diagnostic and treatment procedures were performed in cases of CC compared to CAU (OR 2.25). Although more procedures make CC more costly, several studies have shown that CC improves quality of life for

acceptable additional costs [38, 75]. This has been shown to be the case for depression comorbid with various chronic medical conditions [76-81].

In patients with ACS, comorbid depression is associated with poorer quality of life [82]. Therefore, given that CC has a better effect than CAU on depression, the fact that depression treatment in ACS has a known favourable prognostic effect on cardiac outcomes [83], and the fact that CC has a positive effect on physical outcomes, CC should be considered a good treatment option in ACS. CC treatment studies with long-term follow up, and focusing on mortality [84], would be needed.

The effect of CC was better than CAU for the depression outcomes, with an overall small to moderate effect size ($d = 0.35$); there was no statistically significant difference between the chronic medical conditions. It is possible that, even when chronic medical conditions show limited potential for improvement under a generic behavioural intervention -such as self-management- as in CC, depressive symptoms can still be further improved by CC compared to CAU. For example, previous research has shown that patients with epilepsy in usual care may not receive the depression care they require [85]. Given that they suffer from fear of seizures [86], are often stigmatized [87], and often have comorbid depression [88], adequate depression care is important. In view of the limited availability of RCTs exploring treatment of comorbid depression in epilepsy [89], the findings of this study suggest that CC may be a good candidate for further research and implementation in epilepsy patient care.

In our analyses we explored three potential moderating effects. Firstly, the CC model: treatment given by a CM in collaboration with a GP versus a CM in collaboration with a physician and a psychiatrist. Secondly, the setting in which the CC was provided: hospital care setting or primary care setting. And thirdly, the role of the CM: providing only monitoring or providing monitoring and psychotherapy. None of these showed statistically significant differences for physical and depressive treatment effects. It appears that, for improvement in both physical and depressive symptoms, it is more important that the CC model is applied in the first place, than who applies it or whether it is provided in a primary care or general hospital setting. Therefore, addressing both physical and mental health in monitoring, counselling and consultation can be seen as important factors for the success of CC. This is supported by the finding that a reduction in physical symptoms is associated with a reduction in depression [90].

The National Institute for Health and Care Excellence (NICE) states that CC is more (cost) effective than CAU for depression outcomes [34], among patients with depression and a chronic physical health problem. One of the recommendations of the Guideline Development Group of NICE was to research the effectiveness of CC on physical health outcomes for patients with depression and a chronic physical health problem. This review and meta-analysis gives an overview of the RCTs that have done so, using physical

outcomes as the primary outcome measure, and thereby adds to the body of knowledge about the effectiveness of CC for this complex group of patients with both physical and mental problems. Over half of the RCTs we identified in this review were published in 2010 or later, showing the recent interest in this topic.

4.2 Limitations

For our search we used the Pubmed database, the Cochrane database, and the reference lists of selected systematic reviews; only studies in English and Dutch **were included**. We also did a search in an European and an American trial register, and contacted principal investigators of key studies, as well as the members of EURONET-SOMA, for suggestions about other **relevant** trials. Even though we performed this rather extensive search, it is possible that we missed studies. However, the high Failsafe numbers (431) that we found, provide some reassurance.

The small number of good quality studies **for each** chronic medical condition covered in this review is a limitation. To provide better estimates for specific chronic medical conditions, more research is needed. Those studies could provide more information about the content of the CAU that is **implemented**, which is lacking in the current studies.

There is also a lack of research on physical outcomes in chronic medical conditions; **most studies instead focus on** only anxiety outcomes [91, 92]. Therefore, we cannot draw conclusions about this. This is an interesting subject for further research, because the combination of anxiety disorders and chronic medical conditions is highly prevalent [13, 93]; the presence of an anxiety disorder leads to a higher likelihood of disability in chronic medical conditions [15], and a more negative perception of the influence of the chronic medical condition [94]. **It has also been** suggested that the presence of physical symptoms negatively influences the course of the anxiety disorder [22].

4.3 Strengths

To our knowledge, this is the first study that reports on the effectiveness of CC for comorbid depression and chronic medical conditions; **CC can be viewed as** a generic intervention that can be applied in a variety of chronic medical conditions with comorbid depression, or in the case of multi-morbidity. **The current study's** focus on physical outcomes is an important **strength of this work**. In all studies, CAU was provided to the control group. In addition, statistical heterogeneity (I^2) of the results of the **studies included in our meta-analysis** was low to moderate. Further, based on the funnel plot, it is unlikely that studies that would change the results have been missed. Therefore, the results can probably be considered robust. However, the effect sizes were still comparatively small so we may need

more personalised treatment models for patients with somatic-psychiatric comorbidity; this is a potential focus for further research in the future [95-97].

4.4 Conclusions

In conclusion, CC is more effective than CAU in terms of illness burden and physical outcomes for patients with comorbid depressive disorder and chronic medical conditions, particularly in patients with hypertension, HIV, COPD, multi-morbidity, arthritis, cancer and ACS. CC is associated with more diagnostic and treatment intervention procedures, and hence, may play a role in improving quality of life. Compared to CAU, CC is also effective for depression outcomes in patients with chronic medical conditions. More research **investigating multiple chronic medical conditions**, with longer follow-up, and taking mortality into account, **is needed in the future.**

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

- [1] Ezedum S, Kerr DN. Collaborative care of hypertensives, using a shared record. *Br Med J* 1977;2(6099):1402-3.
- [2] Vlasveld MC, Anema JR, Beekman AT, van Mechelen W, Hoedeman R, van Marwijk HW, et al. Multidisciplinary collaborative care for depressive disorder in the occupational health setting: Design of a randomised controlled trial and cost-effectiveness study. *BMC Health Serv Res* 2008;8:99.
- [3] Van der Feltz-Cornelis CM. The depression initiative. Description of a collaborative care model for depression in the primary care setting in the Netherlands. *Clinical Neuropsychiatry* 2011;8(4).
- [4] Walker J, Sharpe M. Depression care for people with cancer: A collaborative care intervention. *Gen Hosp Psychiatry* 2009;31(5):436-41.
- [5] Chwastiak LA, Jackson SL, Russo J, DeKeyser P, Kiefer M, Belyeu B, et al. A collaborative care team to integrate behavioral health care and treatment of poorly-controlled type 2 diabetes in an urban safety net primary care clinic. *Gen Hosp Psychiatry* 2017;44:10-5.
- [6] Beach SR, Walker J, Celano CM, Mastromauro CA, Sharpe M, Huffman JC. Implementing collaborative care programs for psychiatric disorders in medical settings: A practical guide. *Gen Hosp Psychiatry* 2015;37(6):522-7.
- [7] Rossom RC, Solberg LI, Magnan S, Crain AL, Beck A, Coleman KJ, et al. Impact of a national collaborative care initiative for patients with depression and diabetes or cardiovascular disease. *Gen Hosp Psychiatry* 2017;44:77-85.
- [8] Gilbody S, Bower P, Fletcher J, Richards D, Sutton AJ. Collaborative care for depression: A cumulative meta-analysis and review of longer-term outcomes. *Arch Intern Med* 2006;166(21):2314-21.
- [9] van Steenberghe-Weijenburg KM, van der Feltz-Cornelis CM, van Benthem TB, Horn EK, Ploeger R, Brals JW, et al. [Collaborative care for co-morbid major depressive disorder in chronically ill outpatients in a general hospital]. *Tijdschr Psychiatr* 2015;57(4):248-57.
- [10] Carter BL. Primary care physician-pharmacist collaborative care model: Strategies for implementation. *Pharmacotherapy* 2016;36(4):363-73.
- [11] Mynors-Wallis LM, Gath DH, Day A, Baker F. Randomised controlled trial of problem solving treatment, antidepressant medication, and combined treatment for major depression in primary care. *BMJ* 2000;320(7226):26-30.
- [12] van der Feltz-Cornelis CM, Nuyen J, Stoop C, Chan J, Jacobson AM, Katon W, et al. Effect of interventions for major depressive disorder and significant depressive

- symptoms in patients with diabetes mellitus: A systematic review and meta-analysis. *Gen Hosp Psychiatry* 2010;32(4):380-95.
- [13] van Eck van der Sluijs J, Ten Have M, Rijnders C, van Marwijk H, de Graaf R, van der Feltz-Cornelis C. Medically unexplained and explained physical symptoms in the general population: Association with prevalent and incident mental disorders. *PloS one* 2015;10(4):e0123274.
 - [14] Penninx BW, van Dyck R. [Depression and somatic comorbidity]. *Ned Tijdschr Geneesk* 2010;154:A1784.
 - [15] Sareen J, Cox BJ, Clara I, Asmundson GJ. The relationship between anxiety disorders and physical disorders in the U.S. National Comorbidity Survey. *Depress Anxiety* 2005;21(4):193-202.
 - [16] Chong SA, Abidin E, Nan L, Vaingankar JA, Subramaniam M. Prevalence and impact of mental and physical comorbidity in the adult Singapore population. *Ann Acad Med Singapore* 2012;41(3):105-14.
 - [17] Hsiang H, Karen MT, Joseph MC, Nahida A, Amritha B, Rachel K. Collaborative care for women with depression: A systematic review. *Psychosomatics* 2017;58(1):11-8.
 - [18] Hoedeman R, Krol B, Blankenstein N, Koopmans PC, Groothoff JW. Severe MUPS in a sick-listed population: A cross-sectional study on prevalence, recognition, psychiatric co-morbidity and impairment. *BMC Public Health* 2009;9:440.
 - [19] Dewa CS, Lin E. Chronic physical illness, psychiatric disorder and disability in the workplace. *Soc Sci Med* 2000;51(1):41-50.
 - [20] Goetzel RZ, Long SR, Ozminkowski RJ, Hawkins K, Wang S, Lynch W. Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting U.S. employers. *J Occup Environ Med* 2004;46(4):398-412.
 - [21] Kang HJ, Kim SY, Bae KY, Kim SW, Shin IS, Yoon JS, et al. Comorbidity of depression with physical disorders: Research and clinical implications. *Chonnam Med J* 2015;51(1):8-18.
 - [22] Huijbregts KM, van Marwijk HW, de Jong FJ, Schreuders B, Beekman AT, van der Feltz-Cornelis CM. Adverse effects of multiple physical symptoms on the course of depressive and anxiety symptoms in primary care. *Psychother Psychosom* 2010;79(6):389-91.
 - [23] Huijbregts KM, van der Feltz-Cornelis CM, van Marwijk HW, de Jong FJ, van der Windt DA, Beekman AT. Negative association of concomitant physical symptoms with the course of major depressive disorder: A systematic review. *Journal of psychosomatic research* 2010;68(6):511-9.

- [24] Huijbregts KM, de Jong FJ, van Marwijk HW, Beekman AT, Ader HJ, van der Feltz-Cornelis CM. A high physical symptom count reduces the effectiveness of treatment for depression, independently of chronic medical conditions. *Journal of psychosomatic research* 2013;74(3):179-85.
- [25] de Jong FJ, van Steenbergen-Weijenburg KM, Huijbregts KM, Vlasveld MC, Van Marwijk HW, Beekman AT, et al. The Depression Initiative. Description of a collaborative care model for depression and of the factors influencing its implementation in the primary care setting in the Netherlands. *Int J Integr Care* 2009;9:e81.
- [26] Jacob V, Chattopadhyay SK, Sipe TA, Thota AB, Byard GJ, Chapman DP, et al. Economics of collaborative care for management of depressive disorders: A community guide systematic review. *Am J Prev Med* 2012;42(5):539-49.
- [27] Goorden M, Huijbregts KM, van Marwijk HW, Beekman AT, van der Feltz-Cornelis CM, Hakkaart-van Roijen L. Cost-utility of collaborative care for major depressive disorder in primary care in the Netherlands. *Journal of psychosomatic research* 2015;79(4):316-23.
- [28] Huijbregts KM, de Jong FJ, van Marwijk HW, Beekman AT, Ader HJ, Hakkaart-van Roijen L, et al. 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. *J Affect Disord* 2013;146(3):328-37.
- [29] Wetherell JL, Petkus AJ, Thorp SR, Stein MB, Chavira DA, Campbell-Sills L, et al. Age differences in treatment response to a collaborative care intervention for anxiety disorders. *Br J Psychiatry* 2013;203(1):65-72.
- [30] Archer J, Bower P, Gilbody S, Lovell K, Richards D, Gask L, et al. Collaborative care for depression and anxiety problems. *Cochrane Database Syst Rev* 2012;10:CD006525.
- [31] Woltmann E, Grogan-Kaylor A, Perron B, Georges H, Kilbourne AM, Bauer MS. Comparative effectiveness of collaborative chronic care models for mental health conditions across primary, specialty, and behavioral health care settings: Systematic review and meta-analysis. *Am J Psychiatry* 2012;169(8):790-804.
- [32] Baumeister H, Hutter N. Collaborative care for depression in medically ill patients. *Curr Opin Psychiatry* 2012;25(5):405-14.
- [33] Muntingh AD, van der Feltz-Cornelis CM, van Marwijk HW, Spinhoven P, van Balkom AJ. Collaborative care for anxiety disorders in primary care: A systematic review and meta-analysis. *BMC Family Practice* 2016;17(1):62.
- [34] NICE. Depression in Adults with a chronic physical health problem: Treatment and Management. Leicester: British Psychological Society; 2010.

- [35] Li M, Kennedy EB, Byrne N, Gerin-Lajoie C, Katz MR, Keshavarz H, et al. Systematic review and meta-analysis of collaborative care interventions for depression in patients with cancer. *Psychooncology* 2016.
- [36] Tully PJ, Baumeister H. Collaborative care for comorbid depression and coronary heart disease: A systematic review and meta-analysis of randomised controlled trials. *BMJ Open* 2015;5(12):e009128.
- [37] Katon W, Guico-Pabia CJ. Improving quality of depression care using organized systems of care: A review of the literature. *The primary care companion for CNS disorders* 2011;13(1).
- [38] van Steenberg-Weijnenburg KM, van der Feltz-Cornelis CM, Horn EK, van Marwijk HW, Beekman AT, Rutten FF, et al. Cost-effectiveness of collaborative care for the treatment of major depressive disorder in primary care. A systematic review. *BMC Health Serv Res* 2010;10:19.
- [39] Wanat M, Walker J, Hodges L, Richardson A, Sharpe M. Selecting, training and supervising nurses to treat depression in the medically ill: Experience and recommendations from the SMaRT oncology collaborative care trials. *Gen Hosp Psychiatry* 2015;37(6):518-21.
- [40] Huffman JC, Niazi SK, Rundell JR, Sharpe M, Katon WJ. Essential articles on collaborative care models for the treatment of psychiatric disorders in medical settings: A publication by the academy of psychosomatic medicine research and evidence-based practice committee. *Psychosomatics* 2014;55(2):109-22.
- [41] <http://www.crd.york.ac.uk/PROSPERO/DisplayPDF.php?ID=CRD42016035553>.
- [42] <https://www.clinicaltrialsregister.eu>.
- [43] <https://clinicaltrials.gov>.
- [44] Weigel A, Husing P, Kohlmann S, Lehmann M, Shedden-Mora M, Toussaint A, et al. A European research network to improve diagnosis, treatment and care for patients with persistent somatic symptoms: Work report of the EURONET-SOMA conference series. *Journal of psychosomatic research* 2017;97:136-8.
- [45] The Cochrane Collaboration. *Cochrane handbook for systematic reviews of interventions*. 2011.
- [46] American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th ed. text revision. DC: Author; 2000.
- [47] DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled clinical trials* 1986;7(3):177-88.
- [48] Borenstein MHL, Higgins JPT, Rothstein HR. *Comprehensive meta-analysis*. Englewood (NJ): Biostat; 2005.

- [49] Guy W. ECDEU assessment manual for psychopharmacology—Revised. DHEW Publ No ADM 76-338. Rockville (MD); 1976.
- [50] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-60.
- [51] Cooper H, Hedges LV (Eds). *The handbook of research synthesis*. New York: Russel Sage Foundation; 1994.
- [52] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to meta-analysis (statistics in practice)*. Chichester (UK): Wiley & Sons; 2009.
- [53] Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 2009;6(7):e1000097.
- [54] Berger D. Double-blinding and bias in medication and cognitive-behavioral therapy trials for major depressive disorder. *F1000Research* 2015;4:638.
- [55] Davidson KW, Rieckmann N, Clemow L, Schwartz JE, Shimbo D, Medina V, et al. Enhanced depression care for patients with acute coronary syndrome and persistent depressive symptoms: Coronary psychosocial evaluation studies randomized controlled trial. *Arch Intern Med* 2010;170(7):600-8.
- [56] Huffman JC, Mastromauro CA, Sowden G, Fricchione GL, Healy BC, Januzzi JL. Impact of a depression care management program for hospitalized cardiac patients. *Circ Cardiovasc Qual Outcomes* 2011;4(2):198-205.
- [57] Huffman JC, Mastromauro CA, Beach SR, Celano CM, DuBois CM, Healy BC, et al. Collaborative care for depression and anxiety disorders in patients with recent cardiac events: The Management of Sadness and Anxiety in Cardiology (MOSAIC) randomized clinical trial. *JAMA Intern Med* 2014;174(6):927-35.
- [58] Davidson KW, Bigger JT, Burg MM, Carney RM, Chaplin WF, Czajkowski S, et al. Centralized, stepped, patient preference-based treatment for patients with post-acute coronary syndrome depression: CODIACS vanguard randomized controlled trial. *JAMA Intern Med* 2013;173(11):997-1004.
- [59] Carney RM, Freedland KE, Steinmeyer BC, Rubin EH, Ewald G. Collaborative care for depression symptoms in an outpatient cardiology setting: A randomized clinical trial. *Int J Cardiol* 2016;219:164-71.
- [60] Lin EH, Katon W, Von Korff M, Tang L, Williams JW, Jr., Kroenke K, et al. Effect of improving depression care on pain and functional outcomes among older adults with arthritis: A randomized controlled trial. *JAMA* 2003;290(18):2428-9.
- [61] Kroenke K, Bair MJ, Damush TM, Wu J, Hoke S, Sutherland J, et al. Optimized antidepressant therapy and pain self-management in primary care patients with

- depression and musculoskeletal pain: A randomized controlled trial. *JAMA* 2009;301(20):2099-110.
- [62] Dwight-Johnson M, Ell K, Lee PJ. Can collaborative care address the needs of low-income Latinas with comorbid depression and cancer? Results from a randomized pilot study. *Psychosomatics* 2005;46(3):224-32.
 - [63] Ell K, Xie B, Quon B, Quinn DI, Dwight-Johnson M, Lee PJ. Randomized controlled trial of collaborative care management of depression among low-income patients with cancer. *J Clin Oncol* 2008;26(27):4488-96.
 - [64] Strong V, Waters R, Hibberd C, Murray G, Wall L, Walker J, et al. Management of depression for people with cancer (SMaRT oncology 1): A randomised trial. *Lancet* 2008;372(9632):40-8.
 - [65] Alexopoulos GS, Kiess DN, Sirey JA, Kanellopoulos D, Seirup JK, Novitch RS, et al. Untangling therapeutic ingredients of a personalized intervention for patients with depression and severe COPD. *Am J Geriatr Psychiatry* 2014;22(11):1316-24.
 - [66] Katon WJ, Von Korff M, Lin EH, Simon G, Ludman E, Russo J, et al. The Pathways Study: A randomized trial of collaborative care in patients with diabetes and depression. *Arch Gen Psychiatry* 2004;61(10):1042-9.
 - [67] Bogner HR, de Vries HF. Integrating type 2 diabetes mellitus and depression treatment among African Americans: A randomized controlled pilot trial. *Diabetes Educ* 2010;36(2):284-92.
 - [68] Ell K, Katon W, Xie B, Lee PJ, Kapetanovic S, Guterman J, et al. Collaborative care management of major depression among low-income, predominantly Hispanic subjects with diabetes: A randomized controlled trial. *Diabetes Care* 2010;33(4):706-13.
 - [69] Ciechanowski P, Chaytor N, Miller J, Fraser R, Russo J, Unutzer J, et al. PEARLS depression treatment for individuals with epilepsy: A randomized controlled trial. *Epilepsy Behav* 2010;19(3):225-31.
 - [70] Pyne JM, Fortney JC, Curran GM, Tripathi S, Atkinson JH, Kilbourne AM, et al. Effectiveness of collaborative care for depression in human immunodeficiency virus clinics. *Arch Intern Med* 2011;171(1):23-31.
 - [71] Bogner HR, de Vries HF. Integration of depression and hypertension treatment: A pilot, randomized controlled trial. *Ann Fam Med* 2008;6(4):295-301.
 - [72] Rollman BL, Belnap BH, LeMenager MS, Mazumdar S, Houck PR, Counihan PJ, et al. Telephone-delivered collaborative care for treating post-CABG depression: A randomized controlled trial. *JAMA* 2009;302(19):2095-103.

- [73] Katon WJ, Lin EH, Von Korff M, Ciechanowski P, Ludman EJ, Young B, et al. Collaborative care for patients with depression and chronic illnesses. *N Engl J Med* 2010;363(27):2611-20.
- [74] van Steenbergen-Weijenburg KM, van Puffelen AL, Horn EK, Nuyen J, van Dam PS, van Benthem TB, et al. More co-morbid depression in patients with type 2 diabetes with multiple complications. An observational study at a specialized outpatient clinic. *Diabet Med* 2011;28(1):86-9.
- [75] Katon W, Russo J, Lin EH, Schmittdiel J, Ciechanowski P, Ludman E, et al. Cost-effectiveness of a multicondition collaborative care intervention: A randomized controlled trial. *Arch Gen Psychiatry* 2012;69(5):506-14.
- [76] Camacho EM, Ntais D, Coventry P, Bower P, Lovell K, Chew-Graham C, et al. Long-term cost-effectiveness of collaborative care (vs usual care) for people with depression and comorbid diabetes or cardiovascular disease: A Markov model informed by the COINCIDE randomised controlled trial. *BMJ Open* 2016;6(10):e012514.
- [77] Celano CM, Healy B, Suarez L, Levy DE, Mastromauro C, Januzzi JL, et al. Cost-effectiveness of a collaborative care depression and anxiety treatment program in patients with acute cardiac illness. *Value Health* 2016;19(2):185-91.
- [78] Duarte A, Walker J, Walker S, Richardson G, Holm Hansen C, Martin P, et al. Cost-effectiveness of integrated collaborative care for comorbid major depression in patients with cancer. *Journal of Psychosomatic Research* 2015;79(6):465-70.
- [79] Johnson JA, Lier DA, Soprovich A, Al Sayah F, Qiu W, Majumdar SR. Cost-effectiveness evaluation of collaborative care for diabetes and depression in primary care. *Am J Prev Med* 2016;51(1):e13-20.
- [80] Painter JT, Fortney JC, Gifford AL, Rimland D, Monson T, Rodriguez-Barradas MC, et al. Cost-effectiveness of collaborative care for depression in HIV clinics. *J Acquir Immune Defic Syndr* 2015;70(4):377-85.
- [81] Goorden M, van der Feltz-Cornelis CM, van Steenbergen-Weijenburg KM, Horn EK, Beekman AT, Hakkaart-van Roijen L. Cost-utility of collaborative care for the treatment of comorbid major depressive disorder in outpatients with chronic physical conditions. A randomized controlled trial in the general hospital setting (CC-DIM). *Neuropsychiatric disease and treatment* 2017;13:1881-93.
- [82] Kim JM, Stewart R, Bae KY, Kang HJ, Kim SW, Shin IS, et al. Effects of depression co-morbidity and treatment on quality of life in patients with acute coronary syndrome: The Korean depression in ACS (K-DEPACS) and the escitalopram for depression in ACS (EsDEPACS) study. *Psychol Med* 2015;45(8):1641-52.

- [83] de Jonge P, Honig A, van Melle JP, Schene AH, Kuyper AM, Tulner D, et al. Nonresponse to treatment for depression following myocardial infarction: Association with subsequent cardiac events. *Am J Psychiatry* 2007;164(9):1371-8.
- [84] Miloyan B, Fried E. A reassessment of the relationship between depression and all-cause mortality in 3,604,005 participants from 293 studies. *World Psychiatry* 2017;16(2):219-20.
- [85] Hermann BP, Seidenberg M, Bell B. Psychiatric comorbidity in chronic epilepsy: Identification, consequences, and treatment of major depression. *Epilepsia* 2000;41(Suppl 2):S31-41.
- [86] Beyenburg S, Mitchell AJ, Schmidt D, Elger CE, Reuber M. Anxiety in patients with epilepsy: Systematic review and suggestions for clinical management. *Epilepsy Behav* 2005;7(2):161-71.
- [87] England MJ, Austin JK, Beck V, Escoffery C, Hesdorffer DC. Erasing epilepsy stigma: Eight key messages. *Health Promot Pract* 2014;15(3):313-8.
- [88] Gaus V, Kiep H, Holtkamp M, Burkert S, Kendel F. Gender differences in depression, but not in anxiety in people with epilepsy. *Seizure* 2015;32:37-42.
- [89] van der Feltz-Cornelis C. Treatment of interictal psychiatric disorder in epilepsy. I. Affective and anxiety disorders. *Acta Neuropsychiatrica* 2002;14(1):39-43.
- [90] Chowdhury S, Burton C. Associations of treatment effects between follow-up times and between outcome domains in interventions for somatoform disorders: Review of three Cochrane reviews. *Journal of Psychosomatic Research* 2017;98:10-8.
- [91] Muntingh A, van der Feltz-Cornelis C, van Marwijk H, Spinhoven P, Assendelft W, de Waal M, et al. Effectiveness of collaborative stepped care for anxiety disorders in primary care: A pragmatic cluster randomised controlled trial. *Psychother Psychosom* 2014;83(1):37-44.
- [92] Kronish IM, Chaplin WF, Rieckmann N, Burg MM, Davidson KW. The effect of enhanced depression care on anxiety symptoms in acute coronary syndrome patients: Findings from the COPES trial. *Psychother Psychosom* 2012;81(4):245-7.
- [93] Walters P, Barley EA, Mann A, Phillips R, Tylee A. Depression in primary care patients with coronary heart disease: Baseline findings from the UPBEAT UK study. *PloS one* 2014;9(6):e98342.
- [94] Smith A, Fortune Z, Phillips R, Walters P, Lee GA, Mann A, et al. UPBEAT study patients' perceptions of the effect of coronary heart disease on their lives: A cross-sectional sub-study. *Int J Nurs Stud* 2014;51(11):1500-6.
- [95] Elfeddali I, van der Feltz-Cornelis CM, van Os J, Knappe S, Vieta E, Wittchen HU, et al. Horizon 2020 priorities in clinical mental health research: Results of a consensus-

- based ROAMER expert survey. *International Journal of Environmental Research and Public Health* 2014;11(10):10915-39.
- [96] van der Feltz-Cornelis CM, van Os J, Knappe S, Schumann G, Vieta E, Wittchen HU, et al. Towards Horizon 2020: Challenges and advances for clinical mental health research - Outcome of an expert survey. *Neuropsychiatric Disease and Treatment* 2014;10:1057-68.
- [97] Haro JM, Ayuso-Mateos JL, Bitter I, Demotes-Mainard J, Leboyer M, Lewis SW, et al. ROAMER: Roadmap for mental health research in Europe. *Int J Methods Psychiatr Res* 2014;23(Suppl 1):1-14.

Illness burden and physical outcomes associated with collaborative care in patients with comorbid depressive disorder in chronic medical conditions: a systematic review and meta-analysis

Abbreviated running title: Collaborative care: review and meta-analysis

Jonna van Eck van der Sluijs, MD^{1,2,3}; Hilde Castelijns, MD⁴; Vera Eijsbroek, MD³; Cees Rijnders, MD, PhD³; Prof. Harm van Marwijk, ^{5,6}; Prof. Christina van der Feltz-Cornelis^{1,2,*}

1 Clinical Centre of excellence for Body, Mind and Health, GGz Breburg, Tilburg, The Netherlands

2 Tranzo department, Tilburg University, Tilburg, the Netherlands

3 Department of Residency training, GGz Breburg, Tilburg, The Netherlands

4 Centre for mental health care, PsyQ Tilburg-Parnassia Groep, Tilburg, The Netherlands

5 Centre for Primary Care, Institute of Population Health, University of Manchester, Manchester, United Kingdom

6 Department of General Practice & Elderly Care Medicine and the EMGO+-Institute for Health and Care Research of VU University medical centre (VUmc), Amsterdam, The Netherlands

*Corresponding author. Clinical Centre of excellence for Body, Mind and Health, GGz Breburg, P.O. Box 770, 5000AT Tilburg, The Netherlands. Tel +31(0)88-0161555; Fax +31(0)88-0161699; Email C.M.vdrFeltz@uvt.nl

Total number of words: 8050 (Introduction through Discussion: 4816)

Number of tables: 3

Number of figures: 7

Conflicts of Interest:

JES, HC, VE, CR and HvM declare that they have no conflicts of interest. In the last 3 years, CFC received unrestricted grants from Eli Lilly and from ACHMEA for unrelated research.

Illness burden and physical outcomes associated with collaborative care in patients
A systematic review and meta-analysis

Abbreviated running title: Collaborative care: Review and meta-analysis

Total number of words: 8111 (Introduction through to Discussion: 4908)

Number of Tables: 3

Number of Figures: 7

Abstract

Objective

Collaborative care (CC) improves depressive symptoms in people with comorbid depressive disorder in chronic medical conditions, but its effect on physical symptoms has not yet systematically been reviewed. This study aims to do so.

Methods

Systematic review and meta-analysis was conducted using Pubmed, the Cochrane Library, and the European and US Clinical Trial Registers. Eligible studies included randomized controlled trials (RCTs) of CC compared to care as usual (CAU), in primary care and general hospital setting, reporting on physical and depressive symptoms as outcomes. Overall treatment effects were estimated for illness burden, physical outcomes and depression, respectively.

Results

Twenty RCTs were included, with N = 4774 patients. The overall effect size of CC versus CAU for illness burden was OR 1.64 (95%CI 1.47;1.83), d = 0.27 (95%CI 0.21;0.33). Overall, depression outcomes were better for CC than for CAU. Moderator analyses did not yield statistically significant differences.

Conclusions

CC is more effective than CAU in terms of illness burden, physical outcomes and depression, in patients with comorbid depression in chronic medical conditions. More research covering multiple physical disorders is needed.

Protocol Registration Number

The protocol for this systematic review and meta-analysis has been registered at the International Prospective Register of Systematic Reviews (PROSPERO) on February 19th 2016: <http://www.crd.york.ac.uk/PROSPERO/DisplayPDF.php?ID=CRD42016035553>.

Keywords

Collaborative care; Depression; Chronic medical disorder; Systematic review; Meta-analysis; Physical outcomes; Illness burden; Comorbidity.

Acronyms used in text

95%CI 95% Confidence Interval

ACS Acute Coronary Syndrome

CAU Care As Usual
CC Collaborative Care
CM Care Manager
COPD Chronic Obstructive Pulmonary Disease
DM Diabetes Mellitus
GP General Practitioner
HIV Human Immunodeficiency Virus
MACE Major Adverse Cardiac Events
RCT Randomized Controlled Trial

1. Introduction

1.1 Rationale

Collaborative care (CC) is a protocolled multi-disciplinary chronic care treatment for depressive and anxiety disorders, originally developed in general health care in the 1970s [1]. It involves a number of health care professionals working together with a patient [2]. Besides primary care settings, CC is now also provided in other settings, such as in general hospitals and occupational health care [3-7]. It involves a care manager (CM) with training in the treatment of depression and anxiety disorders, a medical doctor (a general practitioner (GP), medical specialist or a pharmacist) and a psychiatrist [8-10]. The CM has regular contact with the patient and organises care in collaboration with the medical doctor and/or pharmacist, supervised by the psychiatrist the focus is on monitoring of symptoms and adherence to medication and other treatments. In addition to patient monitoring, the CM may also provide psychotherapy. The form of psychotherapy that is most often provided in CC is Problem Solving Treatment (PST) [11]. PST is a brief form of evidence-based psychotherapy that addresses problems that cause and maintain psychiatric symptoms.

Although CC has been found to have some, although small, effects on physical outcomes in comorbid depression and diabetes [12], as compared to CAU, it remains unclear what effect CC has on physical outcomes in chronic medical conditions. In general, physical outcomes are not the main focus of such research.

Individuals with chronic medical conditions frequently suffer from comorbid depression or anxiety disorders [13-17]. This comorbidity can easily lead to disability, such as in the workplace [18, 19], and high health care costs [20]. In such cases of comorbidity, the treatment response is often less favourable for both the chronic medical condition [21] and the mental disorder [22-24]. Research has shown that CC is effective, as compared to CAU, in improving depressive and anxiety disorders [25-33], and the National Institute for Health and Care Excellence (NICE) guidelines on depression recommend CC as the treatment of choice in the case of concomitant chronic medical disorders [34]. Previous systematic reviews have examined mental health outcomes in comorbid diabetes and depression [12], cancer and depression [35], and coronary heart disease and depression [36]. However, no systematic review has assessed physical outcomes in comorbid depression and chronic medical conditions. This study aims to do so, with a focus on illness burden and physical symptoms.

1.2 Objectives

Literature reviews to date have focused on the effectiveness of CC for psychological outcomes, rather than for physical outcomes, in patients with comorbid physical conditions

and depressive or anxiety disorders [30-32]. Therefore, the objective of this study was to perform a systematic review and meta-analysis that extends beyond previous reviews by reporting the effects of CC on physical disorders and co-morbid depressive or anxiety disorders, with a primary focus on physical outcomes and illness burden. We will explore results for each chronic medical condition separately, as well as multi-morbidity. Results as provided in the trials will be included.

1.3 Hypotheses

We hypothesize that the effect of CC on illness burden and physical outcomes will be superior to the effect of CAU. Further, we hypothesize that the effect might not be the same for all chronic medical conditions. The effect might be larger when self-management can improve the physical outcomes, such as in case of type II diabetes mellitus, chronic obstructive pulmonary disease (COPD) and multi-morbidity, where patients need to take their medication on a regular basis and make far-reaching lifestyle adaptations, for which CC may be of particular help. The effect may be weaker for cancer, for example, where most of the treatment is provided in hospital.

We hypothesize that CC will be associated with more somatic diagnostic and treatment interventions, and with lower illness burden, as compared to CAU. We hypothesize that this effect will be larger in patients with somatic multi-morbidity than in patients with one chronic medical condition, as the integrative effect of CC is predicted to be most effective. In view of ongoing discussions regarding whether CC is more effective in the case of two or three collaborating professionals [37], if CC is more effective when provided in the primary care setting or the general hospital setting [9, 38], and if CMs in CC should provide monitoring only or psychotherapeutic treatment as well [39, 40], we will explore if CC model, setting and the role of the CM differentiates the effect of CC in patients with depressive or anxiety disorder and comorbid chronic physical disorders.

2. Methods

2.1 Eligibility criteria, information sources and search

The protocol for this systematic review and meta-analysis has been registered at the International Prospective Register of Systematic Reviews (PROSPERO) [41]. The search terms and selection criteria for this review were defined by three researchers: J.E.S., H.C. and C.F.C. We identified potentially eligible randomized controlled trials (RCTs) and systematic reviews by searching the PubMed database and the Cochrane database of systematic reviews. MeSH terms and free text terms for “somatic”, “depression” and

“collaborative care” were used to search PubMed, with limitations of “RCT” or “systematic review”. We included only articles published in English or Dutch language. Appendix A presents the detailed and complete search strategy. The reference lists of three recent published systematic reviews on CC [30-32] were checked for studies missed by the PubMed search. Further, the search terms of the systematic review [18] were used to update the search up until August 1st 2017.

We also did a search in two clinical trial registers to find ongoing trials. We used the European Union Clinical Trials Register [42] and ‘ClinicalTrials.gov; a service of the U.S. National Institutes of Health’ [43] for this purpose. We searched for interventional studies of CC for depression or anxiety disorder with comorbid chronic medical conditions. Furthermore, we contacted principal investigators of key studies as well as the members of EURONET-SOMA [44] for suggestions about other trials that might be eligible for inclusion.

2.2 Study selection

In this systematic review, we included RCTs or cluster RCTs that evaluated the effect on physical outcomes of collaborative care for patients with physical disorders and co-morbid depression or anxiety disorders. We defined CC a priori as a treatment intervention performed by a multidisciplinary team with at least two of the following three care givers involved: a medical doctor (a GP or medical specialist), a consultant psychiatrist and/or a CM, who is most often a nurse. The trials were required to compare a treatment intervention, meeting the CC definition, with CAU, and assess the effect of CC on both physical symptoms as well as on depressive/anxiety symptoms. The outcomes could be either self-reported or clinician rated. Because CC is used in specialist care settings as well as in primary care, both settings were investigated. Exclusion criteria were:

- Articles that focused on only physical symptoms or only depressive-anxiety symptoms, instead of on both;
- Interventions that did not meet our criteria for CC;
- High risk of bias, indicated by a score of <4 out of 7 in the risk of bias assessment performed according to the Cochrane quality criteria [45].

Two researchers (J.E.S and H.C.) independently and in duplicate judged whether each study met our inclusion criteria based on the title and abstract of the article. After this, three researchers (J.E.S., H.C. and V.E.) independently and in triplicate ascertained whether each study met the inclusion criteria on the basis of the full text. Disagreement regarding inclusion status was discussed and we tried to come to a mutual agreement. If no consensus could be met, a fourth researcher (C.F.C.) was consulted. Consensus was reached in all cases. C.F.C. approached the EURONET-SOMA members for relevant publications. J.E.S. checked the trial registers.

2.3 Data collection process

Relevant outcome analyses were independently extracted from full text articles by three researchers (J.E.S., H.C. and V.E.). Extracted data included the following study characteristics and outcome analyses: setting, mean age, somatic diagnosis, psychiatric diagnosis, treatment model, treatment provided by CM, and main outcome (mean) compared to CAU. We pilot tested the standardized form to check if it was complete. In the case of multiple follow-up assessments, we chose the short-medium term main finding of the study (3-12 months). Authors were contacted to clarify data when there were uncertainties.

2.4 Data items

2.4.1 Type of study design

Randomized controlled trials.

2.4.2 Definition of the treatment interventions included in the review

The treatment intervention had to be carefully examined to assess whether the intervention was in accordance with our definition of CC. The characteristics of the treatment intervention had to be specified in terms of the roles of the GP/medical specialist, psychiatrist and CM.

2.4.3 Patient groups

Trials were included if the patients were aged 18 years or older, and were diagnosed with both a chronic medical condition and a depressive disorder and/or anxiety disorder. Trials were eligible for inclusion if they were conducted in primary care and/or specialist settings. Depressive disorders were classified as minor depressive disorder, major depressive disorder, dysthymic disorder or persistent depressive symptoms. Anxiety disorders were classified as any possible anxiety disorder according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision (DSM-IV-TR) [46]. Both depressive disorders and anxiety disorders had to be assessed by a validated questionnaire with a cut-off score, or by a diagnostic interview.

2.4.4 Outcome definitions

An important difference between our study and earlier reviews is that our study focuses explicitly on illness burden and physical outcomes of CC. Outcomes should define the impact of the intervention on the clinical condition of the patients in terms of physical symptomatology and depressive or anxiety symptom severity. Outcomes should report objective measures of such dimensions, as assessed by validated questionnaires or methods. These questionnaires could either be self-report or clinician rated. Table 1 describes all physical and depressive outcome measures included in the current study.

-Insert Table 1-

2.5 Risk of bias in individual studies

Risk of bias in each trial was assessed in triplicate, at study and outcome level, by three independent assessors (J.E.S., V.E. and H.C.). The assessment was performed according to the Cochrane quality criteria [45], of which randomisation was the most important [45].

2.6 Summary measures

We performed a random effects meta-analysis [47]. We calculated effect sizes for each study using the statistical program Comprehensive Meta-Analysis, version 2 [48]. If studies reported on multiple outcomes, we chose to take the following outcome measures into account for the analysis:

- (1) Physical symptom severity, such as cardiac symptoms, HbA1c, hypertension parameters and pain.
- (2) Somatic diagnostic and intervention procedures such as hospital admission, surgery, diagnostic procedures, etc.
- (3) Physical functioning.
- (4) Depressive symptom severity, i.e., treatment response or remission.

2.7 Synthesis of results and risk of bias across studies

2.7.1 Illness burden

We calculated and pooled effect sizes for both physical and depressive outcomes using a weighted mean that should indicate a combined assessment of illness burden. Illness burden was defined as an outcome that establishes the impact of the intervention at hand on the general clinical condition of the patients. The clinical meaning of this is envisioned to be somewhat similar to the Clinical Global Impression [49], an assessment that a clinician would make in order to indicate if the general clinical condition of a patient has improved in terms of symptoms and functioning. For this, we used the physical and depressive outcomes mentioned above, and translated these measures to a standardized pooled combined equally weighted effect size, using a statistical program that corrects for possible inflated weights, in the case of multiple outcomes, using equal weighting [48]. The program computes the weighted mean for each study, by taking the average of the outcomes, and the average variance. This computation assumes that the correlation between the outcomes of one study is 1.0. In fact, this correlation is almost certainly lower than 1.0, so this is a conservative approach, aimed at correcting for possible inflated weight caused by multiple outcomes. We gave the outcomes similar weights in the weighted means calculation. This

meant that illness burden was pooled consistent with previous research [12]. We calculated an Odds Ratio (OR) with a 95% confidence interval (95%CI).

2.7.2 Physical and depressive outcomes

In the next step, we pooled the effects for physical outcomes and for depressive outcomes separately. Then, we pooled across different chronic medical conditions, and multi-morbidity, as the effect of CC might be of a general nature and thus, comparison of the general effect of CC across several chronic medical conditions would be of high clinical relevance. Disease control parameters were included if they were reported in the study.

In the random-effects meta-analysis [47], between-study heterogeneity was assessed using the Q-statistic, which reflects the observed dispersion [50]. In order to quantify this dispersion, the I^2 statistic was used, which describes the percentage of total variation across the studies that is the result of heterogeneity rather than of chance. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity; I^2 values of 25%, 50% and 75% correspond to low, moderate and high heterogeneity, respectively [50]. Publication bias was examined by constructing a Begg [51] funnel plot and performing the fail-safe N [52]. High-resolution plots are presented. All statistical pooling was conducted using the program Comprehensive Meta-Analysis, version 2. Missing data were not imputed. The findings are presented according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [53]. The PRISMA checklist is included in Appendix B.

2.8 Additional pre-envisioned subgroup analyses

Additional pre-defined analyses were performed for three theoretically interesting potential moderating effects, based on the literature [32], and as described in our hypotheses. As such, we explored if CC model, setting and the role of the CM differentiates the effect of CC in chronic medical conditions with comorbid depressive or anxiety disorder. The effect sizes were calculated and pooled.

Results

3.1 Study selection

The initial search resulted in 1109 hits, from which 199 studies were selected based on the titles. Of those, 63 studies were selected based on the abstracts, and of those studies, 20 met the inclusion criteria based on the full text. The search of the clinical trial registers

yielded one trial that fitted our inclusion criteria, but this was a trial we had already included (see Flow diagram 1) [53].

-Insert Flow diagram-

The total N from these 20 studies was 4774. All studies focused on depression; only one study investigated the combination of depression and anxiety. Because we found only one study that reported physical outcomes of CC compared to CAU for the combination of a chronic medical condition with comorbid anxiety disorder, we will only report the effects of CC for the combination of a chronic medical condition and depressive disorder.

3.2 Study characteristics

In accordance with our inclusion criteria, all studies were RCTs. The sample sizes varied between 55 and 1001 patients. All 20 studies reported on a physical outcome separately. Follow-up periods were short to medium term, i.e., 3-12 months.

3.3 Risk of bias

The risk of bias was rated for all 20 studies; see Table 2. All 20 studies were rated as acceptable (low risk of bias) for inclusion in the meta-analysis. In this review, a cut-off score of at least four out of a possible seven scores was chosen to define a low risk of bias, according to the Cochrane quality criteria [45]. In all 20 studies, blinding of clinical personnel to treatment type provided was not possible due to the study design. However, this is a common feature of psychotherapeutic intervention studies [54].

-Insert Table 2-

3.4 Results of individual studies

The characteristics of the 20 included studies, as well as the effect sizes calculated in our meta-analyses, are shown in Table 3. A variety of chronic medical conditions were reported in the 20 studies: Acute coronary syndrome (ACS) (five studies; n = 866) [55-59], arthritis (two studies; n = 1251) [60, 61], cancer (three studies; n = 727) [62-64], chronic obstructive pulmonary disease (COPD) (one study; n = 138) [65], diabetes mellitus (DM) type II (three studies; n = 774) [66-68], epilepsy (one study; n = 80) [69], human immunodeficiency virus (HIV) (one study; n = 276) [70], hypertension (one study; n = 64) [71], and multi-morbidity (three studies; n = 598) [9, 72, 73]. Fifteen of the 20 studies diagnosed depression or dysthymia using self-report, 12 of the studies used the Patient Health Questionnaire (PHQ-9), two studies used the Beck Depression Inventory (BDI) and

one study used the Centre for Epidemiologic Studies Depression Scale (CES-D). Four studies used the Structured Clinical Interview for DSM-IV to diagnose major or minor depression, and one study use the Mini International Neuropsychiatric Interview (MINI). The mean age of the patients in the studies varied between 44 and 72 years.

-Insert Table 3-

3.5 Synthesis of results

3.5.1 Illness burden

In order to establish the illness burden based on the physical and depressive outcomes, a meta-analysis was performed across the 20 studies, N = 4774. The overall OR was 1.64 (95%CI 1.47;1.83) with an overall effect size of CC of $d = 0.27$, 95%CI 0.21;0.33 for CC compared to CAU. The results showed low statistical heterogeneity ($I^2 < 0.1$) and no significant dispersion (Q value of 14.95, df = 19, $p = 0.75$). The Forest plot is shown below in Figure 2.

-Insert Figure 2-

3.5.2 Combined physical outcomes

In order to establish the effect of CC on physical outcomes, a meta-analysis was performed for the combined physical outcomes: HbA1C, number of epileptic seizures, HIV-symptom severity, pain and physical functioning. This yielded 19 studies, total N = 4692. The overall OR for CC compared to CAU was 1.46 (95%CI 1.28;1.67) and the overall effect size of CC was $d = 0.21$, 95%CI 0.14;0.26. The results showed low statistical heterogeneity ($I^2 = 25.21$) and no significant dispersion (Q value of 24.07, df = 18, $p = 0.15$). The Forest plot is shown in Figure 3.

-Insert Figure 3-

The results varied depending on the chronic medical condition investigated. Results were significant for all chronic medical conditions, except for DM type II and epilepsy. The highest OR (4.18, 95%CI 2.18;8.04; $d = 0.79$, 95%CI 0.43;1.15) was found for hypertension. HIV, COPD, multi-morbidity, arthritis, cancer and ACS showed smaller ORs, with no significant dispersion (Q = 16,854, df = 8, $p = 0.032$). The results are shown in Figure 4.

-Insert Figure 4-

3.5.3 Somatic diagnostic or treatment procedures

Four studies reported on somatic diagnostic or treatment procedures; see the Forest plot in Figure 5. Low heterogeneity was found between the different treatment studies in terms of somatic diagnoses and intervention procedures ($I^2 < 0.01$; Q value = 0.25, df = 3, p = 0.97). Overall, more procedures were performed in CC compared to CAU, with an OR of 2.25 (95%CI 1.67;3.04), d = 0.45 (95%CI 0.28;0.61).

-Insert Figure 5-

3.5.4 Depression outcomes

All studies reported a depression outcome [9, 55-57, 60, 62-66, 69, 70, 72]. There was no statistical heterogeneity across the studies ($I^2 < 0.01$, Q value = 18.63, df = 19, p = 0.48). Results for CC were significantly better than for CAU, with an overall OR of 1.90 (95%CI 1.70;2.12). The overall effect size of CC for the depression outcome was small to moderate (d = 0.35, 95%CI 0.29;0.42). No statistically significant differences in depression outcomes were found across the different physical disorders. With the exception of COPD (based on only one study), all depressive outcomes were statistically significant. The results are shown in Figure 6.

-Insert Figure 6-

3.5.5 Additional analyses: Moderating effects

We selected three potential moderating effects, as indicated in Table 3. The first potential moderating effect was the CC model. CC treatment was either given by a CM in collaboration with a GP (four studies) or by a CM in collaboration with a physician and a psychiatrist (16 studies). The second hypothesised potential moderating effect was the setting in which the CC was provided; hospital care setting (11 studies) or primary care setting (nine studies). The third potential moderating effect was the role of the CM: monitoring only (10 studies) or providing psychotherapy as well as monitoring (10 studies). For all three moderating effects, no statistically significant differences in treatment effects of CC compared to CAU were found for illness burden, physical outcomes, somatic and diagnostic interventions and depression outcomes.

3.6 Risk of bias across studies: Publication bias

Based on the funnel plot (see Figure 6) is it unlikely that any studies that have been missed would change the results, given that we would have to include at least 431 extra studies to find diverging results.

4. Discussion

4.1 Summary of evidence

This is the first systematic review that focusses explicitly on the effects of CC on illness burden and physical outcomes for patients with comorbid depression in chronic medical conditions. We performed a systematic review and meta-analysis of 20 RCTs, and reported overall illness burden, physical outcomes and depressive outcomes for CC compared to CAU, among patients with comorbid depressive disorder and chronic medical conditions. Our results indicate that CC is more effective than CAU in terms of illness burden, physical outcomes and depression outcomes. Although the overall effect sizes for illness burden and combined physical outcomes were small ($d = 0.27$ and 0.21 , respectively), the effect sizes for the specific chronic medical conditions varied considerably. This confirms our first hypothesis and makes a strong case for providing CC in chronic medical conditions with comorbid depression.

Regarding our second hypothesis, the largest statistically significant effect size in terms of chronic medical conditions was found for hypertension, followed by HIV, COPD, multi-morbidity, arthritis, cancer and ACS. The strong effects of CC on hypertension can most likely be explained by the greater adherence of these patients to both antidepressant and antihypertensive medications [71]. However, given that there is only one study in our meta-analysis that investigated hypertension, more studies are needed to draw firm conclusions. We hypothesized that CC would be most effective in chronic medical conditions with high potential for self-management and case-management; our findings partly confirm this. However, our finding that physical outcomes are not significantly better for CC compared to CAU for patients with type II DM shows that self-management may not be enough to improve outcomes for patients with type II DM, especially in terms of HbA1c.

In patients with DM, the presence of multiple complications is significantly associated with comorbid depression [74], meaning that more complex patients are at higher risk for depression. Because of this complexity, both enhanced coordination of care and self-management are important. Both are an important part of CC, and a possible explanation for the success of CC in complex patient groups. For patients with DM, CC has previously been recommended as a method to provide all necessary aspects of depression care, such as psychotherapy and anti-depressant medication [12]. The results of our meta-analysis support this recommendation. We also found that more somatic diagnostic and treatment procedures were performed in cases of CC compared to CAU (OR 2.25). Although more procedures make CC more costly, several studies have shown that CC improves quality of life for

acceptable additional costs [38, 75]. This has been shown to be the case for depression comorbid with various chronic medical conditions [76-81].

In patients with ACS, comorbid depression is associated with poorer quality of life [82]. Therefore, given that CC has a better effect than CAU on depression, the fact that depression treatment in ACS has a known favourable prognostic effect on cardiac outcomes [83], and the fact that CC has a positive effect on physical outcomes, CC should be considered a good treatment option in ACS. CC treatment studies with long-term follow up, and focusing on mortality [84], would be needed.

The effect of CC was better than CAU for the depression outcomes, with an overall small to moderate effect size ($d = 0.35$); there was no statistically significant difference between the chronic medical conditions. It is possible that, even when chronic medical conditions show limited potential for improvement under a generic behavioural intervention -such as self-management- as in CC, depressive symptoms can still be further improved by CC compared to CAU. For example, previous research has shown that patients with epilepsy in usual care may not receive the depression care they require [85]. Given that they suffer from fear of seizures [86], are often stigmatized [87], and often have comorbid depression [88], adequate depression care is important. In view of the limited availability of RCTs exploring treatment of comorbid depression in epilepsy [89], the findings of this study suggest that CC may be a good candidate for further research and implementation in epilepsy patient care.

In our analyses we explored three potential moderating effects. Firstly, the CC model: treatment given by a CM in collaboration with a GP versus a CM in collaboration with a physician and a psychiatrist. Secondly, the setting in which the CC was provided: hospital care setting or primary care setting. And thirdly, the role of the CM: providing only monitoring or providing monitoring and psychotherapy. None of these showed statistically significant differences for physical and depressive treatment effects. It appears that, for improvement in both physical and depressive symptoms, it is more important that the CC model is applied in the first place, than who applies it or whether it is provided in a primary care or general hospital setting. Therefore, addressing both physical and mental health in monitoring, counselling and consultation can be seen as important factors for the success of CC. This is supported by the finding that a reduction in physical symptoms is associated with a reduction in depression [90].

The National Institute for Health and Care Excellence (NICE) states that CC is more (cost) effective than CAU for depression outcomes [34], among patients with depression and a chronic physical health problem. One of the recommendations of the Guideline Development Group of NICE was to research the effectiveness of CC on physical health outcomes for patients with depression and a chronic physical health problem. This review and meta-analysis gives an overview of the RCTs that have done so, using physical

outcomes as the primary outcome measure, and thereby adds to the body of knowledge about the effectiveness of CC for this complex group of patients with both physical and mental problems. Over half of the RCTs we identified in this review were published in 2010 or later, showing the recent interest in this topic.

4.2 Limitations

For our search we used the Pubmed database, the Cochrane database, and the reference lists of selected systematic reviews; only studies in English and Dutch were included. We also did a search in an European and an American trial register, and contacted principal investigators of key studies, as well as the members of EURONET-SOMA, for suggestions about other relevant trials. Even though we performed this rather extensive search, it is possible that we missed studies. However, the high Failsafe numbers (431) that we found, provide some reassurance.

The small number of good quality studies for each chronic medical condition covered in this review is a limitation. To provide better estimates for specific chronic medical conditions, more research is needed. Those studies could provide more information about the content of the CAU that is implemented, which is lacking in the current studies.

There is also a lack of research on physical outcomes in chronic medical conditions; most studies instead focus on only anxiety outcomes [91, 92]. Therefore, we cannot draw conclusions about this. This is an interesting subject for further research, because the combination of anxiety disorders and chronic medical conditions is highly prevalent [13, 93]; the presence of an anxiety disorder leads to a higher likelihood of disability in chronic medical conditions [15], and a more negative perception of the influence of the chronic medical condition [94]. It has also been suggested that the presence of physical symptoms negatively influences the course of the anxiety disorder [22].

4.3 Strengths

To our knowledge, this is the first study that reports on the effectiveness of CC for comorbid depression and chronic medical conditions; CC can be viewed as a generic intervention that can be applied in a variety of chronic medical conditions with comorbid depression, or in the case of multi-morbidity. The current study's focus on physical outcomes is an important strength of this work. In all studies, CAU was provided to the control group. In addition, statistical heterogeneity (I^2) of the results of the studies included in our meta-analysis was low to moderate. Further, based on the funnel plot, it is unlikely that studies that would change the results have been missed. Therefore, the results can probably be considered robust. However, the effect sizes were still comparatively small so we may need

more personalised treatment models for patients with somatic-psychiatric comorbidity; this is a potential focus for further research in the future [95-97].

4.4 Conclusions

In conclusion, CC is more effective than CAU in terms of illness burden and physical outcomes for patients with comorbid depressive disorder and chronic medical conditions, particularly in patients with hypertension, HIV, COPD, multi-morbidity, arthritis, cancer and ACS. CC is associated with more diagnostic and treatment intervention procedures, and hence, may play a role in improving quality of life. Compared to CAU, CC is also effective for depression outcomes in patients with chronic medical conditions. More research investigating multiple chronic medical conditions, with longer follow-up, and taking mortality into account, is needed in the future.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

- [1] Ezedum S, Kerr DN. Collaborative care of hypertensives, using a shared record. *Br Med J* 1977;2(6099):1402-3.
- [2] Vlasveld MC, Anema JR, Beekman AT, van Mechelen W, Hoedeman R, van Marwijk HW, et al. Multidisciplinary collaborative care for depressive disorder in the occupational health setting: Design of a randomised controlled trial and cost-effectiveness study. *BMC Health Serv Res* 2008;8:99.
- [3] Van der Feltz-Cornelis CM. The depression initiative. Description of a collaborative care model for depression in the primary care setting in the Netherlands. *Clinical Neuropsychiatry* 2011;8(4).
- [4] Walker J, Sharpe M. Depression care for people with cancer: A collaborative care intervention. *Gen Hosp Psychiatry* 2009;31(5):436-41.
- [5] Chwastiak LA, Jackson SL, Russo J, DeKeyser P, Kiefer M, Belyeu B, et al. A collaborative care team to integrate behavioral health care and treatment of poorly-controlled type 2 diabetes in an urban safety net primary care clinic. *Gen Hosp Psychiatry* 2017;44:10-5.
- [6] Beach SR, Walker J, Celano CM, Mastromauro CA, Sharpe M, Huffman JC. Implementing collaborative care programs for psychiatric disorders in medical settings: A practical guide. *Gen Hosp Psychiatry* 2015;37(6):522-7.
- [7] Rossom RC, Solberg LI, Magnan S, Crain AL, Beck A, Coleman KJ, et al. Impact of a national collaborative care initiative for patients with depression and diabetes or cardiovascular disease. *Gen Hosp Psychiatry* 2017;44:77-85.
- [8] Gilbody S, Bower P, Fletcher J, Richards D, Sutton AJ. Collaborative care for depression: A cumulative meta-analysis and review of longer-term outcomes. *Arch Intern Med* 2006;166(21):2314-21.
- [9] van Steenberghe-Weijnenburg KM, van der Feltz-Cornelis CM, van Benthem TB, Horn EK, Ploeger R, Brals JW, et al. [Collaborative care for co-morbid major depressive disorder in chronically ill outpatients in a general hospital]. *Tijdschr Psychiatr* 2015;57(4):248-57.
- [10] Carter BL. Primary care physician-pharmacist collaborative care model: Strategies for implementation. *Pharmacotherapy* 2016;36(4):363-73.
- [11] Mynors-Wallis LM, Gath DH, Day A, Baker F. Randomised controlled trial of problem solving treatment, antidepressant medication, and combined treatment for major depression in primary care. *BMJ* 2000;320(7226):26-30.
- [12] van der Feltz-Cornelis CM, Nuyen J, Stoop C, Chan J, Jacobson AM, Katon W, et al. Effect of interventions for major depressive disorder and significant depressive

- symptoms in patients with diabetes mellitus: A systematic review and meta-analysis. *Gen Hosp Psychiatry* 2010;32(4):380-95.
- [13] van Eck van der Sluijs J, Ten Have M, Rijnders C, van Marwijk H, de Graaf R, van der Feltz-Cornelis C. Medically unexplained and explained physical symptoms in the general population: Association with prevalent and incident mental disorders. *PloS one* 2015;10(4):e0123274.
 - [14] Penninx BW, van Dyck R. [Depression and somatic comorbidity]. *Ned Tijdschr Geneesk* 2010;154:A1784.
 - [15] Sareen J, Cox BJ, Clara I, Asmundson GJ. The relationship between anxiety disorders and physical disorders in the U.S. National Comorbidity Survey. *Depress Anxiety* 2005;21(4):193-202.
 - [16] Chong SA, Abidin E, Nan L, Vaingankar JA, Subramaniam M. Prevalence and impact of mental and physical comorbidity in the adult Singapore population. *Ann Acad Med Singapore* 2012;41(3):105-14.
 - [17] Hsiang H, Karen MT, Joseph MC, Nahida A, Amritha B, Rachel K. Collaborative care for women with depression: A systematic review. *Psychosomatics* 2017;58(1):11-8.
 - [18] Hoedeman R, Krol B, Blankenstein N, Koopmans PC, Groothoff JW. Severe MUPS in a sick-listed population: A cross-sectional study on prevalence, recognition, psychiatric co-morbidity and impairment. *BMC Public Health* 2009;9:440.
 - [19] Dewa CS, Lin E. Chronic physical illness, psychiatric disorder and disability in the workplace. *Soc Sci Med* 2000;51(1):41-50.
 - [20] Goetzel RZ, Long SR, Ozminkowski RJ, Hawkins K, Wang S, Lynch W. Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting U.S. employers. *J Occup Environ Med* 2004;46(4):398-412.
 - [21] Kang HJ, Kim SY, Bae KY, Kim SW, Shin IS, Yoon JS, et al. Comorbidity of depression with physical disorders: Research and clinical implications. *Chonnam Med J* 2015;51(1):8-18.
 - [22] Huijbregts KM, van Marwijk HW, de Jong FJ, Schreuders B, Beekman AT, van der Feltz-Cornelis CM. Adverse effects of multiple physical symptoms on the course of depressive and anxiety symptoms in primary care. *Psychother Psychosom* 2010;79(6):389-91.
 - [23] Huijbregts KM, van der Feltz-Cornelis CM, van Marwijk HW, de Jong FJ, van der Windt DA, Beekman AT. Negative association of concomitant physical symptoms with the course of major depressive disorder: A systematic review. *Journal of psychosomatic research* 2010;68(6):511-9.

- [24] Huijbregts KM, de Jong FJ, van Marwijk HW, Beekman AT, Ader HJ, van der Feltz-Cornelis CM. A high physical symptom count reduces the effectiveness of treatment for depression, independently of chronic medical conditions. *Journal of psychosomatic research* 2013;74(3):179-85.
- [25] de Jong FJ, van Steenbergen-Weijenburg KM, Huijbregts KM, Vlasveld MC, Van Marwijk HW, Beekman AT, et al. The Depression Initiative. Description of a collaborative care model for depression and of the factors influencing its implementation in the primary care setting in the Netherlands. *Int J Integr Care* 2009;9:e81.
- [26] Jacob V, Chattopadhyay SK, Sipe TA, Thota AB, Byard GJ, Chapman DP, et al. Economics of collaborative care for management of depressive disorders: A community guide systematic review. *Am J Prev Med* 2012;42(5):539-49.
- [27] Goorden M, Huijbregts KM, van Marwijk HW, Beekman AT, van der Feltz-Cornelis CM, Hakkaart-van Roijen L. Cost-utility of collaborative care for major depressive disorder in primary care in the Netherlands. *Journal of psychosomatic research* 2015;79(4):316-23.
- [28] Huijbregts KM, de Jong FJ, van Marwijk HW, Beekman AT, Ader HJ, Hakkaart-van Roijen L, et al. 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. *J Affect Disord* 2013;146(3):328-37.
- [29] Wetherell JL, Petkus AJ, Thorp SR, Stein MB, Chavira DA, Campbell-Sills L, et al. Age differences in treatment response to a collaborative care intervention for anxiety disorders. *Br J Psychiatry* 2013;203(1):65-72.
- [30] Archer J, Bower P, Gilbody S, Lovell K, Richards D, Gask L, et al. Collaborative care for depression and anxiety problems. *Cochrane Database Syst Rev* 2012;10:CD006525.
- [31] Woltmann E, Grogan-Kaylor A, Perron B, Georges H, Kilbourne AM, Bauer MS. Comparative effectiveness of collaborative chronic care models for mental health conditions across primary, specialty, and behavioral health care settings: Systematic review and meta-analysis. *Am J Psychiatry* 2012;169(8):790-804.
- [32] Baumeister H, Hutter N. Collaborative care for depression in medically ill patients. *Curr Opin Psychiatry* 2012;25(5):405-14.
- [33] Muntingh AD, van der Feltz-Cornelis CM, van Marwijk HW, Spinhoven P, van Balkom AJ. Collaborative care for anxiety disorders in primary care: A systematic review and meta-analysis. *BMC Family Practice* 2016;17(1):62.
- [34] NICE. Depression in Adults with a chronic physical health problem: Treatment and Management. Leicester: British Psychological Society; 2010.

- [35] Li M, Kennedy EB, Byrne N, Gerin-Lajoie C, Katz MR, Keshavarz H, et al. Systematic review and meta-analysis of collaborative care interventions for depression in patients with cancer. *Psychooncology* 2016.
- [36] Tully PJ, Baumeister H. Collaborative care for comorbid depression and coronary heart disease: A systematic review and meta-analysis of randomised controlled trials. *BMJ Open* 2015;5(12):e009128.
- [37] Katon W, Guico-Pabia CJ. Improving quality of depression care using organized systems of care: A review of the literature. *The primary care companion for CNS disorders* 2011;13(1).
- [38] van Steenberg-Weijnenburg KM, van der Feltz-Cornelis CM, Horn EK, van Marwijk HW, Beekman AT, Rutten FF, et al. Cost-effectiveness of collaborative care for the treatment of major depressive disorder in primary care. A systematic review. *BMC Health Serv Res* 2010;10:19.
- [39] Wanat M, Walker J, Hodges L, Richardson A, Sharpe M. Selecting, training and supervising nurses to treat depression in the medically ill: Experience and recommendations from the SMaRT oncology collaborative care trials. *Gen Hosp Psychiatry* 2015;37(6):518-21.
- [40] Huffman JC, Niazi SK, Rundell JR, Sharpe M, Katon WJ. Essential articles on collaborative care models for the treatment of psychiatric disorders in medical settings: A publication by the academy of psychosomatic medicine research and evidence-based practice committee. *Psychosomatics* 2014;55(2):109-22.
- [41] <http://www.crd.york.ac.uk/PROSPERO/DisplayPDF.php?ID=CRD42016035553>.
- [42] <https://www.clinicaltrialsregister.eu>.
- [43] <https://clinicaltrials.gov>.
- [44] Weigel A, Husing P, Kohlmann S, Lehmann M, Shedden-Mora M, Toussaint A, et al. A European research network to improve diagnosis, treatment and care for patients with persistent somatic symptoms: Work report of the EURONET-SOMA conference series. *Journal of psychosomatic research* 2017;97:136-8.
- [45] The Cochrane Collaboration. *Cochrane handbook for systematic reviews of interventions*. 2011.
- [46] American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th ed. text revision. DC: Author; 2000.
- [47] DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled clinical trials* 1986;7(3):177-88.
- [48] Borenstein MHL, Higgins JPT, Rothstein HR. *Comprehensive meta-analysis*. Englewood (NJ): Biostat; 2005.

- [49] Guy W. ECDEU assessment manual for psychopharmacology—Revised. DHEW Publ No ADM 76-338. Rockville (MD); 1976.
- [50] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-60.
- [51] Cooper H, Hedges LV (Eds). *The handbook of research synthesis*. New York: Russel Sage Foundation; 1994.
- [52] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to meta-analysis (statistics in practice)*. Chichester (UK): Wiley & Sons; 2009.
- [53] Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 2009;6(7):e1000097.
- [54] Berger D. Double-blinding and bias in medication and cognitive-behavioral therapy trials for major depressive disorder. *F1000Research* 2015;4:638.
- [55] Davidson KW, Rieckmann N, Clemow L, Schwartz JE, Shimbo D, Medina V, et al. Enhanced depression care for patients with acute coronary syndrome and persistent depressive symptoms: Coronary psychosocial evaluation studies randomized controlled trial. *Arch Intern Med* 2010;170(7):600-8.
- [56] Huffman JC, Mastromauro CA, Sowden G, Fricchione GL, Healy BC, Januzzi JL. Impact of a depression care management program for hospitalized cardiac patients. *Circ Cardiovasc Qual Outcomes* 2011;4(2):198-205.
- [57] Huffman JC, Mastromauro CA, Beach SR, Celano CM, DuBois CM, Healy BC, et al. Collaborative care for depression and anxiety disorders in patients with recent cardiac events: The Management of Sadness and Anxiety in Cardiology (MOSAIC) randomized clinical trial. *JAMA Intern Med* 2014;174(6):927-35.
- [58] Davidson KW, Bigger JT, Burg MM, Carney RM, Chaplin WF, Czajkowski S, et al. Centralized, stepped, patient preference-based treatment for patients with post-acute coronary syndrome depression: CODIACS vanguard randomized controlled trial. *JAMA Intern Med* 2013;173(11):997-1004.
- [59] Carney RM, Freedland KE, Steinmeyer BC, Rubin EH, Ewald G. Collaborative care for depression symptoms in an outpatient cardiology setting: A randomized clinical trial. *Int J Cardiol* 2016;219:164-71.
- [60] Lin EH, Katon W, Von Korff M, Tang L, Williams JW, Jr., Kroenke K, et al. Effect of improving depression care on pain and functional outcomes among older adults with arthritis: A randomized controlled trial. *JAMA* 2003;290(18):2428-9.
- [61] Kroenke K, Bair MJ, Damush TM, Wu J, Hoke S, Sutherland J, et al. Optimized antidepressant therapy and pain self-management in primary care patients with

- depression and musculoskeletal pain: A randomized controlled trial. *JAMA* 2009;301(20):2099-110.
- [62] Dwight-Johnson M, Ell K, Lee PJ. Can collaborative care address the needs of low-income Latinas with comorbid depression and cancer? Results from a randomized pilot study. *Psychosomatics* 2005;46(3):224-32.
 - [63] Ell K, Xie B, Quon B, Quinn DI, Dwight-Johnson M, Lee PJ. Randomized controlled trial of collaborative care management of depression among low-income patients with cancer. *J Clin Oncol* 2008;26(27):4488-96.
 - [64] Strong V, Waters R, Hibberd C, Murray G, Wall L, Walker J, et al. Management of depression for people with cancer (SMaRT oncology 1): A randomised trial. *Lancet* 2008;372(9632):40-8.
 - [65] Alexopoulos GS, Kiess DN, Sirey JA, Kanellopoulos D, Seirup JK, Novitch RS, et al. Untangling therapeutic ingredients of a personalized intervention for patients with depression and severe COPD. *Am J Geriatr Psychiatry* 2014;22(11):1316-24.
 - [66] Katon WJ, Von Korff M, Lin EH, Simon G, Ludman E, Russo J, et al. The Pathways Study: A randomized trial of collaborative care in patients with diabetes and depression. *Arch Gen Psychiatry* 2004;61(10):1042-9.
 - [67] Bogner HR, de Vries HF. Integrating type 2 diabetes mellitus and depression treatment among African Americans: A randomized controlled pilot trial. *Diabetes Educ* 2010;36(2):284-92.
 - [68] Ell K, Katon W, Xie B, Lee PJ, Kapetanovic S, Guterman J, et al. Collaborative care management of major depression among low-income, predominantly Hispanic subjects with diabetes: A randomized controlled trial. *Diabetes Care* 2010;33(4):706-13.
 - [69] Ciechanowski P, Chaytor N, Miller J, Fraser R, Russo J, Unutzer J, et al. PEARLS depression treatment for individuals with epilepsy: A randomized controlled trial. *Epilepsy Behav* 2010;19(3):225-31.
 - [70] Pyne JM, Fortney JC, Curran GM, Tripathi S, Atkinson JH, Kilbourne AM, et al. Effectiveness of collaborative care for depression in human immunodeficiency virus clinics. *Arch Intern Med* 2011;171(1):23-31.
 - [71] Bogner HR, de Vries HF. Integration of depression and hypertension treatment: A pilot, randomized controlled trial. *Ann Fam Med* 2008;6(4):295-301.
 - [72] Rollman BL, Belnap BH, LeMenager MS, Mazumdar S, Houck PR, Counihan PJ, et al. Telephone-delivered collaborative care for treating post-CABG depression: A randomized controlled trial. *JAMA* 2009;302(19):2095-103.

- [73] Katon WJ, Lin EH, Von Korff M, Ciechanowski P, Ludman EJ, Young B, et al. Collaborative care for patients with depression and chronic illnesses. *N Engl J Med* 2010;363(27):2611-20.
- [74] van Steenbergen-Weijenburg KM, van Puffelen AL, Horn EK, Nuyen J, van Dam PS, van Benthem TB, et al. More co-morbid depression in patients with type 2 diabetes with multiple complications. An observational study at a specialized outpatient clinic. *Diabet Med* 2011;28(1):86-9.
- [75] Katon W, Russo J, Lin EH, Schmittdiel J, Ciechanowski P, Ludman E, et al. Cost-effectiveness of a multicondition collaborative care intervention: A randomized controlled trial. *Arch Gen Psychiatry* 2012;69(5):506-14.
- [76] Camacho EM, Ntais D, Coventry P, Bower P, Lovell K, Chew-Graham C, et al. Long-term cost-effectiveness of collaborative care (vs usual care) for people with depression and comorbid diabetes or cardiovascular disease: A Markov model informed by the COINCIDE randomised controlled trial. *BMJ Open* 2016;6(10):e012514.
- [77] Celano CM, Healy B, Suarez L, Levy DE, Mastromauro C, Januzzi JL, et al. Cost-effectiveness of a collaborative care depression and anxiety treatment program in patients with acute cardiac illness. *Value Health* 2016;19(2):185-91.
- [78] Duarte A, Walker J, Walker S, Richardson G, Holm Hansen C, Martin P, et al. Cost-effectiveness of integrated collaborative care for comorbid major depression in patients with cancer. *Journal of Psychosomatic Research* 2015;79(6):465-70.
- [79] Johnson JA, Lier DA, Soprovich A, Al Sayah F, Qiu W, Majumdar SR. Cost-effectiveness evaluation of collaborative care for diabetes and depression in primary care. *Am J Prev Med* 2016;51(1):e13-20.
- [80] Painter JT, Fortney JC, Gifford AL, Rimland D, Monson T, Rodriguez-Barradas MC, et al. Cost-effectiveness of collaborative care for depression in HIV clinics. *J Acquir Immune Defic Syndr* 2015;70(4):377-85.
- [81] Goorden M, van der Feltz-Cornelis CM, van Steenbergen-Weijenburg KM, Horn EK, Beekman AT, Hakkaart-van Roijen L. Cost-utility of collaborative care for the treatment of comorbid major depressive disorder in outpatients with chronic physical conditions. A randomized controlled trial in the general hospital setting (CC-DIM). *Neuropsychiatric disease and treatment* 2017;13:1881-93.
- [82] Kim JM, Stewart R, Bae KY, Kang HJ, Kim SW, Shin IS, et al. Effects of depression co-morbidity and treatment on quality of life in patients with acute coronary syndrome: The Korean depression in ACS (K-DEPACS) and the escitalopram for depression in ACS (EsDEPACS) study. *Psychol Med* 2015;45(8):1641-52.

- [83] de Jonge P, Honig A, van Melle JP, Schene AH, Kuyper AM, Tulner D, et al. Nonresponse to treatment for depression following myocardial infarction: Association with subsequent cardiac events. *Am J Psychiatry* 2007;164(9):1371-8.
- [84] Miloyan B, Fried E. A reassessment of the relationship between depression and all-cause mortality in 3,604,005 participants from 293 studies. *World Psychiatry* 2017;16(2):219-20.
- [85] Hermann BP, Seidenberg M, Bell B. Psychiatric comorbidity in chronic epilepsy: Identification, consequences, and treatment of major depression. *Epilepsia* 2000;41(Suppl 2):S31-41.
- [86] Beyenburg S, Mitchell AJ, Schmidt D, Elger CE, Reuber M. Anxiety in patients with epilepsy: Systematic review and suggestions for clinical management. *Epilepsy Behav* 2005;7(2):161-71.
- [87] England MJ, Austin JK, Beck V, Escoffery C, Hesdorffer DC. Erasing epilepsy stigma: Eight key messages. *Health Promot Pract* 2014;15(3):313-8.
- [88] Gaus V, Kiep H, Holtkamp M, Burkert S, Kendel F. Gender differences in depression, but not in anxiety in people with epilepsy. *Seizure* 2015;32:37-42.
- [89] van der Feltz-Cornelis C. Treatment of interictal psychiatric disorder in epilepsy. I. Affective and anxiety disorders. *Acta Neuropsychiatrica* 2002;14(1):39-43.
- [90] Chowdhury S, Burton C. Associations of treatment effects between follow-up times and between outcome domains in interventions for somatoform disorders: Review of three Cochrane reviews. *Journal of Psychosomatic Research* 2017;98:10-8.
- [91] Muntingh A, van der Feltz-Cornelis C, van Marwijk H, Spinhoven P, Assendelft W, de Waal M, et al. Effectiveness of collaborative stepped care for anxiety disorders in primary care: A pragmatic cluster randomised controlled trial. *Psychother Psychosom* 2014;83(1):37-44.
- [92] Kronish IM, Chaplin WF, Rieckmann N, Burg MM, Davidson KW. The effect of enhanced depression care on anxiety symptoms in acute coronary syndrome patients: Findings from the COPES trial. *Psychother Psychosom* 2012;81(4):245-7.
- [93] Walters P, Barley EA, Mann A, Phillips R, Tylee A. Depression in primary care patients with coronary heart disease: Baseline findings from the UPBEAT UK study. *PloS one* 2014;9(6):e98342.
- [94] Smith A, Fortune Z, Phillips R, Walters P, Lee GA, Mann A, et al. UPBEAT study patients' perceptions of the effect of coronary heart disease on their lives: A cross-sectional sub-study. *Int J Nurs Stud* 2014;51(11):1500-6.
- [95] Elfeddali I, van der Feltz-Cornelis CM, van Os J, Knappe S, Vieta E, Wittchen HU, et al. Horizon 2020 priorities in clinical mental health research: Results of a consensus-

- based ROAMER expert survey. *International Journal of Environmental Research and Public Health* 2014;11(10):10915-39.
- [96] van der Feltz-Cornelis CM, van Os J, Knappe S, Schumann G, Vieta E, Wittchen HU, et al. Towards Horizon 2020: Challenges and advances for clinical mental health research - Outcome of an expert survey. *Neuropsychiatric Disease and Treatment* 2014;10:1057-68.
- [97] Haro JM, Ayuso-Mateos JL, Bitter I, Demotes-Mainard J, Leboyer M, Lewis SW, et al. ROAMER: Roadmap for mental health research in Europe. *Int J Methods Psychiatr Res* 2014;23(Suppl 1):1-14.

Appendix A

Search strategy Pubmed

("Somatic Symptoms"[ALL] OR "somatic"[ALL] OR "physical symptoms"[ALL]) AND (("Depression"[Mesh] OR "Affective Symptoms"[Mesh] OR "Depressive Disorder"[Mesh] OR "Depressive Disorder, Major"[Mesh] OR "Dysthymic Disorder"[Mesh] OR "Mood Disorders"[Mesh] OR "depress"[tiab] OR "affective symptom"[tiab] OR "dysthymi"[tiab] OR "affective disorder"[tiab] OR "mood disorder"[tiab]) OR ("Anxiety Disorders"[Mesh] OR "Stress Disorders, Post-Traumatic"[Mesh] OR "Panic Disorder"[Mesh] OR "Obsessive-Compulsive Disorder"[Mesh] OR "Generalized anxiety disorder"[tiab] OR "Anxiety"[tiab] OR "Anxious"[tiab] OR "phobia"[tiab] OR "post-trauma"[tiab] OR "post trauma"[tiab] OR "postrauma"[tiab] OR "panic"[tiab] OR "obsessive"[tiab] OR "compulsive"[tiab])) AND (("Collaborative"[ALL] OR "coordinate"[ALL] OR "co-ordinate"[ALL] OR "shared"[ALL] OR "integrat"[ALL] OR "stepped"[ALL] OR "systematic"[ALL]) AND ("care"[ALL] OR "healthcare"[ALL] OR "health care"[ALL] OR "working"[ALL] OR "intervention"[ALL] OR "service"[ALL] OR "model"[ALL] OR "effort"[ALL] OR "manage"[ALL]) OR ((augment[ALL] or enhance[ALL]) AND (care[ALL] or healthcare[ALL] or "health care"[ALL] or communicat[ALL])) OR ("care manage*" OR "case manage*" OR "chronic care"[ALL] OR "complex intervention"[ALL] OR "cooperative behave*" OR "co-operative behave*" OR "joint working"[ALL] OR pathway[ALL] OR interprofessional[ALL] OR inter-professional[ALL] OR interdisciplinary[ALL] OR interdisciplinary[ALL] OR multidiscipline[ALL] OR multi-disciplin* OR multiprofession* OR multi-profession* OR transdisciplin* OR trans-disciplin* OR multifacet[ALL] OR multifacet[ALL] OR "complex intervention"[ALL] OR "multiple intervention"[ALL] OR multi-intervention[ALL] OR "organisational intervention"[ALL] OR "organizational intervention"[ALL] OR "interpersonal relation*" OR " interpersonal relation*" OR "interinstitutional relation*" OR "inter-institutional relation*" OR "consultation liaison"[ALL] OR algorithm[ALL] OR "treatment guideline"[ALL] OR "treatment protocol"[ALL] OR "treatment delivery"[ALL] OR "treatment model"[ALL] OR adherence[ALL] OR compliance[ALL] OR concordance[ALL] OR "patient care team"[ALL] OR "patient care management"[ALL] OR "patient care planning"[ALL] OR "casemanagement"[ALL] OR "managed care program"[ALL] OR "delivery of healthcare"[ALL] OR "continuity of patient care"[ALL] OR "professional-patient relations"[ALL] OR "interprofessional relations"[ALL])) AND ((Clinical trial[pt] OR Randomized Controlled Trial[pt] OR randomize*[tiab] OR randomis* [tiab] OR placebo[tiab] OR clinical trials[mh] OR randomly[tiab] OR trial[ti]) NOT (animals[mh] NOT (animals[mh] AND humans[mh])) OR ("meta-analysis"[pt]) OR (metaanal*[tw]) OR (metaanal*[tw]) OR (quantitative*[tw] AND review*[tw]) OR (quantitative*[tw] AND overview*[tw]) OR (systematic*[tw] AND review*[tw]) OR (systematic*[tw] AND overview*[tw]) OR (methodologic*[tw] AND review*[tw]) OR (methodologic*[tw] AND overview*[tw]) OR ("review"[pt] AND "medline"[tw]))

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2,3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4,5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5,6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	7
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Appendix A
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7,8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8,9
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	9,10
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	10
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	10,11
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	11,12

PRISMA 2009 Checklist

Page 1 of 2

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	11,12
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	12
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	13 + Flow diagram 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	13
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	13,14
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	14 + Table 2 + Forest plot 1-5
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	14-16
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	17
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	16,17
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	18-20
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	20-21
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	21-22
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	22

PRISMA 2009 Checklist

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

Appendix C Risk of bias assessment

Davidson et al, 2010 [55]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	'At each site, eligible patients were randomized on a 1:1 basis within randomly ordered blocks of 4 or 6 patients according to a table of assignments prepared in advance by the trial statistician.'
Allocation concealment (selection bias)	+	'Using a Web-based program, project coordinators specified the strata, initials, and study identification number of the person to be randomized, and the program issued the group assignment.'
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	'Interviewers and those collecting medical outcome data were blinded to intervention assignment.'
Incomplete outcome data addressed (attrition bias)	+	Drop-out is presented in a flowchart: 10/80 in the intervention group and 6/77 in the usual care group at 9 months. An intention-to-treat analyses was performed.
Selective reporting (reporting bias)	+	All expected outcomes are reported.
Other	+	13 of the 80 patients randomized to treatment never attended a first depression care visit, which will also happen when a treatment is implemented.

Davidson et al, 2013 [58]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	'Participants were randomized in a 1:1 ratio according to a computer-generated, permuted-block randomization scheme, stratified by site and current antidepressant use'
Allocation concealment (selection bias)	+	'Participants were randomized ... that local research assistants accessed by telephone'
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	'Research assistants who were masked to randomization administered the BDI and ascertained adverse events, mental health and hospitalization information, and medication use at baseline and 6 months during in-person interviews and at 2 and 4 months by telephone.'
Incomplete outcome data addressed (attrition bias)	+	There were no statistical differences in baseline characteristics between the two groups. Attrition in the intervention group was 6/73 = 8%; and in the CAU group 6/77 = 8%.
Selective reporting (reporting bias)	+	All collected outcomes are reported.
Other	+	N/A

Huffman et al, 2011 [56]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	After enrollment, subjects were randomly assigned via random number generator to collaborative care or usual

		care.
Allocation concealment (selection bias)	Unclear	Insufficient information to permit judgement. The method of concealment is not described.
Blinding personnel (performance bias)	-	Not possible for practice staff and patients. 'unit staff members were aware that a depression study was ongoing, and usual care subjects' physicians were informed of the subjects' depression".
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	Blinded study staff reviewed subjects' medical records for demographic information, cardiac risk factors and admission medication. Study outcome measures were obtained by study staff blinded to subjects' treatment condition.
Incomplete outcome data addressed (attrition bias)	+	Of the 175 patients at baseline, for 164 patients at least 1 follow-up assessment was available. Reasons: deceased patients (Int: n=2/UC n=2), declined continued participation (Int: n=0/UC: n=2), lost to follow-up (Int: n=4/UC: n=1). Patients in the intervention and usual care groups were similar on all baseline demographic and clinical characteristics, no significant between group differences in baseline. A total of 127 subjects (77.4% of 164) completed 6-week follow-up. 138 (86.3%) completed the 12-week follow-up, 137 (86.2%) completed the 6-months follow-up.
Selective reporting (reporting bias)	+	All outcomes specified in the article are reported.
Other	+	-Study was performed in a single academic medical center in a largely white population. -This is a much less intense follow-up intervention than most successful collaborative care programs (3 phone contacts from a care manager over the 12 week postdischarge period and no contacts thereafter.

Huffman et al, 2014 [57]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	"Participants were randomized to CC or UC using computerized random number generator".
Allocation concealment (selection bias)	+	"numbers placed in sealed envelopes into study binders".
Blinding personnel (performance bias)	-	Not possible for practice staff and patients
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	"Baseline sociodemographic and medical data were collected from electronic medical record by blinded study staff and from patients prior to randomisation". Longitudinal postdischarge outcome assessments were completed by telephone by staff blind to study condition".
Incomplete outcome data addressed (attrition bias)	+	Outcome was assessed at 24 weeks. In the intervention group 3 patients are lost to medical and/or mortality follow-up, 3 patients are lost to contact. In the control group 5 patients were lost to contact.
Selective reporting (reporting bias)	+	All outcomes specified in the article are reported.
Other	+	- it occurred in an academic medical center among

	<p>mostly white patients</p> <ul style="list-style-type: none"> – the intervention was delivered by research clinicians who have experience with the population and the CC program – ‘there was only 1 CM and our results may have been primarily due to the proficiency of this clinician’ – psychiatric diagnosis were not made using formal structured clinical interviews
--	--

Carney et al, 2016 [59]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	‘...patients... were randomly assigned in a 1:1 allocation ratio. Participants were randomized in permuted blocks...’
Allocation concealment (selection bias)	+	‘Treatment assignments were concealed in sequentially numbered, opaque envelopes (one set per stratum) and opened by the study coordinator after the baseline evaluation.’
Blinding personnel (performance bias)	–	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	‘The interviews, questionnaires, and medical chart reviews were performed by research staff who were blinded to the group status of the participants.’
Incomplete outcome data addressed (attrition bias)	+	92% of the participants completed the 6-months assessment.
Selective reporting (reporting bias)	+	All collected outcomes are reported.
Other	+	N/A

Kroenke et al, 2009 [61]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	‘Individuals were randomized to ... with randomization stratified by pain location (back vs leg) and clinic site (university vs Veterans Affairs medical center).’
Allocation concealment (selection bias)	+	‘Randomization lists were computer generated and treatment assignments were supplied in sealed opaque envelopes.’
Blinding personnel (performance bias)	–	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	‘All baseline and follow-up outcome assessments were conducted by a research assistant blinded to treatment allocation and uninvolved in the care management of the participants.’
Incomplete outcome data addressed (attrition bias)	+	‘there were no significant baseline differences between the intervention and usual care groups.’ Methods: ‘Enrolling 250 individuals allowed up to a 25% attrition rate.’ Attrition in the intervention group was 21/123 =17%; and in the CAU group 24/127=19%.
Selective reporting (reporting bias)	+	All collected outcomes are reported.
Other	+	‘Because only one-third of eligible patients agreed to enroll, the extent to which the benefits we found are

generalizable to patients not desiring participation in a trial is uncertain.'

Lin et al, 2003 [60]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	"The random assignment was stratified by recruitment method (screening or referral) or clinic. Within each stratum, participants were assigned according to a random number sequence that was developed using a computer random number generator at the coordinating center".
Allocation concealment (selection bias)	+	"Random assignment information was contained in a set of numbered, sealed envelopes for each stratum that were used sequentially for newly enrolled patients at each clinic".
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	"Baseline interviews were conducted by trained lay interviewers using structured computerized interviews before randomisation; thus, the interviewers were blind to study assignment. Blind follow-up interviews were performed at 3,6 and 12 months".
Incomplete outcome data addressed (attrition bias) (Short-term outcomes (2-6 weeks))	+	Outcome was assessed at 3-6-12 months. A total of 484 intervention patients (n=506) (95.7%) and 480 usual care patients (n=495) (97.0%) completed all 3 assessments (3, 6 and 12 months). Reasons for missing outcome data were: non-respondents and deceased. "We used a multiple imputation technique to account for item- and unit-level nonresponse and uncertainty in imputed values". "Intervention and control group comparisons showed no significant difference in clinical or demographic characteristics at baseline (P>0.1)".
Selective reporting (reporting bias)	+	All outcomes specified in the article are reported.
Other	+	"Lack of information on analgesic medication use at follow-up limits interpretation of the findings, because we can not preclude the possibility that intervention patients received more aggressive analgesic treatment".

Dwight-Johnson et al, 2005 [62]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	'computer-generated random assignment to the treatment condition'
Allocation concealment (selection bias)	+	'the study recruiter opened a sealed envelope'
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	The baseline, 4-, and 8-month assessments were conducted by telephone in English and Spanish by bilingual interviewers blind to intervention status.

Incomplete outcome data addressed (attrition bias)	+	'Thirty-six patients completed either one or both of the follow-up assessments at 4 months (collaborative intervention: N=15, usual care: N=15) and 8 months (N=17 and 12, respectively). The only significant difference was that dropouts were four times as likely as completers to report having important personal or family problems at baseline (OR=4.3, 95% CI=1.1-16.8; p=0.03). Dropout rates did not significantly differ between the collaborative intervention and usual care groups.'
Selective reporting (reporting bias)	+	All expected outcomes are reported.
Other	+	Relatively small pilot study.

Ell et al, 2008 [63]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	'...randomly determined by computer algorithm.'
Allocation concealment (selection bias)	+	'Patients were assigned to intervention or enhanced usual care (EUC) via the recruiter providing the patient with a choice from five sealed envelopes that contained one sheet of paper indicating a study group...'
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	'The baseline interview was conducted before random assignment.' 'Blinded Assessment' at 6 and 12 months.
Incomplete outcome data addressed (attrition bias)	+	The rates of mortality and patients lost to follow-up were not significantly different between groups (Fig 1). By 12 months, total attrition had increased to 214 patients (45.3%), with 68 known patient deaths (14.4%), three patients (0.6%) in hospice or palliative care, and 143 patients (30.3%) lost to follow-up (with 62 patients declining further study participation and 81 patients being unable to be located). 'In this article, we present results with all available data.' 'Intervention effects on depression and quality-of-life outcomes analyzed with imputed data were consistent with the results analyzed with all available data.' 'The rates of death and patients lost to follow-up were high but were not significantly different between the intervention and EUC groups. Moreover, comparison of patients who did and did not remain in the trial found no significant difference in baseline depression severity or quality-of-life outcomes.'
Selective reporting (reporting bias)	+	All expected outcomes are reported.
Other	+	N/A

Strong et al, 2008 [64]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	-	Patients were randomly assigned to either usual care alone, or usual care supplemented with the intervention. No method of randomization is described.
Allocation concealment (selection bias)	+	A computer programme was used to allocate patients to the two groups, with minimisation for sex, age, primary cancer site, and extent of disease.

Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	Persistence of a diagnosis of major depressive disorder was assessed by interviewers who were unaware of treatment allocation. To compensate for possible bias caused by the interviewers discovering treatment allocations during interviews, all the interviews were audio-recorded and edited to remove such information. An independent investigator reassessed these recordings.
Incomplete outcome data addressed (attrition bias)	+	200 patients randomly assigned. 99 usual care group. At 3 months follow up: 99 (100%). 80 with data at all timepoints (19,2% missing). 101 intervention group. At 3 months follow up: 97. 85 with data at all timepoints (15.8% missing).
Selective reporting (reporting bias)	+	Data collected on 3, 6 and 12 months. Reported data for 3, 6 and 12 months.
Other	+	- For ethical reasons usual care was optimised by informing primary care doctors that their patients had major depressive disorder. Therefore prescription of antidepressants was higher in both groups than previously reported for patients who receive usual care alone. Possible selection bias: patients with poor prognosis were excluded.

Alexopoulos et al, 2014 [65]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	'...participants were randomized (1:1) into ... in blocks of five using random numbers...'
Allocation concealment (selection bias)	+	'...participants were randomized ... using random numbers provided by our Biostatistics Unit.'
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	'Assessments were conducted by trained assistants blinded to the nature of the intervention, study goals, and randomization. One of the authors (P.J.R.), also blinded to treatment assignment, provided training, conducted rating reliability studies every 3 months, and reviewed all assessments before assigning final ratings.'
Incomplete outcome data addressed (attrition bias)	+	Short term: 'With the exception of education, there were no significant differences among participants assigned to PID-C or UC.' 'Eighteen percent of PID-C participants and 17% of UC participants died during the intervention phase (discharge-28 weeks). Other attrition was 25% in the PID-C arm and 17% in the UC arm.' Longer term: 'Mortality was lower during follow-up (28-52 weeks). Attrition from other causes was 24% compared with 26%. There were no differences in overall rates of attrition between the two arms at 28 weeks. There were no significant differences in age, education, severity of depression, or dyspnea-related disability between those who withdrew for reasons other than death and those who remained in the study.'
Selective reporting (reporting bias)	+	All collected outcomes are reported.
Other	+	N/A

Bogner et al, 2010 [67]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Unclear	'patients ... were randomly assigned'. Insufficient information about the sequence generation process to permit judgement.
Allocation concealment (selection bias)	Unclear	Insufficient information to permit judgement: the method of concealment is not described.
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	Unclear	The study did not address this outcome.
Incomplete outcome data addressed (attrition bias)	+	Characteristics of patients who refused to participate in the study were compared with those who participated, no significant differences were found. All participants attended the required sessions at baseline, 2, 4, and 6 weeks. All participants attended the final follow up at 12 weeks.
Selective reporting (reporting bias)	+	All expected outcomes are reported.
Other	+	N/A

Ell et al, 2010 [68]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	"trained bilingual study recruiters identified diabetic patients from medical charts". Randomisation was conducted via a computer-generated random number in blocks of 10"
Allocation concealment (selection bias)	+	"with random assignment enclosed in sealed envelopes".
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+/-	Insufficient information to permit judgement. The study did not address this outcome.
Incomplete outcome data addressed (attrition bias) (Short-term outcomes (2-6 weeks))	+/-	Follow-up study attrition did not vary between groups (21.6% in EUC group vs 21,8% in the INT group at 6 months, 28.4% in the EUC group vs 26.4% in the INT group at 12 months, and 29.4% in the EUC group vs 25.4% in the INT group at 18 months) Total 106 (27%) excluded from analysis, "enrolling 350 patients allowed a 30% attrition rate". Baseline depression, quality of life clinical outcomes and most demographics were similar between attrition and retention groups
Selective reporting (reporting bias)	+	All outcomes specified in the article are reported.
Other	-	- The study design may have biased the comparisons in favor of the EUC group. Increased prescription of antidepressant medications over baseline among usual care patients may be attributable in part to the enhanced usual care design.

- The same practitioners treated both intervention and UC patients which give a spill-over effect on quality of depression treatment”.
- The reliance on self-reports of chronic medical conditions can be questioned
- Potentially under reporting of use of antidepressants and psychotherapy.

Katon et al, 2004 [66]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	“Using a computerized algorithm, patients were randomized to the intervention or usual care group”.
Allocation concealment (selection bias)	Unclear	Insufficient information to permit judgement. The method of concealment is not described.
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	“After randomisation, telephone interviews were provided at 3,6 and 12 months by a telephone survey team who were blinded tot intervention status”.
Incomplete outcome data addressed (attrition bias) (Short-term outcomes (2–6 weeks))	+	Outcome was assessed at 3–6–12 months. A total of 132 intervention patients (80.5%) and 131 usual care patients (79.4%) completed all 3 assessments (3, 6 and 12 months).
Selective reporting (reporting bias)	+	All outcomes specified in the article are reported.
Other	+	This RCT was completed in 1 large health care system, limiting generalizability. The relatively low baseline HbA1c levels in this primary care population may have limited the effectiveness of the intervention on glycemic control

Ciechanowski et al, 2010 [69]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	‘We used computer-generated random assignment and block randomized to a 50:50 allocation ratio in each group depending on whether or not seizures were reported in the preceding 6 months.’
Allocation concealment (selection bias)	+	‘Concealment of assignment was ensured until randomization by having assistants without an intervention role generate randomization sequence, enroll patients, and assign patients to treatment groups.’
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	‘Outcome assessments were conducted in subjects’ homes at baseline, and by telephone at 6 and 12 months by blinded interviewers’
Incomplete outcome data addressed (attrition bias)	+	‘Ten subjects did not have 6- or 12-month assessments and were compared with the 70 subjects on all baseline variables. Those not completing follow-up assessments were younger (M=36.8, SD=5.2) than those who did (M=44.9, SD=11.2) [t(78)=3.82, P=0.001].’
Selective reporting (reporting bias)	+	All expected outcomes are reported.
Other	+	N/A

Pyne et al, 2011 [70]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	Participants were randomized to the intervention or to usual care in a 1:1 ratio according to a computer-generated random assignment sequence stratified by clinic and generated in advance.
Allocation concealment (selection bias)	+	Research assistants at each clinic were provided envelopes labeled by participant number and containing randomized assignment.
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	Baseline and 6- and 12-month data were collected by telephone interviewers who were blinded to treatment assignment and used scripted computer-based assessments.
Incomplete outcome data addressed (attrition bias)	+	276 patients. Usual care group n=138, 3 months follow-up n=126 (missing 12%). intervention group n=138. 3 months follow-up n=123 (missing 10.1%). 12 months follow up n=110 (missing 20.3%). Intervention group n=138. 12 months follow-up n=105 (missing 23.9%).
Selective reporting (reporting bias)	+	Baseline and 6- and 12-month data were collected. In the article data on baseline, 6 and 12 months are reported.
Other	+	Weaknesses included the potential lack of generalizability from VA to other treatment settings and the use of self-reported medication adherence data.

Bogner et al, 2008 [71]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+/-	'patients ... were randomly assigned'. Insufficient information about the sequence generation process to permit judgement.
Allocation concealment (selection bias)	+/-	Insufficient information to permit judgement: the method of concealment is not described.
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+/-	Insufficient information to permit judgement.
Incomplete outcome data addressed (attrition bias)	+	Characteristics of patients who refused to participate in the study were compared with those who participated, no significant differences were found. All participants attended the required sessions at baseline, 2, 4, and 6 weeks.
Selective reporting (reporting bias)	+	All expected outcomes are reported.
Other	+	N/A

Katon et al, 2010 [73]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	"Patients were assigned to a treatment group with the use of a permuted block design, with randomly selected

		block sizes of 4, 6 and 8 patients".
Allocation concealment (selection bias)	Unclear	Insufficient information to permit judgement. The method of concealment is not described.
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	"research assistants who were unaware of the intervention status implemented study procedures"
Incomplete outcome data addressed (attrition bias)	+	88% completed all 6 months telephone and laboratory assessments and 83% completed all 12 months assessment. The characteristics of the patients in the intervention group and the usual care group were similar at baseline.
Selective reporting (reporting bias)	+	The initial primary outcome was the percentage of patients achieving disease control at 12 months on all three disease-control measures (glycated hemoglobin level, systolic blood pressure and LDL cholesterol level). This outcome has changed in august 2009, before study data became available for analysis in november 2009. All other expected outcomes are reported.
Other	+	Usual care was enhanced in the trial by notification of primary care physicians about depressive illness and the results of baseline, 6 and 12 months laboratory tests

Rollman et al, 2009 [72]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	'Patients were randomized to either the intervention or "usual care" group in a 1:1 ratio in blocks of 4 according to a computer-generated random assignment sequence stratified by hospital site, prepared in advance by our statistician, entered into the computer-assisted telephone interview program used by our assessors, and concealed until after the 2-week call. A study nurse or project coordinator then informed patients of their treatment assignment and notified their PCP.'
Allocation concealment (selection bias)	+	computer generated randomization, concealed until the 2-weeks call
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	+	Telephone assessors were blinded as to patients randomization and baseline depression status and they cautioned subjects at the beginning of each call not to divulge their treatment assignment.
Incomplete outcome data addressed (attrition bias)	+	150 intervention group, 2 month 121 assessed (missing 19.3%), 152 usual care 136 assessed (missing 10.5%), comparison group n=151, 135 assessed (missing 10.6%). 8 month assessment: intervention group 126/150 assessed (missing 16%), usual care group 126/152 assessed (missing 17%), comparison group 141/151 assessed (missing 6.6%)
Selective reporting (reporting bias)	-	Data collected at 2 weeks, 2, 4 and 8 month follow-up. In article only data at 8 month follow up is reported.

Other	+	'Nurse-recruiters were required to obtain patients' prior consent through a hospital staff member before they could even approach to obtain consent to administer the PHQ-2. This potentially introduced a selection bias if severely depressed patients were less likely to participate in our screening procedures.'
-------	---	--

van Steenbergen et al, 2015 [34]

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	+	Randomization was conducted by investigators with a specialized computer system.
Allocation concealment (selection bias)	+	After enrolling participants, the computer allocated a randomisationgroup.
Blinding personnel (performance bias)	-	Not possible for practice staff and patients.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	-	Insufficient information to permit judgement.
Incomplete outcome data addressed (attrition bias)	-	Outcome was assessed at 3-6-9-12 months. Reasons for missing data are not mentioned. Missing data>27%. There were no significant differences in demographic and clinical characteristics
Selective reporting (reporting bias)	+	All collected outcomes are reported.
Other	+	N/A

Forest plot 1 Meta-analysis burden of disease based on the combined physical and depressive outcomes

Outcome: see data extraction table for specific outcome measures; CAU = care as usual; CC = collaborative care.

Forest plot 2 Meta-analysis combined physical outcomes

Outcome: see data extraction table for specific outcome measures; CAU = care as usual; CC = collaborative care.

Forest plot 3 Meta-analysis combined physical outcomes per chronic medical condition

Forest plot 4 Meta-analysis somatic diagnostic and treatment interventions

Forest plot 5 Meta-analysis depressive outcome

Funnel plot 1

Flow diagram 1

Table 1 Outcomes in meta-analysis

Outcome	Included in Meta-analysis as	
Physical outcomes		
Somatic diagnostic or intervention procedures	Somatic diagnostic or intervention procedures	
Symptom load related to chronic medical condition: <i>Cardiac events:</i> MACE Angina pectoris Post infarct <i>Arthritis:</i> Pain Physical functioning <i>Cancer:</i> Pain Physical functioning <i>COPD:</i> Dyspnea related disability <i>Diabetes Mellitus type 2:</i> HbA1C <i>Epilepsy:</i> Seizures <i>HIV:</i> HIV symptom severity <i>Hypertension:</i> Blood pressure <i>Multimorbidity:</i> Abovementioned outcomes corresponding with relevant chronic medical conditions combined Physical functioning	Combined physical outcomes	Burden of disease
Depression		
Depressive symptoms	Depressive symptoms	

Table 2 Data extraction table Collaborative Care (CC) versus Care As Usual (CAU), 20 RCTs, N = 4774

Study	N (completers) Mean age	Measure depression at baseline	Setting CC model Role caremanager	Outcome assessment Follow-up	Effect size	Comments
<i>Acute Coronary Syndrome (ACS) (Myocardial infarction or unstable angina) N = 866</i>					Physical outcome: 0.15 / Depression: 0.35	
Davidson et al, (2010), USA [55]	N = 157: CC 80, CAU 77 Mean age : CC 59.3, CAU 61.1	Persistent depressive symptoms: ≥ 10 on the Beck Depression Inventory (BDI).	5 general hospitals. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 2.89 MACE: Rate CC vs CAU OR: 3.46 Depression: OR: 2.41 FU 6 – 9 months	BOD: Δ : 0.58 MACE: Δ : 0.69 Depression: Δ : 0.48	Significant improvement in BOD. Not significant improvement in MACE. Significant improvement in depressive symptoms.
Davidson et al, (2013), USA, [58]	N = 150: CC 73, CAU 77 Mean age: 59.2 (SD 9.7)	Depression: BDI score ≥ 10 on 2 screening occasions or ≥ 15 on 1 occasion.	2 private and 5 academic ambulatory centers. CM + physician + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 1.66 Physical functioning: SF-36 physical subscale OR: 1.34 Depression: OR: 2.06 FU 6 months	BOD: Δ : 0.28 Physical functioning: Δ : 0.16 Depression: Δ : 2.06	Not significant improvement in BOD. Not significant improvement in physical functioning. Significant improvement in depressive symptoms.
Huffman et al., (2011), USA, [56]	N = 175: CC 90, CAU 85 Mean age CC: 62.1 (SD 12.8), CAU: 62.6 (SD 12.2)	Depression: PHQ-9 ≥ 10 .	3 inpatient cardiac units at an urban academic medical centre for acute cardiac disease. CM + GP + psychiatrist. Monitoring.	BOD: OR: 1.57 Physical functioning: SF-12 PCS OR: 1.04 Depression: OR: 2.37	BOD: Δ : 0.25 Physical functioning: Δ : 0.02 Depression: Δ : 0.48	Not significant improvement in BOD. Not significant improvement in physical functioning. Significant

				FU 6 months		improvement in depressive symptoms.
Huffman et al., (2014), USA, [57]	N = 183: CC 92, CAU 91 Mean age: 60.5 (12.7)	Major depression: PHQ-9 $\geq$ 10	3 cardiac units at an urban academic medical centre. CM + GP + psychiatrist. Monitoring.	BOD: OR: 1.41 Combined physical outcomes: MACE and physical functioning (SF-12 PCS) OR: 1.28 Depression: OR: 1.72 FU 24 weeks	BOD: Δ : 0.19 Combined physical outcomes: Δ : 0.14 Depression: Δ : 0.30	Not significant improvement in BOD. Not significant improvement in combined physical outcomes. Significant improvement in depressive symptoms.
Carney et al, (2016), USA, [59]	N = 201: CC 100, CAU 101 Mean age CC: 63.1 (SD 10.0), CAU: 63.0 (SD 9.5)	Depression: PHQ-9 $\geq$ 10.	1 outpatient cardiology service at the university and 1 suburban satellite facility. CM + GP + psychiatrist. Monitoring.	BOD: OR: 1.36 Physical functioning: PROMIS OR: 1.34 Depression: OR: 1.38 FU 6 months	BOD: Δ : 0.17 Physical functioning: Δ : 0.16 Depression: Δ : 0.18	Not significant improvement in BOD. Not significant improvement in physical functioning. Not significant improvement in depressive symptoms.
<i>Arthritis N = 1251</i>					Physical outcome: 0.25 / Depression: 0.29	
Kroenke et al, (2009), USA, [61]	N = 250: CC 123, CAU 127 Mean age CC: 55.2 (SD 12.6); CAU 55.8 (SD 11.1)	Depression: PHQ-9 $\geq$ 10.	2 primary care clinical systems. CM + physician + psychiatrist. Monitoring.	BOD: OR: 2.05 Combined physical outcomes: pain (graded chronic pain scale severity score) and physical functioning (disability score) OR: 2.00 Depression: OR: 2.15 FU 6–12 months	BOD: Δ : 0.40 Combined physical outcomes: Δ : 0.38 Depression: Δ : 0.42	Significant improvement in BOD. Significant improvement in combined physical functioning. Significant improvement in depressive symptoms.

Lin et al., (2003), USA, [60]	N = 1001: CC 506, CAU 495 Mean age 72 (SD 7.4)	Major depression and dysthymia: SCID I and severity of depression symptoms by SCL-20.	18 primary care clinics from 8 health care organisations. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 1.40 Combined physical outcomes: pain (pain intensity) and physical functioning (interference of arthritis with daily functioning) OR: 1.37 Depression: OR: 1.46 FU 12 months	BOD: Δ: 0.19 Combined physical outcomes: Δ: 0.17 Depression: Δ: 0.21	Significant improvement in BOD. Significant improvement in combined physical functioning. Significant improvement in depressive symptoms.
<i>Cancer, N = 727</i>					Physical outcome: 0.16 / Depression: 0.40	
Dwight-Johnson et al, (2005), USA, [62]	N = 55: CC 28, CAU 27 Mean age CC 47.7 (SD 11.9), CAU 46.8 (SD 10.8)	Major depression or dysthymia or persistent depressive symptoms: PHQ- 9 (no minimum score provided), three additional questions from PRIME-MD were used to assess dysthymia.	Outpatient breast and gynaecology clinics of 1 medical centre. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 2.56 Physical functioning: functional assessment of cancer therapy OR: 1.48 Depression: OR: 4.44 FU 8 months	BOD: Δ: 0.52 Physical functioning: Δ: 0.22 Depression: Δ: 0.82	Not significant improvement in BOD. Not significant improvement in physical functioning. Significant improvement in depressive symptoms.
Ell et al, (2008), USA, [63]	N = 472: CC 242, CAU 230 Mean age and range unknown, 49.4% ≥ 50 years	Major depression or dysthymia: PHQ-9 ≥ 10, or dysthymia according to the SCID-IV.	Medical oncology clinics. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 1.55 Combined physical outcomes: pain (BPI) and physical functioning (SF-12 physical component scale) OR: 1.43 Depression: OR: 1.83 FU 12 months	BOD: Δ: 0.24 Combined physical outcomes: Δ: 0.20 Depression: Δ: 0.33	Significant improvement in BOD. Significant improvement in combined physical functioning. Significant improvement in depressive symptoms.
Strong et al,	N = 200: CC 101,	Major depressive	1 regional tertiary	BOD: OR: 1.58	BOD: Δ: 0.25	Not significant

(2008), Scotland (UK), [64]	CAU 99 Mean age CC 56.6 (SD 11.4), CAU 56.6 (SD 12.3)	disorder: Hospital Anxiety and Depression Scale score >15, SCID-IV: major depression, Symptom Checklist-20.	cancer centre, primary care physicians. CM +GP. Monitoring.	Combined physical outcomes: pain and physical functioning (both: EORTC QLQ C30) OR: 1.14 Somatic diagnostic and treatment procedures Depression: OR: 2.23 FU 3 months	Combined physical outcomes: Δ : 0.07 Somatic diagnostic and treatment procedures Depression: Δ : 0.44	improvement in BOD. Not significant improvement in combined physical outcomes. Significant improvement in depressive symptoms.
<i>COPD, N = 138</i>					Physical outcome: 0.35 / Depression: 0.25	
Alexopoulos et al, (2014), USA, [65]	N = 138: CC 67, CAU 71 . Mean age CC 70.9 (54-86), CAU 71.0 (54-86)	Unipolar major depression: Structured Clinical Interview for DSM-IV, and a score of ≥ 14 on the 17-item Hamilton Depression Rating Scale.	First session of intervention in an acute inpatient pulmonary rehabilitation unit, remaining sessions in primary care setting. CM +GP. Monitoring.	BOD: OR: 1.72 Physical functioning: dyspnea-related disability OR: 1.87 Depression: OR:1.58 FU 28 weeks	BOD: Δ : 0.30 Physical functioning: Δ : 0.35 Depression: Δ : 0.25	Not significant improvement in BOD. Significant improvement in physical functioning. Not significant improvement in depressive symptoms.
<i>Diabetes Mellitus type 2, N = 774</i>					Physical outcome: 0.15 / Depression: 0.36	
Bogner et al, (2010), USA, [67]	N = 58: CC 29, CAU 29 Mean age 60.2 (50-80).	Depression: CES-D.	1 primary care practice. CM + GP. Monitoring.	BOD: OR: 2.97 HbA1C: OR: 3.16 Depression: OR: 2.80 FU 12 weeks	BOD: Δ : 0.60 HbA1C: Δ : 0.63 Depression: Δ : 0.57	Significant improvement in BOD. Significant improvement in HbA1C. Significant improvement in depressive symptoms.
Ell et al, (2010), USA, [68]	N = 387: CC 193, CAU 194	Depression: PHQ-9 ≥ 10 .	2 public safety-net clinics.	BOD: OR: 1.38	BOD: Δ : 0.18	Not significant improvement in

	≥ 18 years; mean age unknown		CM + GP + psychiatrist. Monitoring and psychotherapy.	HbA1C: OR: 1.04 Depression: OR: 1.84 FU 6 months	HbA1C: Δ: 0.02 Depression: Δ: 0.34	BOD. Not significant improvement in HbA1C. Significant improvement in depressive symptoms.
Katon et al, (2004), USA, [66]	N = 329: CC 164, CAU 165 Mean age CC: 58.6 (SD 11.8), CAU 58.1 (12.0)	Major depression and/or dysthymia: PHQ-9 ≥ 10, SCL-90 depression mean item score > 1.1.	9 primary care clinics. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 1.51 HbA1C: OR: 1.20 Depression: OR: 1.91 FU 12 months	BOD: Δ: 0.23 HbA1C: Δ: 0.10 Depression: Δ: 0.36	Not significant improvement in BOD. Not significant improvement in HbA1C. Significant improvement in depressive symptoms.
<i>Epilepsy, N = 80</i>					Physical outcome: 0.01 / Depression: 0.79	
Ciechanowski et al, (2010), USA, [69]	N = 80: CC 40, CAU 40 Mean age: 43.9 (SD 11.0)	Minor or major depression or dysthymic disorder: PHQ-9 ≥ 10, or SCID items indicating depressive symptoms more days than not for 2 years.	Regional epilepsy center and neurology clinics CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 1.61 Seizures: OR: 1.01 Depression: OR: 4.12 FU 6– 12 months	BOD: Δ: 0.26 Seizures: Δ: 0.01 Depression: Δ: 0.79	Not significant improvement in BOD. Not significant improvement in seizures. Significant improvement in depressive symptoms.
<i>HIV, N = 276</i>					Physical outcome: 0.42 / Depression: 0.47	
Pyne et al, (2011), USA, [70]	N = 276: CC 123, CAU 126 Mean age: 49.8 (SD 10.5)	Depression: PHQ-9 ≥ 10.	3 Veterans Affairs HIV treatment facilities. CM + GP + psychiatrist. Monitoring.	BOD: OR: 2.22 HIV symptom severity: OR: 2.15 Somatic diagnostic and treatment procedures: Depression: OR: 2.36	BOD: Δ: 0.44 HIV symptom severity: Δ: 0.42 Somatic diagnostic and treatment procedures: Depression: Δ: 0.47	Significant improvement in BOD. Significant improvement in HIV symptom severity. Significant improvement in

				FU 6 months		depressive symptoms.
<i>Hypertension, N = 64</i>					Physical outcome: 0.79 / Depression: 0.71	
Bogner et al, (2008), USA, [71]	N = 64: CC 32, CAU 32 Mean age: 58.6 (50–80).	Depression: MINI and CES-D.	1 primary care practice. CM + GP. Monitoring.	BOD: OR: 3.99 Combined physical outcomes: systolic and diastolic blood pressure OR: 4.19 Depression: OR: 3.63 FU 6 weeks	BOD: Δ : 0.76 Combined physical outcomes: Δ : 0.79 Depression: Δ : 0.71	Significant improvement in BOD. Not significant improvement in combined physical outcomes. Significant improvement in depressive symptoms.
<i>Multimorbidity, N = 598</i>					Physical outcome: 0.28 / Depression: 0.46	
Katon et al, (2010), USA, [73]	N = 214: CC 106, CAU 108 Mean age CC: 57.4 (SD 10.5), CAU: 56.3 (SD 12.1)	Depression: PHQ-9 $\geq$ 10.	14 primary care clinics. CM + physician + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 2.09 Combined physical outcomes: HbA1C, systolic and diastolic blood pressure OR: 1.95 Depression: OR: 2.39 FU 6 – 12 months	BOD: Δ : 0.41 Combined physical outcomes: Δ : 0.37 Depression: Δ : 0.48	Significant improvement in BOD. Significant improvement in combined physical outcomes. Significant improvement in depressive symptoms.
Rollman et al, (2009), USA, [72]	N = 302: CC 150, CAU 152 Mean age CC: 64 (SD 10.8), CAU: 64 (11.2)	Depression: PHQ-9 $\geq$ 10.	2 university-affiliated and 5 community hospitals. CM + GP + psychiatrist. Monitoring.	BOD: OR: 1.97 Physical functioning: DASI 1.36 Depression: OR: 2.38 FU 8 months	BOD: Δ : 0.37 Physical functioning: Δ : 0.17 Depression: Δ : 0.48	Significant improvement in BOD. Significant improvement in physical functioning. Significant improvement in depressive

						symptoms.
van Steenberg et al, (2015), The Netherlands, [34]	N = 82: CC 42, CAU 42 Mean age CC 57 (SD 11.6), CAU 60 (11.6)	Major depression: PHQ-9 ≥ 15 . Depression: PHQ-9 ≥ 10 .	Outpatients clinics in 5 hospitals. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 2.04 Somatic diagnostic and treatment procedures : OR: 2.68 Depression: OR: 1.56 FU 3 months	BOD: Δ: 0.39 Somatic diagnostic and treatment procedures: Δ : 0.54 Depression: Δ: 0.24	Not significant improvement in BOD. Not significant improvement in somatic diagnostic and treatment procedures. Not significant improvement in depressive symptoms.

FU = follow-up; CC = collaborative care; CAU = care as usual; BOD = burden of disease combined for physical and depression outcome; MACE: Major Adverse Cardiac Event; EORTC QLQ C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; PHQ-9 = Patient Health Questionnaire-9; MINI = Mini Internationaal Neuropsychiatrisch Interview; CES-D = Center for Epidemiologic Studies Depression Scale; SF-12 PCS = 12-Item Short Form Health Survey Physical Component Summary; SF-36 PCS = 36-Item Short Form Health Survey Physical Component Summary; PROMIS = Common Fund's Patient-Reported Outcomes Measurement Information System; DASI = Duke Activity Status Index.

Table 3 Data extraction table Collaborative Care (CC) versus Care As Usual (CAU), 20 RCTs, N = 4774

Study	N (completers) Mean age	Measure depression at baseline	Setting CC model Role caremanager	Outcome assessment Follow-up	Effect size	Comments
Cardiac events (ACS, myocardial infarction, unstable angina, heart failure, arrhythmia) N = 866					Physical outcome: 0.15 / Depression: 0.35	
Davidson et al, (2010), USA [55]	N = 157: CC 80, CAU 77 Mean age : CC 59.3, CAU 61.1	Persistent depressive symptoms: ≥ 10 on the Beck Depression Inventory (BDI).	5 general hospitals. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 2.89 MACE: Rate CC vs CAU OR: 3.46 Depression: OR: 2.41 FU 6 – 9 months	BOD: Δ : 0.58 MACE: Δ : 0.69 Depression: Δ : 0.48	Significant improvement in BOD. Not significant improvement in MACE. Significant improvement in depressive symptoms.
Davidson et al, (2013), USA, [58]	N = 150: CC 73, CAU 77 Mean age: 59.2 (SD 9.7)	Depression: BDI score ≥ 10 on 2 screening occasions or ≥ 15 on 1 occasion.	2 private and 5 academic ambulatory centers. CM + physician + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 1.66 Physical functioning: SF-36 physical subscale OR: 1.34 Depression: OR: 2.06 FU 6 months	BOD: Δ : 0.28 Physical functioning: Δ : 0.16 Depression: Δ : 2.06	Not significant improvement in BOD. Not significant improvement in physical functioning. Significant improvement in depressive symptoms.
Huffman et al., (2011), USA, [56]	N = 175: CC 90, CAU 85 Mean age CC: 62.1 (SD 12.8), CAU: 62.6 (SD 12.2)	Depression: PHQ-9 ≥ 10 .	3 inpatient cardiac units at an urban academic medical centre for acute cardiac disease. CM + GP + psychiatrist. Monitoring.	BOD: OR: 1.57 Physical functioning: SF-12 PCS OR: 1.04 Depression: OR: 2.37	BOD: Δ : 0.25 Physical functioning: Δ : 0.02 Depression: Δ : 0.48	Not significant improvement in BOD. Not significant improvement in physical functioning. Significant

				FU 6 months		improvement in depressive symptoms.
Huffman et al., (2014), USA, [57]	N = 183: CC 92, CAU 91 Mean age: 60.5 (12.7)	Major depression: PHQ-9 $\geq$ 10	3 cardiac units at an urban academic medical centre. CM + GP + psychiatrist. Monitoring.	BOD: OR: 1.41 Combined physical outcomes: MACE and physical functioning (SF-12 PCS) OR: 1.28 Depression: OR: 1.72 FU 24 weeks	BOD: Δ : 0.19 Combined physical outcomes: Δ : 0.14 Depression: Δ : 0.30	Not significant improvement in BOD. Not significant improvement in combined physical outcomes. Significant improvement in depressive symptoms.
Carney et al, (2016), USA, [59]	N = 201: CC 100, CAU 101 Mean age CC: 63.1 (SD 10.0), CAU: 63.0 (SD 9.5)	Depression: PHQ-9 $\geq$ 10.	1 outpatient cardiology service at the university and 1 suburban satellite facility. CM + GP + psychiatrist. Monitoring.	BOD: OR: 1.36 Physical functioning: PROMIS OR: 1.34 Depression: OR: 1.38 FU 6 months	BOD: Δ : 0.17 Physical functioning: Δ : 0.16 Depression: Δ : 0.18	Not significant improvement in BOD. Not significant improvement in physical functioning. Not significant improvement in depressive symptoms.
<i>Arthritis N = 1251</i>					Physical outcome: 0.25 / Depression: 0.29	
Kroenke et al, (2009), USA, [61]	N = 250: CC 123, CAU 127 Mean age CC: 55.2 (SD 12.6); CAU 55.8 (SD 11.1)	Depression: PHQ-9 $\geq$ 10.	2 primary care clinical systems. CM + physician + psychiatrist. Monitoring.	BOD: OR: 2.05 Combined physical outcomes: pain (graded chronic pain scale severity score) and physical functioning (disability score) OR: 2.00 Depression: OR: 2.15 FU 6–12 months	BOD: Δ : 0.40 Combined physical outcomes: Δ : 0.38 Depression: Δ : 0.42	Significant improvement in BOD. Significant improvement in combined physical functioning. Significant improvement in depressive symptoms.

Lin et al., (2003), USA, [60]	N = 1001: CC 506, CAU 495 Mean age 72 (SD 7.4)	Major depression and dysthymia: SCID I and severity of depression symptoms by SCL-20.	18 primary care clinics from 8 health care organisations. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 1.40 Combined physical outcomes: pain (pain intensity) and physical functioning (interference of arthritis with daily functioning) OR: 1.37 Depression: OR: 1.46 FU 12 months	BOD: Δ : 0.19 Combined physical outcomes: Δ : 0.17 Depression: Δ : 0.21	Significant improvement in BOD. Significant improvement in combined physical functioning. Significant improvement in depressive symptoms.
<i>Cancer, N = 727</i>					Physical outcome: 0.16 / Depression: 0.40	
Dwight-Johnson et al, (2005), USA, [62]	N = 55: CC 28, CAU 27 Mean age CC 47.7 (SD 11.9), CAU 46.8 (SD 10.8)	Major depression or dysthymia or persistent depressive symptoms: PHQ- 9 (no minimum score provided), three additional questions from PRIME-MD were used to assess dysthymia.	Outpatient breast and gynaecology clinics of 1 medical centre. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 2.56 Physical functioning: functional assessment of cancer therapy OR: 1.48 Depression: OR: 4.44 FU 8 months	BOD: Δ : 0.52 Physical functioning: Δ : 0.22 Depression: Δ : 0.82	Not significant improvement in BOD. Not significant improvement in physical functioning. Significant improvement in depressive symptoms.
Ell et al, (2008), USA, [63]	N = 472: CC 242, CAU 230 Mean age and range unknown, 49.4% $\geq$ 50 years	Major depression or dysthymia: PHQ-9 $\geq$ 10, or dysthymia according to the SCID-IV.	Medical oncology clinics. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 1.55 Combined physical outcomes: pain (BPI) and physical functioning (SF-12 physical component scale) OR: 1.43 Depression: OR: 1.83 FU 12 months	BOD: Δ : 0.24 Combined physical outcomes: Δ : 0.20 Depression: Δ : 0.33	Significant improvement in BOD. Significant improvement in combined physical functioning. Significant improvement in depressive symptoms.
Strong et al,	N = 200: CC 101,	Major depressive	1 regional tertiary	BOD: OR: 1.58	BOD: Δ : 0.25	Not significant

(2008), Scotland (UK), [64]	CAU 99 Mean age CC 56.6 (SD 11.4), CAU 56.6 (SD 12.3)	disorder: Hospital Anxiety and Depression Scale score >15, SCID-IV: major depression, Symptom Checklist-20.	cancer centre, primary care physicians. CM +GP. Monitoring.	Combined physical outcomes: pain and physical functioning (both: EORTC QLQ C30) OR: 1.14 Somatic diagnostic and treatment procedures Depression: OR: 2.23 FU 3 months	Combined physical outcomes: Δ : 0.07 Somatic diagnostic and treatment procedures Depression: Δ : 0.44	improvement in BOD. Not significant improvement in combined physical outcomes. Significant improvement in depressive symptoms.
<i>COPD, N = 138</i>					Physical outcome: 0.35 / Depression: 0.25	
Alexopoulos et al, (2014), USA, [65]	N = 138: CC 67, CAU 71 . Mean age CC 70.9 (54-86), CAU 71.0 (54-86)	Unipolar major depression: Structured Clinical Interview for DSM-IV, and a score of ≥ 14 on the 17-item Hamilton Depression Rating Scale.	First session of intervention in an acute inpatient pulmonary rehabilitation unit, remaining sessions in primary care setting. CM +GP. Monitoring.	BOD: OR: 1.72 Physical functioning: dyspnea-related disability OR: 1.87 Depression: OR:1.58 FU 28 weeks	BOD: Δ : 0.30 Physical functioning: Δ : 0.35 Depression: Δ : 0.25	Not significant improvement in BOD. Significant improvement in physical functioning. Not significant improvement in depressive symptoms.
<i>Diabetes Mellitus type 2, N = 774</i>					Physical outcome: 0.15 / Depression: 0.36	
Bogner et al, (2010), USA, [67]	N = 58: CC 29, CAU 29 Mean age 60.2 (50-80).	Depression: CES-D.	1 primary care practice. CM + GP. Monitoring.	BOD: OR: 2.97 HbA1C: OR: 3.16 Depression: OR: 2.80 FU 12 weeks	BOD: Δ : 0.60 HbA1C: Δ : 0.63 Depression: Δ : 0.57	Significant improvement in BOD. Significant improvement in HbA1C. Significant improvement in depressive symptoms.
Ell et al, (2010), USA, [68]	N = 387: CC 193, CAU 194	Depression: PHQ-9 ≥ 10 .	2 public safety-net clinics.	BOD: OR: 1.38	BOD: Δ : 0.18	Not significant improvement in

	≥ 18 years; mean age unknown		CM + GP + psychiatrist. Monitoring and psychotherapy.	HbA1C: OR: 1.04 Depression: OR: 1.84 FU 6 months	HbA1C: Δ: 0.02 Depression: Δ: 0.34	BOD. Not significant improvement in HbA1C. Significant improvement in depressive symptoms.
Katon et al, (2004), USA, [66]	N = 329: CC 164, CAU 165 Mean age CC: 58.6 (SD 11.8), CAU 58.1 (12.0)	Major depression and/or dysthymia: PHQ-9 ≥ 10, SCL-90 depression mean item score > 1.1.	9 primary care clinics. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 1.51 HbA1C: OR: 1.20 Depression: OR: 1.91 FU 12 months	BOD: Δ: 0.23 HbA1C: Δ: 0.10 Depression: Δ: 0.36	Not significant improvement in BOD. Not significant improvement in HbA1C. Significant improvement in depressive symptoms.
<i>Epilepsy, N = 80</i>					Physical outcome: 0.01 / Depression: 0.79	
Ciechanowski et al, (2010), USA, [69]	N = 80: CC 40, CAU 40 Mean age: 43.9 (SD 11.0)	Minor or major depression or dysthymic disorder: PHQ-9 ≥ 10, or SCID items indicating depressive symptoms more days than not for 2 years.	Regional epilepsy center and neurology clinics CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 1.61 Seizures: OR: 1.01 Depression: OR: 4.12 FU 6– 12 months	BOD: Δ: 0.26 Seizures: Δ: 0.01 Depression: Δ: 0.79	Not significant improvement in BOD. Not significant improvement in seizures. Significant improvement in depressive symptoms.
<i>HIV, N = 276</i>					Physical outcome: 0.42 / Depression: 0.47	
Pyne et al, (2011), USA, [70]	N = 276: CC 123, CAU 126 Mean age: 49.8 (SD 10.5)	Depression: PHQ-9 ≥ 10.	3 Veterans Affairs HIV treatment facilities. CM + GP + psychiatrist. Monitoring.	BOD: OR: 2.22 HIV symptom severity: OR: 2.15 Somatic diagnostic and treatment procedures: Depression: OR: 2.36	BOD: Δ: 0.44 HIV symptom severity: Δ: 0.42 Somatic diagnostic and treatment procedures: Depression: Δ: 0.47	Significant improvement in BOD. Significant improvement in HIV symptom severity. Significant improvement in

				FU 6 months		depressive symptoms.
<i>Hypertension, N = 64</i>					Physical outcome: 0.79 / Depression: 0.71	
Bogner et al, (2008), USA, [71]	N = 64: CC 32, CAU 32 Mean age: 58.6 (50–80).	Depression: MINI and CES-D.	1 primary care practice. CM + GP. Monitoring.	BOD: OR: 3.99 Combined physical outcomes: systolic and diastolic blood pressure OR: 4.19 Depression: OR: 3.63 FU 6 weeks	BOD: Δ : 0.76 Combined physical outcomes: Δ : 0.79 Depression: Δ : 0.71	Significant improvement in BOD. Not significant improvement in combined physical outcomes. Significant improvement in depressive symptoms.
<i>Multimorbidity, N = 598</i>					Physical outcome: 0.28 / Depression: 0.46	
Katon et al, (2010), USA, [73]	N = 214: CC 106, CAU 108 Mean age CC: 57.4 (SD 10.5), CAU: 56.3 (SD 12.1)	Depression: PHQ-9 $\geq$ 10.	14 primary care clinics. CM + physician + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 2.09 Combined physical outcomes: HbA1C, systolic and diastolic blood pressure OR: 1.95 Depression: OR: 2.39 FU 6 – 12 months	BOD: Δ : 0.41 Combined physical outcomes: Δ : 0.37 Depression: Δ : 0.48	Significant improvement in BOD. Significant improvement in combined physical outcomes. Significant improvement in depressive symptoms.
Rollman et al, (2009), USA, [72]	N = 302: CC 150, CAU 152 Mean age CC: 64 (SD 10.8), CAU: 64 (11.2)	Depression: PHQ-9 $\geq$ 10.	2 university-affiliated and 5 community hospitals. CM + GP + psychiatrist. Monitoring.	BOD: OR: 1.97 Physical functioning: DASI 1.36 Depression: OR: 2.38 FU 8 months	BOD: Δ : 0.37 Physical functioning: Δ : 0.17 Depression: Δ : 0.48	Significant improvement in BOD. Significant improvement in physical functioning. Significant improvement in depressive

						symptoms.
van Steenberg et al, (2015), The Netherlands, [34]	N = 82: CC 42, CAU 42 Mean age CC 57 (SD 11.6), CAU 60 (11.6)	Major depression: PHQ- 9 $\geq$ 15. Depression: PHQ-9 $\geq$ 10.	Outpatients clinics in 5 hospitals. CM + GP + psychiatrist. Monitoring and psychotherapy.	BOD: OR: 2.04 Somatic diagnostic and treatment procedures : OR: 2.68 Depression: OR: 1.56 FU 3 months	BOD: Δ: 0.39 Somatic diagnostic and treatment procedures: Δ : 0.54 Depression: Δ: 0.24	Not significant improvement in BOD. Not significant improvement in somatic diagnostic and treatment procedures. Not significant improvement in depressive symptoms.

FU = follow-up; CC = collaborative care; CAU = care as usual; BOD = burden of disease combined for physical and depression outcome; ACS = Acute Coronary Syndrome; MACE: Major Adverse Cardiac Event; EORTC QLQ C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; PHQ-9 = Patient Health Questionnaire-9; MINI = Mini Internationaal Neuropsychiatrisch Interview; CES-D = Center for Epidemiologic Studies Depression Scale; SF-12 PCS = 12-Item Short Form Health Survey Physical Component Summary; SF-36 PCS = 36-Item Short Form Health Survey Physical Component Summary; PROMIS = Common Fund's Patient-Reported Outcomes Measurement Information System; DASI = Duke Activity Status Index.