

# BMJ Open How often are interventions in cluster-randomised controlled trials of complex interventions in general practices effective and reasons for potential shortcomings? Protocol and results of a feasibility project for a systematic review

Andrea Siebenhofer,<sup>1,2</sup> Stefanie Erckenbrecht,<sup>3</sup> Gudrun Pregartner,<sup>4</sup> Andrea Berghold,<sup>4</sup> Christiane Muth<sup>1</sup>

**To cite:** Siebenhofer A, Erckenbrecht S, Pregartner G, *et al.* How often are interventions in cluster-randomised controlled trials of complex interventions in general practices effective and reasons for potential shortcomings? Protocol and results of a feasibility project for a systematic review. *BMJ Open* 2016;**6**:e009414. doi:10.1136/bmjopen-2015-009414

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/bmjopen-2015-009414>).

AS and SE contributed equally.

Received 16 July 2015  
Revised 11 September 2015  
Accepted 23 October 2015



CrossMark

For numbered affiliations see end of article.

## Correspondence to

Professor Andrea Siebenhofer; [andrea.siebenhofer@medunigraz.at](mailto:andrea.siebenhofer@medunigraz.at)

## ABSTRACT

**Introduction:** Most studies conducted at general practices investigate complex interventions and increasingly use cluster-randomised controlled trial (c-RCT) designs to do so. Our primary objective is to evaluate how frequently complex interventions are shown to be more, equally or less effective than routine care in c-RCTs with a superior design. The secondary aim is to discover whether the quality of a c-RCT determines the likelihood of the complex intervention being effective.

**Methods and analysis:** All c-RCTs of any design that have a patient-relevant primary outcome and with a duration of at least 1 year will be included. The search will be performed in three electronic databases (MEDLINE, EMBASE and the Cochrane Database of Systematic Reviews (CDSR)). The screening process, data collection, quality assessment and statistical data analyses (if suitably similar and of adequate quality) will be performed in accordance with requirements of the Cochrane Handbook for Systematic Reviews of Interventions. A feasibility project was carried out that was restricted to a search in MEDLINE and the CCTR for c-RCTs published in 1 of the 8 journals that are most relevant to general practice. The process from trial selection to data collection, assessment and results presentation was piloted. Of the 512 abstracts identified during the feasibility search, 21 studies examined complex interventions in a general practice setting. Extrapolating the preliminary search to include all relevant c-RCTs in three databases, about 5000 abstracts and 150 primary studies are expected to be identified in the main study. 14 studies included in the feasibility project (67%) did not show a positive effect on a primary patient-relevant end point.

**Ethics and dissemination:** Ethical approval is not being sought for this review. Findings will be disseminated via peer-reviewed journals that frequently

## Strengths and limitations of this study

- Study selection, data extraction, and the assessment of risk of bias will be conducted by two authors independently.
- A comprehensive feasibility check was carried out for a systematic review and the results provided important information on how to design the study.
- It will be difficult to pool data because the target population is very variable and not limited to specific conditions or diseases. There is also likely to be considerable variation in patient-relevant primary outcomes.
- No search in trial registries to indicate possible publication bias is planned.

publish articles on the results of c-RCTs and through presentations at international conferences.

**Trial registration number:** PROSPERO CRD201400923.

## INTRODUCTION

Primary care delivered at general practices is critical in any healthcare system, and its importance is increasing due to the rising prevalence of chronic diseases and multimorbidity in an ageing population.<sup>1</sup> At the same time, the need to control healthcare costs makes it particularly important that healthcare professionals use interventions with proven effectiveness.

Most studies conducted at general practices focus on the behaviour of patients and health professionals, or on organisational

change<sup>2</sup> and concentrate on interventions such as disease management programmes, vaccination programmes, and screenings. As such interventions are generally complex, cluster-randomised controlled trials (c-RCTs) are increasingly used to evaluate them. However, c-RCTs have certain methodological shortcomings, a common example of which is inadequate concealment of the treatment allocation. As a result, the CONSORT statement was updated and extended to c-RCTs in 2004 and 2010, and now includes specific advice on how to meet various quality standards.<sup>3 4</sup> In addition, the 2013 Ottawa Statement describes the ethical issues that should be considered when conducting c-RCTs and provides guidance and key recommendations for researchers and ethics committees.<sup>5</sup>

The present manuscript describes the protocol of a methodological systematic review on the basis of a feasibility project. It has the primary objective of evaluating how frequently complex interventions are shown to be more, equally or less effective than routine care in c-RCTs that use a superior design. The secondary objective of the review is to discover whether the quality of a c-RCT determines the likelihood that the complex intervention will be proven to be effective.

## METHODS

### Criteria for inclusion in this review

#### Eligibility criteria

All c-RCTs involving adults, adolescents, and children in a general practice setting will be included. The trials must investigate a complex intervention in accordance with the recommendations of the latest Medical Research Council (MRC) guidance: we have included all interventions that involve interacting components in the experimental group—such as treatment, changes in behaviour required by those delivering and/or receiving interventions, and/or changes in the number of organisational levels targeted by the intervention.<sup>6</sup> To avoid additional heterogeneity between studies arising from active comparators, the control group must have continued to receive treatment as usual (routine care). For inclusion in our review, trialists must either have explicitly defined primary outcome(s) as primary or main outcome(s), have used such outcome(s) in a power and sample size calculation, or have listed it (these) as the main outcome(s) in their trial's objectives.<sup>7</sup> In addition, the primary outcome(s) has (have) to be patient relevant, and detailed criteria for the assessment of the patient-relevant end points should have been determined in accordance with the Institute for Quality and Efficiency in Health Care (IQWiG) methods V.4.2, which give a concise and literature-based definition of what is meant by patient relevance.<sup>8</sup> In this connection, 'patient relevant' refers to how a patient feels, functions, or survives—that is, whether indicators of mortality, morbidity, health-related quality of life, hospitalisation and/or treatment satisfaction are provided. If a study reports

on more than one primary end point, only the patient-relevant end points will be included in this study. As we want to evaluate the long-term benefit of an intervention, only studies of at least 12 months' duration will be considered. Inclusion and exclusion criteria are shown in [table 1](#).

### Outcome measures

1. Summarising the evidence from c-RCTs to describe the distribution of estimates of treatment effect with respect to direction (in favour of the complex intervention or the routine care treatments), magnitude (size of the effect), and statistical significance (or CI).
2. Evaluating how frequently complex interventions in c-RCTs are more, equally or less effective than routine care.
3. Exploring the extent to which methodological (eg, power calculations and intra-cluster correlation coefficients (ICC)) and other factors (eg, ethical approval, sponsorship, run-in phase) explain differences in the

**Table 1** Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
All c-RCTs of any design (eg, parallel, cross-over, or stepped wedge)	Main results were previously published; design papers; protocols or pilot studies and 'side papers' such as tertiary literature
Superiority trials	There is no reporting on patient-relevant primary outcomes Only practice (general practice level) data are available
c-RCT compares a complex intervention with routine care	No consideration of comparators, such as active controls or sham interventions, will be permitted due to homogeneity of samples
Studies have a patient-relevant primary outcome (no surrogates)	
Study duration >12 months	
Study included all age groups	
Studies examine individual patient data	
Funding is not relevant (commercial, non-commercial, other funding)	
General practice is the cluster	
No language restrictions	
Published studies only	
c-RCT, cluster-randomised controlled trial.	

distribution of c-RCTs that show results that either favour or disadvantage complex interventions.

### Search methods

The search strategy was developed by the Institute of General Practice at Goethe University Frankfurt, Germany, in cooperation with the Centre for Research in Evidence-Based Practice at Bond University, Australia, and was broadly based on a validated approach developed by Taljaard *et al.*<sup>9</sup> Relevant papers will be identified by searching the Central Register of Controlled Trials (CCTR) (last issue), MEDLINE (from 1962 until recently) and EMBASE (from 1988 until recently) without any language restriction. The full search strategy for MEDLINE and CCTR appears in table 2 and will be adapted for EMBASE. The proposed end date for the literature search is September 2015.

The literature search for the feasibility project was carried out in the databases 'Cochrane Central Register of Controlled Trials' (EBM Reviews—CCTR) and 'MEDLINE' between 1946 and April 2014. Following recommendations by Eldridge *et al.*,<sup>10</sup> the search strategy for the feasibility project was restricted to the journals publishing the highest numbers of articles related to general practice, namely the *British Medical Journal*, *British Journal of General Practice*, *Family Practice*, *Preventive Medicine*, *Annals of Internal Medicine*, *Journal of General Internal Medicine and Paediatrics*. The *Canadian Medical Journal (CMJ)* was also included because the initial inspection of a 10% sample of the unrestricted search results showed that this journal also contains a high number of c-RCTs in a general practice setting. For the

full review, the search will not be restricted to articles published in 'general practice'-related journals.

To ensure literature saturation, we will scan references in methodological and relevant secondary literature that were identified in the three electronic databases and reference lists of the included studies and that were published after January 2010. We will also search the authors' personal files—literature collected during the conceptual development of our project idea—to make sure that all relevant material has been captured.

### Expected primary studies

The initial search of the feasibility project identified 512 papers. Of these, 426 were excluded following abstract screening. The full texts of the remaining 86 studies were screened and 21 papers, or 4% of initial findings, ultimately fulfilled the inclusion criteria (see online supplementary file 1). When these results are extrapolated to take account of an unrestricted literature search that includes EMBASE as a third database as well as journals that do not focus on general practice, we expect to have about 5000 findings and 150 papers (3% of initial findings).

### Expected authors' responses

Eighteen authors were contacted for further information on the intra-cluster correlation coefficients (ICCs) assumed in the sample size calculation and observed in the data used in their studies; three authors had already provided the necessary information in their publications. If no response was forthcoming, reminders were sent 4 weeks after the initial contact. Of the 18 authors we contacted, 4 forwarded the relevant information; thus, ICCs were available for about 22% of studies (see online supplementary file 2).

### Selection of trials and data collection

The abstracts, titles and full texts will be independently screened by two reviewers. Data from each study will be assessed independently by the two authors and entered into data extraction templates. Disagreements will be resolved by a third reviewer and relevant missing information will be requested from the original authors of the study.

### Quality assessment

The criteria listed were developed during the feasibility phase and based on the CONSORT statement—extension to cluster randomised trials;<sup>4</sup> the extraction sheets for RCTs used by IQWiG;<sup>11</sup> the Cochrane Handbook for Systematic Reviews of Interventions;<sup>12</sup> and the systematic review of Froud *et al.*<sup>13</sup> The authors extracted all named criteria and then decided which to include in the assessment, based on their frequency and relevance to the research question. Additional criteria, such as number of participating practices and length of observation period for intervention and control groups, were included because these were considered to be important

**Table 2** Search strategy for MEDLINE and CCTR

**Databases: MEDLINE (Ovid) (1946 to April 2014) and EBM Reviews—Cochrane Central Register of Controlled Trials (Ovid) (January 2014)**

- 1 general practitioners/ or physicians, family/ or physicians, primary care/
- 2 Primary Health Care/
- 3 exp General Practice/
- 4 ((family or general or primary) adj3 (practic\* or practition\*)).tw.
- 5 primary care.tw.
- 6 (gp or gps).tw.
- 7 ((family or primary or general) adj3 (physician\* or doctor\* or clinician\*)).tw.
- 8 or/1–7
- 9 (cluster\* adj2 randomi\*).tw.
- 10 exp cluster analysis/
- 11 (practice\* adj5 random\*).tw.
- 12 or/9–11
- 13 exp animals/ not humans.sh.
- 14 12 not 13
- 15 8 and 14
- 16 remove duplicates from 15

quality measures. Finally, the reported criteria were grouped thematically (general information, sample size calculation, randomisation and blinding process, analyses) (see online supplementary file 3) and piloted using the studies identified during the feasibility phase. The results of the preliminary assessment are presented in online supplementary file 3 which includes tables representing the final template for the upcoming full review. As additional information, we will extract from the discussion section of the included studies the authors' own interpretations and explanations as to why their studies did not show a positive effect.

### Data analysis

Data will be summarised (or pooled) statistically where appropriate. We will perform the statistical analyses in accordance with the guidelines provided in the latest version of the Cochrane Handbook for Systematic Reviews of Interventions.<sup>12</sup> In addition, descriptive plots and analyses will be performed to explore the distribution of effect sizes, and the frequency of complex interventions in c-RCTs being more, equally or less effective than routine care. Data analysis will be performed in Cochrane Review Manager 5.1.0. We will use either HR or OR to estimate the individual and overall effects of studies that are presented with a 95% CI. We will also calculate the heterogeneity statistics ( $\chi^2$  and  $I^2$ ), and test the robustness of the results by repeating the analysis using different statistical models (fixed-effect and random effects model). When heterogeneity is found, we will attempt to determine the reasons for this by examining individual study and subgroup characteristics. Subgroup analyses are planned if sufficient RCTs can be identified, for example, on study fields, type of outcome, and type of practice. We will perform sensitivity analyses in order to explore the robustness of our results and visually inspect funnel plots for any indication of publication bias.

The intra-cluster correlation coefficient (ICC) gives a measure of the similarity of observations from the same class. It is usually defined as the proportion of variance accounted for by class variation and can be estimated by analysing variance methods.<sup>14</sup> To assess whether unrealistic assumptions regarding the ICC used in the sample size calculation may have resulted in trials failing to show the superiority of a complex intervention, we will compare the available ICC pairs used for the sample size calculation with the ICCs actually obtained from the data. Furthermore, descriptive statistics such as minimum, maximum and median absolute differences will be stated.

### ETHICS AND DISSEMINATION

Ethics approval will not be required, since this is a protocol for a systematic review utilising published data. Results will provide information on the shortcomings of c-RCTs and help in the design of studies with complex interventions. Once completed, the results from this

systematic review will be published in a peer-reviewed journal, and presented at international and national conferences.

### DISCUSSION

During the course of the development of the final protocol, the feasibility project was extraordinarily useful for judging whether the planned systematic review would be feasible in terms of the numbers of trials expected, the definition of the inclusion and exclusion criteria, the development of data extraction forms, and results presentations. It also helped summarise the quality criteria that need to be collected to deal with the question of which methodological and other factors explain differences in c-RCTs to show results that either favour or disadvantage complex interventions. Using a selective search in journals relevant to general practice, a total of 21 studies examining complex interventions in a general practice setting were identified during the feasibility project. Fourteen of these did not demonstrate a positive effect on a primary patient-relevant end point. This corresponds to 67% of all studies, which considering that a complex intervention study usually requires considerable effort and monetary resources—as well as a large number of patients—is a striking number.

In order to describe and analyse the differences between studies with and without an intervention effect, we developed our own checklist which includes 18 quality aspects based on previously used criteria in other methodological papers. For example, the CONSORT Statement requires that both estimated and observed ICCs are reported, and it is interesting to note that several papers providing lists of ICCs that are relevant to general practice have been published and recommended for use in study design.<sup>15 16</sup> However, in our feasibility project we only found three studies<sup>17–19</sup> that quoted both ICC values, indicating that journals and reviewers do not attach enough importance to this issue and should request this information more rigorously. When the search is expanded to include three databases and with no restriction to general practice-related journals, we expect this methodological review to provide an answer to the research question. A minor limitation for this study that should be noted is that no search in trial registries is planned, and the existence of publication bias can, therefore, not be ruled out. We selected complex interventions as the vast majority of interventions at general practices are multifaceted. The term 'complex intervention' is thus a descriptive element of the study. In a subsequent substudy—based on samples of the included studies—we will appraise the complex interventions themselves in accordance with the recommendations of Möhler *et al.*<sup>20 21</sup>

Diaz-Ordaz<sup>22</sup> published a review based on 73 c-RCTs conducted in residential facilities for older people. Less than 30% of the trials had accounted for clustering in their sample size calculations, and considerable

differences existed between studies with and without an intervention effect. Another recently published systematic review by Ivers *et al*<sup>23</sup>—that dealt with more than 300 cluster-randomised trials published between 2000 and 2008—clearly showed that despite the publication of the CONSORT statement on the reporting and methodological quality of c-RCTs in 2004,<sup>3</sup> very few aspects had been adhered to and further effort was necessary to improve methodological quality. A Cochrane Review, which was not restricted to c-RCTs and was published by Turner in 2013,<sup>24</sup> demonstrated that in RCTs, the factors named in the CONSORT statement were more often fully reported when journals had actively encouraged its use.

Several further reviews published before 2005 show that even though there has been some methodological improvement in terms of appropriate sample size calculations and analyses, weaknesses are still present, especially with regard to blinding and allocation status.<sup>2 10 25</sup> In addition to these methodological reviews, the publication of the extended CONSORT statement<sup>3</sup> and the Ottawa Statement<sup>5</sup> also help researchers to avoid pitfalls and design c-RCTs properly.

To the best of our knowledge, our research question—which aims to evaluate whether, and if so under what circumstances, it is sensible to assess the efficacy of a proposed intervention by conducting a c-RCT in a general practice setting—is novel. Our feasibility project has enabled us to obtain a valid estimate of the proportion of studies that are effective, and this provides the basis for a project that has been approved by the German Federal Ministry of Education and Research, and registered with Prospero (CRD42014009234). By performing a full literature search and exploring the extent to which methodological (eg, power calculation, intra-cluster correlation, handling of missing data<sup>25</sup>) and other factors (such as baseline risk and severity of diseases which influence the effect size,<sup>26</sup> ethical approval and sponsorship) explain differences in the reported effectiveness of c-RCTs, the main project aims to further underscore possible shortcomings, and provide further information and help in the design of studies of complex interventions.

#### Author affiliations

<sup>1</sup>Institute of General Practice, Goethe University, Frankfurt am Main, Germany

<sup>2</sup>Institute of General Practice and Evidence-based Health Services Research, Medical University of Graz, Graz, Austria

<sup>3</sup>AQUA-Institute for Applied Quality Improvement and Research in Health Care, Göttingen, Germany

<sup>4</sup>Institute for Medical Informatics, Statistics and Documentation, Medical University Graz, Graz, Austria

**Acknowledgements** Paul Glasziou, Sarah Thorning and Elaine Beller extensively discussed the project idea during the research fellowship of Andrea Siebenhofer at the Centre for Research in Evidence-Based Practice, Bond University, Australia in 2013.

**Contributors** AS was responsible for the conceptual design of the review. The manuscript was drafted by AS and SE, and was revised by CM. SE was responsible for the feasibility check, and AS critically revised the manuscript.

Additional statistical analysis for the feasibility check was conducted by AB and GP. The final version of this article has been reviewed and approved by all authors.

**Funding** The systematic review will be funded by the German Federal Ministry of Education and Research (FKZ: 01KG1504), Germany.

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

#### REFERENCES

1. Sachverständigenrates zur Begutachtung der Entwicklung im Gesundheitswesen. Gutachten 2009 des Sachverständigenrates zur Begutachtung der Entwicklung im Gesundheitswesen. Koordination und Integration—Gesundheitsversorgung in einer Gesellschaft des längeren Lebens. 2009;16/13770:1–520.
2. Eldridge SM, Ashby D, Feder GS, *et al*. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. *Clin Trials* 2004;1:80–90.
3. Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004;328:702–8.
4. Campbell MK, Piaggio G, Elbourne DR, *et al*. Consort 2010 statement: extension to cluster. *BMJ* 2012;345:e5661.
5. Taljaard M, Weijer C, Grimshaw JM. The Ottawa Statement on the ethical design and conduct of cluster randomised trials: précis for researchers and research ethics committees. *BMJ* 2013;346:f2838.
6. Craig P, Dieppe P, Macintyre S, *et al*. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:979–83.
7. Djulbegovic B, Kumar A, Glasziou PP, *et al*. New treatments compared to established treatments in randomized trials. *Cochrane Database Syst Rev* 2012;10:MR000024.
8. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). General methods: Version 4.1 (28.11.2013). Inst. für Qual. und Wirtschaftlichkeit im Gesundheitswes. 2013;1–202. [https://www.iqwig.de/download/IQWiG\\_General\\_Methods\\_Version\\_4-1.pdf](https://www.iqwig.de/download/IQWiG_General_Methods_Version_4-1.pdf)
9. Taljaard M, McGowan J, Grimshaw JM, *et al*. Electronic search strategies to identify reports of cluster randomized trials in MEDLINE: low precision will improve with adherence to reporting standards. *BMC Med Res Methodol* 2010;10:15.
10. Eldridge S, Ashby D, Bennett C, *et al*. Internal and external validity of cluster randomised trials: systematic review of recent trials. *BMJ* 2008;336:876–80.
11. Institute for Quality and Efficiency in Health Care. Institute for Quality and Efficiency in Health Care. <https://www.iqwig.de/>
12. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. [Last updated March 2011]. 2011. <http://handbook.cochrane.org/>
13. Froud R, Eldridge S, Diaz Ordaz K, *et al*. Quality of cluster randomized controlled trials in oral health: a systematic review of reports published between 2005 and 2009. *Community Dent Oral Epidemiol* 2012;40(Suppl 1):3–14.
14. Donner A, Klar N. *Design and analysis of cluster randomization trials in health research*. London: Arnold, 2000.
15. Smeeth L, Siu-Woon Ng E. Intraclass correlation coefficients for cluster randomized trials in primary care. *Control Clin Trials* 2002;23:409–21.
16. Adams G, Gulliford MC, Ukoumunne OC, *et al*. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiol* 2004;57:785–94.
17. Steventon A, Bardsley M, Billings J. Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. *BMJ* 2012;344:e3874.
18. Griffiths C, Foster G, Barnes N, *et al*. Specialist nurse intervention to reduce unscheduled asthma care in a deprived multiethnic area: the East London Randomised Controlled Trial for High Risk Asthma (ELECTRA). *BMJ* 2004;328:144.
19. Murphy AW, Cupples ME, Smith SM, *et al*. Effect of tailored practice and patient care plans on. *BMJ* 2009;339:b4220.



20. Möhler R, Köpke S, Meyer G. Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare: revised guideline (CReDECI 2). *Trials* 2015;16:1–9.
21. Möhler R, Bartoszek G, Köpke S MG. Proposed Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare (CReDECI): guideline development. *Int J Nurs Stud* 2012;49:40–6.
22. Diaz-Ordaz K, Froud R, Sheehan B, *et al*. A systematic review of cluster randomised trials in residential facilities for older people suggests how to improve quality. *BMC Med Res Methodol* 2013;13:127.
23. Ivers NM, Taljaard M, Dixon S, *et al*. Impact of CONSORT extension for cluster randomised trials on quality of reporting and study methodology: review of random sample of 300 trials, 2000-8. *BMJ* 2011;342:d5886.
24. Turner L, Shamseer L, Altman DG, *et al*. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev* 2013;11:MR000030.
25. Díaz-Ordaz K, Kenward MG, Cohen A, *et al*. Are missing data adequately handled in cluster randomised trials? A systematic review and guidelines. *Clin Trials* 2014;11:590–600.
26. Muth C, van den Akker M, Blom JW, *et al*. The Ariadne principles: how to handle multimorbidity in primary care consultations. *BMC Med* 2014;12:223.