Cluster-randomized controlled trials evaluating complex interventions in general practices are mostly ineffective: a systematic review

Andrea Siebenhofer[a,b,*], Michael A. Paulitsch[a], Gudrun Pregartner[c], Andrea Berghold[c], Klaus Jeitler[b,c], Christiane Muth[a], Jennifer Engler[a]

[a]Institute of General Practice, Goethe University Frankfurt am Main, Theodor-Stern-Kai 7, Frankfurt am Main 60590, Germany
[b]Institute of General Practice and Evidence-based Health Services Research, Medical University of Graz, Auenbruggerplatz 2/9/IV, Graz 8036, Austria
[c]Institute for Medical Informatics, Statistics and Documentation, Medical University Graz, Auenbruggerplatz 2, Graz 8036, Austria

Accepted 17 October 2017; Published online 31 October 2017

Abstract

Objectives: The aim of this study was to evaluate how frequently complex interventions are shown to be superior to routine care in general practice-based cluster-randomized controlled studies (c-RCTs) and to explore whether potential differences explain results that come out in favor of a complex intervention.

Study Design and Setting: We performed an unrestricted search in the Central Register of Controlled Trials, MEDLINE, and EMBASE. Included were all c-RCTs that included a patient-relevant primary outcome in a general practice setting with at least 1-year follow-up. We extracted effect sizes, P-values, intracluster correlation coefficients (ICCs), and 22 quality aspects.

Results: We identified 29 trials with 99 patient-relevant primary outcomes. After adjustment for multiple testing on a trial level, four outcomes (4%) in four studies (14%) remained statistically significant. Of the 11 studies that reported ICCs, in 8, the ICC was equal to or smaller than the assumed ICC. In 16 of the 17 studies with available sample size calculation, effect sizes were smaller than anticipated.

Conclusion: More than 85% of the c-RCTs failed to demonstrate a beneficial effect on a predefined primary endpoint. All but one study were overly optimistic with regard to the expected treatment effect. This highlights the importance of weighing up the potential merit of new treatments and planning prospectively, when designing clinical studies in a general practice setting. © 2017 Elsevier Inc. All rights reserved.

Keywords: Cluster-randomized controlled trial; General practice; Effectiveness; Complex intervention; Shortcomings; Systematic review

1. Introduction

Cluster-randomized controlled trials (c-RCTs) are considered to be a suitable study design for examining patient-relevant clinical questions in primary care. A c-RCT is characterized by the random group assignment of people from, for example, communities, families, or medical practices [1,2]. However, methodological shortcomings are common [3]. Publications such as the extended Consolidated Standards of Reporting Trials (CONSORT) [1,2] and the Ottawa statements [4] were made to overcome this problem.

As most interventions in general practice are multifaceted and often have partly interacting components, they are considered to be complex, meaning that the study team has to consider many different aspects when designing the study [5,6]. Furthermore, study authors can be expected to be more or less convinced of the superiority of the new intervention when planning such an elaborate enterprise. However, in our feasibility project, which was restricted to journals relevant to general practice, only 33% of included studies showed statistically significant effects on a patient-relevant primary endpoint [7]. This was even less than in a recently published systematic review on

What is new?

Key findings
- A very low number of studies evaluating a complex intervention in a general practice setting showed superiority compared to routine care.

- Underestimates of intracluster similarities were unlikely to have been the principal reason behind the large number of negative findings.

- Anticipated effect sizes were mostly higher than those actually obtained in studies examining complex interventions vs. routine care.

What this adds to what was known?
- In a recently published systematic review of randomized controlled trials, only 50–60% demonstrated the superiority of new interventions when compared with standard treatment. In our review, more than 85% of cluster-randomized controlled studies (c-RCTs) failed to even demonstrate a beneficial effect on a predefined primary endpoint.

What is the implication and what should change now?
- Our findings highlight the importance of weighing up the potential merit of new treatments, considering the appropriateness of c-RCTs, or any other study design, and planning prospectively, when designing clinical studies in a general practice setting.

randomized controlled trials (RCTs) which reported that 50–60% had demonstrated the superiority of new interventions compared to standard treatment [8].

At the Institute of General Practice at Goethe University Frankfurt, two recent c-RCTs [9,10] involving complex interventions were unable to demonstrate superiority, even though great effort was put into designing these trials. As a result, the current methodological project was initiated. The primary objective of our systematic and methodological review was to evaluate how frequently and to what extent complex interventions are shown to be superior to routine care in general practice-based c-RCTs. A further aim was to explore whether potential differences in methodological and other factors could explain results that come out in favor of a complex intervention.

2. Methods

The protocol and the results of the feasibility project for this systematic and methodological review were recently

published in BMJ Open [7]. The protocol was registered at PROSPERO [11].

2.1. Eligibility criteria

We conducted a systematic and methodological review of all published and available c-RCTs, independent of patient age, and with the general practice setting as the level of randomization. In a superiority trial, the intervention group had to have investigated a complex intervention in accordance with the recommendations of the latest Medical Research Council guidance [12], and the control group had to have received routine care. To prevent additional heterogeneity between studies arising from active comparators, the control group had to have continued to receive treatment as usual (routine care). For inclusion in our review, trialists either had to have explicitly defined the patient-relevant primary outcome(s) that they used as primary or main outcome(s) in a power and sample size calculation or to have listed it (these) as the main outcome(s) in their trial's objectives [13]. The primary outcome(s) had to be patient relevant, and detailed criteria for the assessment of the patient-relevant endpoints had to have been determined in accordance with the Institute for Quality and Efficiency in Health Care (IQWiG) methods version 4.2, which gives a concise and literature-based definition of what is meant by patient relevance [14]. In this connection, "patient relevant" is considered to refer to how a patient feels, functions, or survives, that is, whether indicators of mortality, morbidity, health-related quality of life, hospitalization, and/or treatment satisfaction are provided. Furthermore, data had to be provided on a patient level, studies had to be of at least 12 months' duration, and patient-relevant primary outcome(s) had to have been measured after a period of at least 12 months had elapsed. Language was not a criterion for exclusion. In addition to those used in the feasibility project, two further criteria were added: psychotherapeutic interventions (e.g., cognitive behavioral therapy) were excluded because they are not typically performed by general practitioners, as were interventions that did not take place in a general practice setting and were not conducted by one or several members of a general practice team.

For further details of the eligibility criteria, see at PROSPERO [11] and the published protocol in BMJ Open 2016 [7].

2.2. Search strategy

As described in the protocol publication [7], we followed the validated strategy recommended by Taljaard et al. [15] and Bland [16], and performed an unrestricted search in the Central Register of Controlled Trials (CCTR, August 2015), MEDLINE (from 1946), and EMBASE (from 1988) until September 14, 2015. A combination of subject headings and text words relating to "general