CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector

Osteoarthritis and Cartilage 23 (2015) 826–838

Osteoarthritis and Cartilage

Review

OARSI Clinical Trials Recommendations: Design and conduct of implementation trials of interventions for osteoarthritis

K.D. Allen † ‡ § *, S.M.A. Bierma-Zeinstra ‖ ¶, N.E. Foster #, Y.M. Golightly † ††, G. Hawker ‡‡ §§

† *Thurston Arthritis Research Center, University of North Carolina, Chapel Hill, NC, USA*
‡ *Department of Medicine, University of North Carolina, Chapel Hill, NC, USA*
§ *Health Services Research and Development, Department of Veterans Affairs Medical Center, Durham, NC, USA*
‖ *Department of General Practice, Erasmus MC – University Medical Center Rotterdam, The Netherlands*
¶ *Department of Orthopaedic Surgery, Erasmus MC – University Medical Center Rotterdam, The Netherlands*
# *Arthritis Research UK Primary Care Centre, Institute of Primary Care and Health Sciences, Keele University, Keele, UK*
†† *Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA*
‡‡ *Department of Medicine, University of Toronto, Canada*
§§ *Women's College Research Institute, Women's College Hospital, University of Toronto, Canada*

ARTICLE INFO

Article history:
Received 8 December 2014
Received in revised form
24 February 2015
Accepted 26 February 2015

Keywords:
Osteoarthritis
Implementation
Clinical trials

SUMMARY

Rigorous implementation research is important for testing strategies to improve the delivery of effective osteoarthritis (OA) interventions. The objective of this manuscript is to describe principles of implementation research, including conceptual frameworks, study designs and methodology, with specific recommendations for randomized clinical trials of OA treatment and management.

This manuscript includes a comprehensive review of prior research and recommendations for implementation trials. The review of literature included identification of seminal articles on implementation research methods, as well as examples of previous exemplar studies using these methods. In addition to a comprehensive summary of this literature, this manuscript provides key recommendations for OA implementation trials.

This review concluded that to date there have been relatively few implementation trials of OA interventions, but this is an emerging area of research. Future OA clinical trials should routinely consider incorporation of implementation aims to enhance translation of findings.

Published by Elsevier Ltd on behalf of Osteoarthritis Research Society International.

Practice gaps in osteoarthritis management

There are a number of evidence-based treatments for osteoarthritis (OA), including behavioral, rehabilitation, pharmacological, and surgical strategies[1–3]. However, many studies show there are gaps in the utilization or uptake of some treatment approaches in real-world clinical practice[4–12]. For example, although exercise is a core component of managing OA, in a recent survey of U.S. physicians, less than one third said they would provide exercise advice for patients with knee OA[13]. Similarly, studies have shown that recommendations for weight management for patients with OA are infrequent in clinical practice[4,9,12]. There is also evidence that physiotherapy is underutilized[5]. For instance, among patients with knee OA in two general practice regions in the UK, only 13% had ever received physiotherapy[14]. With respect to pharmacotherapy, studies have found inadequate attention to safety-related issues, particularly with respect to use of nonsteroidal anti-inflammatory drugs (NSAIDs) among older adults[15]. With regard to surgical interventions, studies document inappropriate use of arthroscopy[16–18] and inequities in joint replacement surgery[19–23]. Overall, studies have shown low "pass rates" for meeting quality indicators for OA-related clinical care[9,15]. In one study of quality of care for community-dwelling adults with OA, quality indicator pass rates ranged from 44 to 73% for recommended non-pharmacological approaches and 27–59% for pharmacological

* Address correspondence and reprint requests to: K.D. Allen, 3300 Thurston Bldg., CB# 7280, Chapel Hill, NC 27599-7280, USA. Tel: 1-919-966-0558.
E-mail addresses: kdallen@email.unc.edu (K.D. Allen), s.bierma-zeinstra@erasmusmc.nl (S.M.A. Bierma-Zeinstra), n.foster@keele.ac.uk (N.E. Foster), golight@email.unc.edu (Y.M. Golightly), g.hawker@utoronto.ca (G. Hawker).

http://dx.doi.org/10.1016/j.joca.2015.02.772
1063-4584/Published by Elsevier Ltd on behalf of Osteoarthritis Research Society International.

approaches[15]. A recent review also found suboptimal management of OA across quality of care domains including effectiveness, safety, access, and support for self-management[10].

There are many factors that may contribute to these gaps in translation of evidence-based OA care, including:

- Provider unfamiliarity with OA treatment recommendations. For instance, one study showed that 2/3 of U.S. primary care physicians were unfamiliar with OA treatment guidelines from the Osteoarthritis Research Society International (OARSI)[11].
- Lack of practical guidance within OA treatment guidelines, such as indicators for when specific treatments may be appropriate for subgroups of patients[24].
- Lack of decision support tools and information technology infrastructure to foster guideline-concordant OA care.
- Lack of incentives or quality measures for OA within health care systems.
- Perception that OA is often seen as a degenerative condition for which current interventions offer little help, therefore relegating OA to a low priority in terms of clinical treatment.
- Provider time constraints, which may particularly impact counseling for behavioral strategies such as weight management and exercise.
- Health care access issues, such as lack of insurance coverage (or high co-payments) for services such as physiotherapy.
- Limitations in many prior studies with respect to external validity and practical application in clinical settings, including:
  - Exclusion of many patients with medical and psychological comorbidities common in patients with OA.
  - Lack of trials that directly compare different interventions, particularly those that compare new interventions to established ones.
  - Few studies conducted in primary care settings, where much of OA care occurs.
  - Lack of studies examining organizational barriers and facilitators to implementing interventions.

Implementation research, defined as research "focused on the adoption or uptake of clinical interventions by providers and/or systems of care[25]," can test strategies to integrate evidence-based OA interventions and improve practice patterns within clinical settings. This manuscript provides an overview of conceptual frameworks, study designs and methodology of implementation research, with specific recommendations for randomized clinical trials (RCTs) that test the implementation of evidence-based OA care. Other implementation study types and intervention taxonomies have been described elsewhere[26].

Methods

This review began with a search of Pubmed using terms of "osteoarthritis" and "implementation." We also conducted general searches for manuscripts covering general RCT implementation methods, covering each of the sub-topics below; from this literature we identified key study frameworks, designs and methods, as well as other manuscripts of high relevance. Authors communicated via a series of teleconferences and email correspondence to identify additional topics and manuscripts for inclusion and to develop and reach concurrence on a set of recommended principles for OA implementation trials. Authors of this review included a health services researcher and exercise physiologist, an academic physiotherapist, an epidemiologist and physical therapist, a rheumatologist, and a clinical researcher in general practice and orthopedics; all had expertise in OA, including studies of care delivery (including RCTs).

Conceptual frameworks to guide implementation trials

For all types of RCTs, conceptual frameworks can help to guide both development and evaluation of intervention strategies. The following conceptual frameworks are examples that represent different paradigms and have been applied to different interventions and implementation studies:

RE-AIM framework

The RE-AIM framework, which stands for Reach, Effectiveness, Adoption, Implementation, and Maintenance, can be used to assess overall effectiveness of an intervention by emphasizing the representativeness of participants and of settings[27,28]. RE-AIM provides a systematic context for the assessment of the impact of health behavior interventions implemented at individual and organizational levels[29–31]. The dimensions of "Reach" and "Effectiveness" focus on the individual level, the "Adoption" and "Implementation" factors examine the organizational level, and the "Maintenance" factor comprises both levels. "Reach" assesses the proportion of the people who receive or are affected by an intervention, as well as the characteristics of the participants to determine representativeness. "Effectiveness" examines positive and negative outcomes of an intervention to ensure that the benefits outweigh the harms when applied to a large population "Adoption" indicates the proportion and types of settings that initiate an intervention, as well as barriers to adoption. "Implementation" refers to how well the intervention is applied to a population as originally planned and is thought to interact with efficacy to determine overall effectiveness. Lastly, "Maintenance" refers to the long-term use or application of an intervention at the individual and organizational levels. In the context of implementation trials, the RE-AIM framework can be used both in the development and evaluation stages. In the trial development phase, the five RE-AIM domains can be used to plan an intervention with high likelihood for public health impact. In the evaluation stage, assessment of all five RE-AIM dimensions can provide comprehensive understanding of long-term implementation potential and areas of weakness. The Practical, Robust Implementation and Sustainability Model (PRISM) is a comprehensive tool that can help researchers and policy makers evaluate how interventions interact with recipients to influence elements in the RE-AIM framework, as well as other models[32].

Knowledge-to-Action (KTA) framework

The KTA framework was developed by Graham and colleagues to facilitate the use of research knowledge by stakeholders including clinicians, policymakers, patients and the general public[33–37]. The KTA process has two main components (Fig. 1): (1) knowledge creation (inner cycle) and (2) action (outer cycle). Both components have multiple phases, and the KTA process is typically complex and iterative; knowledge and action phases can occur sequentially or simultaneously and interact with each other. A key characteristic of the knowledge creation "funnel" (Fig. 1) is that knowledge becomes more distilled, refined, and useful to stakeholders during the process; this is a result of incorporating stakeholder needs in each phase of the knowledge creation process. The action cycle focuses on "deliberately engineering change in health care systems and groups[36]." Notably, stakeholder involvement is also critical for action cycle phases (e.g., adaption and tailoring of knowledge and interventions to the local context; Fig. 1). The action cycle also includes monitoring of knowledge use to determine the effectiveness of strategies and modify them accordingly, as well as planning for sustained knowledge use.

The KTA process has been applied to a number of different health conditions and settings[35,38–42]. For example, Tugwell and

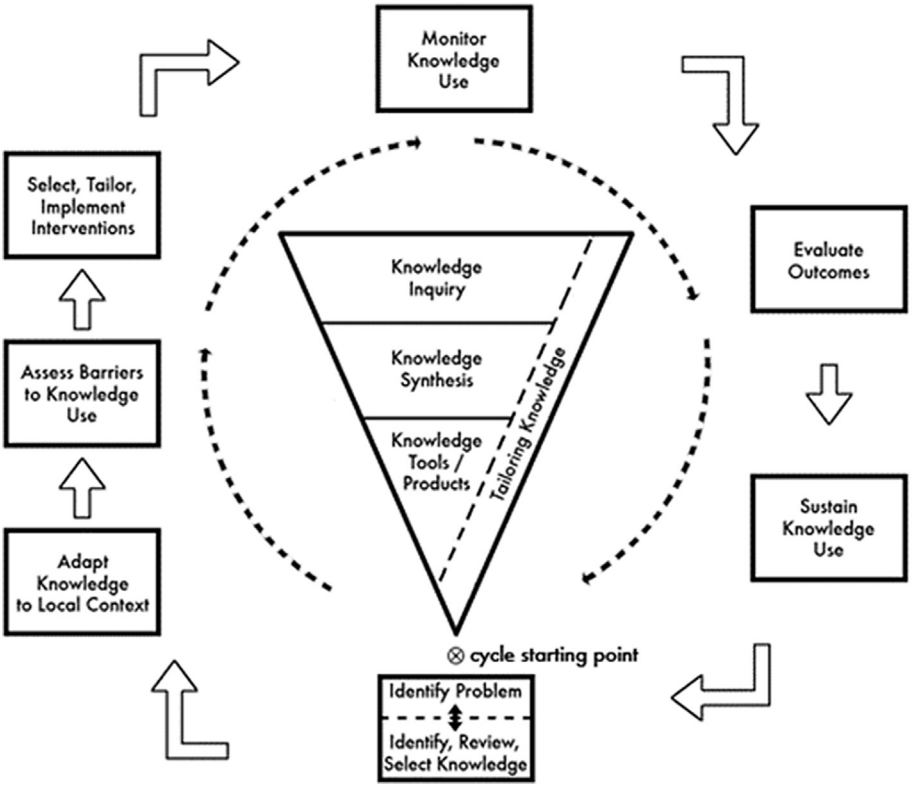

Fig. 1. The KTA Process. From Graham, I. D. *et al.* (2006) *Journal of Continuing Education in the Health Professions*, Vol. 26, No. 1, 13–24.

colleagues provided a detailed description of applying the KTA process to the management of musculoskeletal conditions, beginning with knowledge from systematic reviews and working with consumers through each knowledge and action phase to develop tailored educational products[38]. A current RCT is using the KTA framework to examine the effect of implementing an online program for patients with rheumatoid arthritis[43]. The KTA framework can be applied in the context of different RCT designs for implementation studies (described below), with components of the knowledge creation and action cycles helping to determine the most appropriate methods to answer key questions for consumers.

Theoretical Domains Framework

The Theoretical Domains Framework focuses on psychosocial and organizational theory related to behaviors and behavior change in clinical practice[44]. The designers of this framework concentrated on the clinical behavior of the individual health professional because health professional–patient encounters are frequent, essential components of health care that are important to the quality of care and the success of the health of the patient[44,45]. This framework includes 12 domains for health professional behavior and behavior change: knowledge; skills; social/professional role and identity; beliefs about capabilities; beliefs about consequences; motivation and goals; memory, attention, and decision processes; environmental context and resources; social influences; emotion; behavioral regulation; and nature of the behaviors[45]. French *et al.*[45] describe a four-step method for developing an implementation intervention: (1) identifying target behaviors (the behavior change needed to reduce the evidence-practice gap), (2) using the Theoretical Domains Framework to identify barriers and enablers to change, (3) identifying techniques to change behavior in a feasible and acceptable manner, and (4) conducting an implementation intervention. The Theoretical Domains Framework is particularly relevant for implementation trials of interventions that focus primarily on provider behaviors and patterns related to OA care.

Normalization Process Theory

The Normalization Process Theory provides a framework based on sociological theory to bridge research, policy, and practice. This framework was originally intended to help researchers and clinicians determine and understand factors that encourage or hinder implementation of an intervention[46]. It has expanded to address the participant's understanding, engagement, and support of the intervention, along with the individual's evaluation of the intervention, in an effort to understand social actions (what people do)[46,47]. The Normalization Process Theory consists of four constructs: Coherence, Cognitive Participation, Collective Action, and Reflexive Monitoring[46]. Coherence is the process of sense-making that individuals and organizations experience when adding or restricting a new practice or set of practices. Cognitive Participation is the manner in which individuals and organizations encourage participants to engage in a new practice. Collective Action is the effort of individuals, teams and organizations together to implement a new practice. The last construct, Reflexive Monitoring, is the assessment of the positive and negative outcomes of the new practice after its initiation and during its ongoing use. Normalization Process Theory can be used in the evaluation stage of implementation trials, particularly for complex approaches that consist of both treatment and organizational interventions, to help with understanding factors that affect implementation processes.

Study designs for implementation trials

There are a number of study types and designs that can be applied to implementation trials. Here we describe some of the most common designs, including their general principles, some

unique methodological features and challenges, and when available, examples of OA trials. Fig. 2 shows a summary of these designs.

Hybrid effectiveness-implementation trials

Curran *et al.* have introduced a typology of trial designs that blends components of clinical effectiveness and implementation studies[25]. The aim of this blended approach is to speed the uptake of research findings into real-world clinical practice. Glasgow and others have suggested that the typical time lag between discovery and clinical uptake is partly due to the predominant research pathway that begins with efficacy studies, then effectiveness trials, and finally implementation research[28,48]. Alternatively, these stages of research can be blended in ways that result in more rapid translation while preserving rigorous methodology[25,28]. Here we focus on approaches for blending effectiveness and implementation research, as described in detail by Curran *et al.*[25], which includes three broad types of designs. *Hybrid Type 1* designs test a clinical intervention (e.g., effectiveness) while also gathering information on its delivery (e.g., implementation) and/or its potential for implementation in a real-world situation. Although the focus of Hybrid Type 1 studies is still on the effectiveness of a clinical intervention, they can simultaneously answer many questions important for transitioning to implementation, such as organization level barriers and facilitators. For example, two studies are examining patient and provider interventions for managing OA in primary care in different health care settings; primary hypotheses are related to the effectiveness of the interventions, but implementation-related analyses will evaluate facility/provider-level variation and clinician feedback on the feasibility and potential for integration into routine OA care[49]. *Hybrid Type 2* designs involve simultaneous testing of a clinical intervention and an implementation strategy. By directly blending effectiveness and implementation aims, these study designs can lead to more rapid provision of results that inform intervention delivery processes. The Management of Osteoarthritis in Consultations Study (http://www.controlled-trials.com/ISRCTN06984617) is an example of a Hybrid Type 2 design; this cluster RCT is testing the feasibility, acceptability and impact of implementing a new approach to supporting self-management for OA in primary care in the UK. Study aims include both evaluation of intervention effectiveness and collection of process data and qualitative interview data regarding barriers, facilitators, and other aspects of intervention delivery. *Hybrid Type 3* designs test an implementation intervention while observing/gathering information on the clinical intervention and related outcomes. Here the focus is primarily on implementation of an intervention with established effectiveness, yet it may still be important to evaluate the effects of the intervention when delivered in routine clinical practice, in specific settings or differing conditions. Selection of an appropriate hybrid effectiveness-implementation design depends on many factors, particularly the current level of evidence for an intervention. Hybrid trials can incorporate the specific study designs described below.

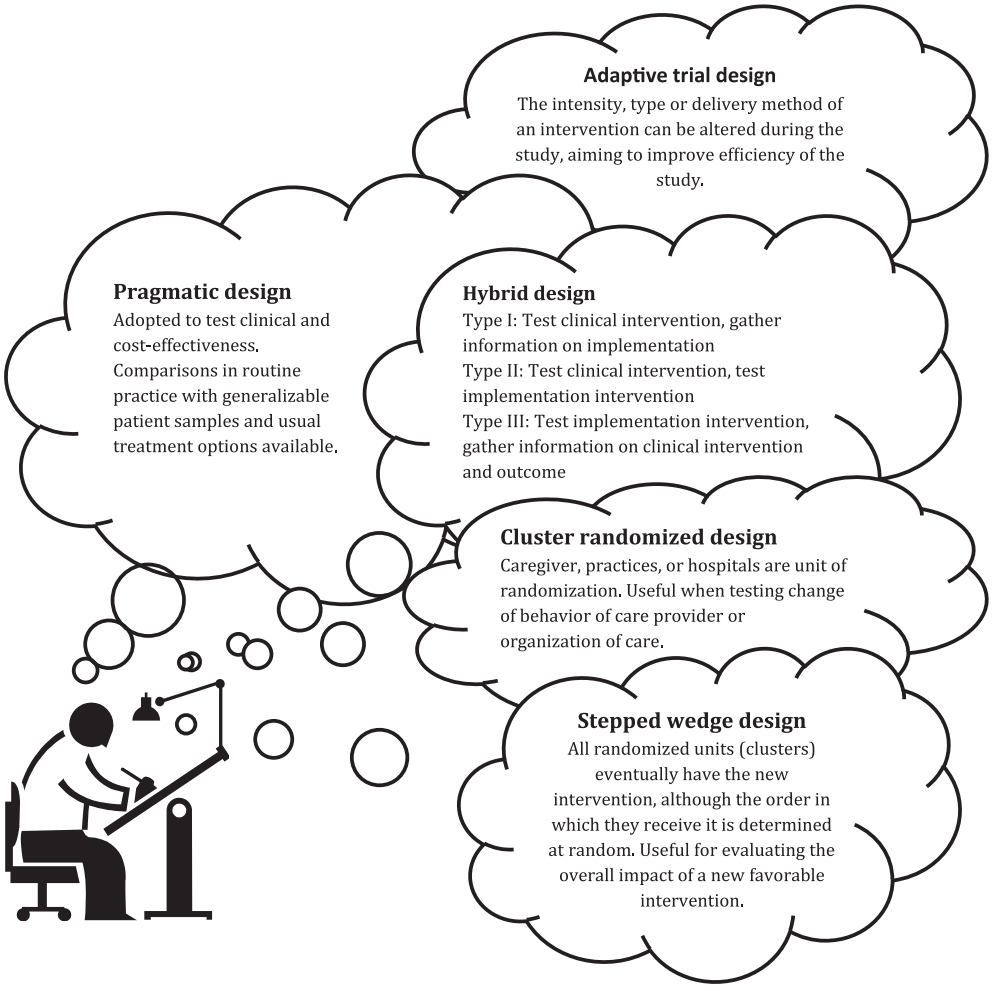

Fig. 2. Summary of implementation trial designs.

Pragmatic trials

A distinction in trial design can be made between explanatory (or efficacy) and pragmatic (or effectiveness) trials[50]. Explanatory trials test causal research hypotheses to determine whether an intervention works in highly controlled (or ideal) conditions, usually with highly selected participants[51]. The comparison intervention in an explanatory trial is often either a 'no intervention' control or a placebo control. The results of explanatory trials may therefore be of questionable generalizability to routine practice. Pragmatic trials compare the effectiveness of interventions in everyday practice with relatively unselected participants and under flexible conditions[50,51]. Therefore they help choose between options for care under the usual conditions in which those options might be offered. The key feature of pragmatic trials is high external validity or generalizability, addressing the question 'does this intervention actually work in real life' despite the complexity of real world clinical services. For this reason pragmatic trials are extremely helpful in informing decisions about routine practice[51]. Evidence from pragmatic trials is most useful to policy makers in deciding on whether to implement and allocate resources to new interventions and services for patients with OA. Originally adopted by Schwartz and Lellouch in 1967[52], the terms explanatory and pragmatic are helpful to distinguish different types of trials, and although the terms suggest a clean dichotomy (a trial is *either* explanatory or pragmatic), in reality, there is a continuous spectrum[50] with many variations in elements including the breadth of eligibility criteria, the flexibility in intervention delivery, expertise of those delivering treatment, degree of standardization of intervention protocol, effort to ensure intervention compliance, and specific approaches to data analyses. In addition, although pragmatic trials are often labeled effectiveness trials, they typically incorporate many elements of implementation research, including evaluation of feasibility and barriers in real-world practice (e.g., Hybrid Type 1 or 2 categorization as described above). The pragmatic-exploratory continuum indicator summary (PRECIS) tool can help researchers assess the degree to which trial design decisions align with its objectives (e.g., pragmatic/decision-making vs explanatory)[50].

Pragmatic trials are being increasingly adopted to test the clinical and cost-effectiveness of implementation interventions in the field of OA, since even well-studied interventions, with high-quality evidence from robustly designed and performed explanatory trials, will probably be less effective when tested in broader populations and clinical services. Many pragmatic trials of interventions for OA are complex, involving a number of separate but interacting components that are likely to be important to the success of the intervention, although the 'active ingredients' are often difficult to specify. The Medical Research Council Framework for the *development* and *evaluation* of complex interventions provides helpful guidance for these types of studies[53,54]. In implementing a novel, complex intervention (e.g., a new educational service comprising nurse-led clinics that focus on exercises and dietary advice for people with OA of the knee), a rigorous process of *development* is needed to inform both the trial design and intervention. What seems to be a sensible plan can easily lead to a superficial, ineffective intervention that is not adopted or sustained in clinical settings; this emphasizes the importance of a strong knowledge creation phase that closely involves key stakeholders, as emphasized by the KTA framework. Ensuring components work well together is vital as it is easy for components to either dilute the effect of each other or work against each other. Pragmatic trials work best when they have large samples (to increase power to detect small differences between the interventions being compared) and simple designs (such as the simple, two-arm parallel design) as this makes the trial easier to plan, conduct and analyze. *Evaluation* of complex interventions is also challenging[53,54], and a key question is to what extent trial investigators attempt to evaluate how a complex intervention works, i.e., to 'tease out' the effective ingredients in the 'black box'. Pragmatic trials investigate the overall performance of the intervention (with all the components working together) and cannot identify the specific components that directly explain the effects observed, unless trial teams make specific additional efforts to do this. For example, nurse-led clinics focused on exercise and diet for OA might potentially work by some or all of the following: improved confidence and competence of the nursing staff in supporting patients with OA to make behavioral changes, increased access for patients to health professional support, decreased fear of physical activity in patients, improved self-efficacy and control over symptoms, increased social interaction and participation, health professionals' time, empathy, attention, and so on. Pragmatic trial teams need to decide whether to measure these factors in order to model an explanatory element, bearing in mind that the greater the measurement burden the more likely selection and attrition bias will operate.

There are some key methodological challenges to consider with pragmatic trials. First, bias or the systematic distortion of the estimates due to poor design, conduct or analysis of a trial[55], is particularly an issue for complex interventions where 'real world' estimates are paramount. Maximizing generalizability and access to the key target group is important, requiring careful attention to how the trial is communicated to potentially eligible patients and using recruitment methods that avoid the recruitment of highly selected participants, such as those that rely only on busy clinicians. Some potential recruitment methods include using consultation-code activated electronic tags or mailed invitation or population screening surveys of registered patients. Second, there are challenges concerning blinding of participants and research personnel. Blinding, or the masking of patients, practitioners, outcome assessors and statisticians about the treatment to which an individual patient has been allocated, is the traditional approach to try to prevent performance and ascertainment bias in trials[55,56]. However, complex implementation interventions, by their nature, may mean that the usual types of blinding (patient and practitioner) are typically not possible. In fully pragmatic trials, in which the specific 'active' ingredient of the intervention is not of interest, placebo interventions are rarely included, yet some blinding is often still possible[50] and feasible[57] particularly of outcome assessment and during analysis.

An example of a pragmatic RCT is that of Gooch *et al.*[58], who tested the effectiveness of a new, evidence-based, clinical pathway, compared to the usual care pathway, for patients undergoing primary hip and knee joint replacement. The new clinical pathway was a clear example of a complex intervention, as it featured central intake clinics, dedicated inpatient resources, care guidelines and efficiency benchmarks. They used a simple design (two-arm parallel RCT), a large sample size (1570 patients) and had a clear primary outcome (WOMAC overall score 12 months following surgery). Results showed that the evidence-informed clinical pathway led to small but significantly superior outcomes.

In summary, pragmatic randomized trials are an increasingly popular design to test implementation interventions. In designing pragmatic trials of complex interventions aimed at improving the evidence-based treatment of OA, there are few 'right' answers. The key is to be clear at the outset of the trial, and to conduct and report the trial robustly[59] to facilitate appropriate interpretation. Clear thinking about the key question and careful decision-making will maximize the chance of the results influencing clinical practice for the better.

Cluster randomized trials

Cluster randomized trials are not unusual in pragmatic intervention research for OA[60–72]. A cluster-randomized trial is one in

which a group is randomized, and not the individual. For example general practices, physiotherapy practices, hospitals, workplaces, or families are randomized. Although this trial design can be used across phases of research, it has particular utility for effectiveness and implementation studies.

Reasons for adopting cluster-randomized trials vary. A scientifically obvious reason is that there might be a *cluster action of the intervention*; an example of this is the treatment of transmittable infections in which the treatment itself also reduces the transmission of the disease[73]. This is not applicable in OA research, but when interventions are given to a cluster (general practice, physiotherapy practice, *etc*.) it might enhance participant motivation and adherence by discussions between participants[74]. Another obvious reason is that some *interventions are targeted at the cluster level* and are aimed at the professional or (multidisciplinary) teams; their care of patients can only be evaluated at cluster level[75]. A frequently communicated reason for adapting cluster-randomized designs is the *avoidance of treatment group contamination*. For example, when an intervention transmits knowledge contamination towards the control or alternative intervention group, this should be avoided by separation of the groups in space or time[75]. Overall, cluster-randomized trials are particularly appropriate for evaluation of interventions that aim to change behavior of either the professional or the patients or aim to change the organization of providing treatments or referrals[76].

There are several issues that require specific attention in cluster-randomized trials. The first of them concerns *informed consent procedures*. This procedure in cluster-randomized trials might differ from the usual procedures when individuals are randomized. Recently, The Ottawa Statement on the ethical design and conduct of cluster randomized controlled trials gave guidance for the informed consent procedure in a range of particular situations[77]. In trials where the intervention is aimed at the professional and contains, for example, a novel management program for OA, one might only seek consent to data collection but not to the intervention or randomization from the patient[78]. In such cases the clinicians or practice leads provide consent to randomization on behalf of their patients. The practice in which the patient participates is already randomized and all patients undergo the management according to randomization. However, information can be provided even if the patient does not have a feasible alternative intervention and cannot withdraw[75]. The value of courteously providing information to patients should not be underestimated, although it is possible that this results in either an increase in goodwill, or an increase in patient concern. Failure to provide this information risks psychological harm, as subsequent discovery of inclusion in an experiment might result in a sense of violation[79,80]. There are no clear rules for the degree to which such information should be given; this depends on the type of intervention studied and the potential to withdraw from the intervention or data collection, and should therefore be covenanted with patient boards, the professionals and ethical committees involved. In some scenarios patient consent may be required, but it is not feasible to obtain consent prior to the clusters being randomized; in these situations the Ottawa statement recommends obtaining consent of patients as soon as is feasible after cluster randomization.

Another issue in cluster randomized trials is *recruitment bias*; there can be lack of blinding to allocation status of those identifying or recruiting individuals into a cluster randomized trial because individuals may be recruited to the trial after the clusters have been randomized[76,81]. The knowledge of whether each cluster is an 'intervention' or 'control' cluster could affect the types of participants recruited. Farrin *et al.* showed differential participant recruitment in a pilot trial of low back pain randomized by primary care practice; a greater number of less severe participants were recruited to the 'active management' practices[82]. The design of the main trial was therefore changed from a randomization on cluster level (practices) towards a randomization on patient level stratified by practice[83]. In some situations, however, it is possible to blind participants and those recruiting participants to the cluster assignment at the time of enrollment[49].

The lack of independence among individuals within the same cluster (intracluster correlation) has implications for the *sample size calculation*. Standard sample size approaches lead to an underpowered study. The inflation in sample size depends on average cluster size and the degree of correlation within clusters, ρ, also known as the intracluster (or intraclass) correlation coefficient (ICC). If m is the cluster size (assumed to be the same for all clusters), then the inflation factor, or "design effect," associated with cluster randomization is $1+(m-1)\rho$[81]. Although typically ρ is small (often <0.05) and is often not known when a trial is planned (and can only be estimated with error after a trial is completed), its impact on the inflation factor can be considerable when clusters are large. In general, the power is increased more easily by increasing the number of clusters rather than the cluster size[84].

Cluster-randomized controlled trials also present special requirements for analysis. The data could be analyzed as if each cluster was a single individual, but this approach ignores the information collected on all participants within a cluster and hence may not use the full richness of the dataset[84]. When analyses are on the patient or participant level, many cluster-randomized trials are analyzed by incorrect statistical methods, not taking the clustering into account. For example, Eldridge *et al.* reviewed 152 cluster-randomized trials in primary care of which 41% did not account for clustering in their analyses[85]. Such analyses create a 'unit of analysis error' and produce over-precise results (the standard error of the estimated intervention effect is too small) and *P* values that are too small. If simple statistical tests are used to compare effects on the patient level, adjustments have to be made to account for the clustering effect. For example, test statistics based on chi-squared or F-tests should be divided by the design effect (as described earlier), while test statistics based on the *t*-test or the *z*-test should be divided by the square root of the design effect[86]. However, several modeling techniques can incorporate patient level data such as mixed linear models, hierarchical linear modeling and generalized estimating equations. These modeling techniques allow the inherent correlation within clusters to be modeled explicitly, and thus a 'correct' model can be obtained[87].

Reporting on cluster-randomized trial requires additional information. In 2012 Campbell *et al.*[84] reported the extension to the CONSORT 2010 statement in which recommendations for the reporting of cluster randomized trials are presented. The main problem associated with their design, conduct, analysis, and interpretation, compared with individually randomized trials, is that two different units of measurement—the cluster and the patient—are used. Each needs to be reported carefully[84].

A clear example of a cluster randomized controlled trial is the ARTIST study reported by Ravaud *et al.*, which compared a novel management program with usual care for knee OA[62]. The care management program involved standardized consultations, provided during three goal oriented visits (education on OA and treatment management; information on physical exercises; information on weight loss). In this study the rheumatologists were randomized and not the individual patients. Each rheumatologist had to include the first two patients who complied with the inclusion criteria. Patients were informed that they were participating in a trial comparing different forms of consultations. They were informed about the content of the consultations to which they were assigned but not the consultation programm the other arm received. The primary outcomes were measured at patient level;

these included patients' weight and time spent on physical exercises at 4 months. Taking the clustering effect into account, the decrease in weight as well as the physical activity score was greater in the standardized consultation group than in the usual care group.

Stepped wedge designs

Stepped wedge randomized trials involve sequential roll-out of a new intervention to the units of randomization (either individuals or, more often, clusters such as primary care practices, hospital clinics or geographical districts), such that each unit 'steps' from offering the control intervention to the new intervention over time. In this design, all randomized units eventually have the new intervention, although the order in which they receive it is determined at random[88]. The stepped wedge design can be described as a type of crossover design in which different clusters cross over at different time points, but only in one direction; from offering the control to offering the new intervention[89]. More than one cluster or unit of randomization may start the new intervention at any given time point, but the time at which the new intervention begins is randomized. For example, a trial including 12 primary care practices might randomly allocate one practice each month for 12 months, such that each month one practice changes from offering usual care for patients with OA (the control phase) to a new evidence-based intervention (the intervention phase). In this example, one primary care practice will have only 1 month of offering usual care before it steps up to offer the new intervention, and at the other end of the trial, one practice will offer usual care for a year before it steps up to deliver the new intervention. Alternatively, the same trial may choose to randomly allocate two or three practices per month to step up to the new intervention, shortening the overall length of the trial. The stepped wedge design necessitates baseline data collection at the time when none of the clusters or units receive the new intervention, as well as data collection at each point in all clusters or units when a new unit or group receives the new intervention[88,89]. The overall effectiveness of the new intervention is determined by comparing the data from the control phase and the intervention phase from all the randomized units.

Stepped wedge designs are thought to be particularly useful for evaluating the overall impact of an intervention (e.g., a new implementation strategy for best practice for OA), where the new intervention or strategy has either already been shown to be effective in an individually randomized trial (the so-called phase IV effectiveness trials)[89] or where there is general consensus by the relevant stakeholders that the new intervention provides more good than harm (i.e., in cases where there is a lack of true collective equipoise about the merits of the new intervention)[88,89]. The design has also been justified in some cases where the intervention under investigation is already a recommended or adopted policy but lacks evidence of effectiveness[90]. It may also be the design of choice if there are financial, logistical or operational reasons that make it impossible to provide the new intervention simultaneously to many randomized units at the same time[88] but rather where phased implementation is preferable. In the example above, it might be the case that the new intervention is supported by only one implementation facilitation team that must travel to each participating primary care practice in turn. In addition, a stepped wedge design might be justified where there are doubts about whether beneficial effects seen by a new intervention in a traditional randomized trial are reproducible in the real world when scaled up to larger populations and communities[91].

There are many purported advantages of the stepped wedge design. The design may increase the motivation of participants or clusters to take part in the trial (as they will all eventually receive the new intervention)[92], it may help address the dilemma of withholding the new intervention when there is no collective equipoise in the scientific or clinical community, and it can be helpful to study the effect of context given that the intervention is implemented in multiple settings with often different characteristics (and the intervention may work in some but not all of these settings). In addition, the phased implementation of the new intervention in the stepped wedge design means that it is possible to improve the intervention or its implementation where necessary before the next unit is randomized[93], the design can help to detect trends in the effectiveness of the intervention over time by conducting a step-by-step comparison, and can increase statistical power as the design involves both within and between cluster comparisons[89].

A key disadvantage is that the timeline for the trial is usually extended, so stepped wedge trials often are lengthier in total than the equivalent simple, two or three arm parallel trial design. It is also not well suited to testing some types of interventions, as it is ideal if there are short time intervals between the application of the intervention and the expected change in key outcomes[90]. Where it is important to also measure longer-term outcomes (follow-up participants beyond the time period in which stepping takes place), this may mean the trial length extends beyond the timelines within scope of many research funders. Multiple data collection points are required, which can be burdensome for participants or clinicians and financially expensive and therefore, ideally, stepped wedge trials make use of routinely collected data. Finally the data analysis of this type of trial is more complicated than other designs[89]. Particular analysis challenges are controlling for temporal trends in outcome variables and accounting for repeated measures on the same participants throughout the trial[88].

There are also some key challenges with this design. Preventing contamination between the clusters receiving the new intervention and those still waiting to step up to the new intervention can be a problem, as can ensuring some level of blinding. Since it is often impossible to blind participants or those delivering the intervention (as both will usually be aware of the step from control to the new intervention), blinding assessors is important[88]. The clusters are randomly allocated to the time at which they move from control to the new intervention, and thus all of the preparatory work must be completed with each unit of randomization such that they are indeed 'ready' to implement the new intervention at the timeslot in which they are randomized to do so. This can be particularly challenging where the new intervention involves significant new training for personnel, where new systems need to be adopted to provide the new intervention in a timely fashion, or indeed where there are multiple other competing priorities that may get in the way of the cluster being able to change to the new intervention at the randomized timeslot (e.g., primary care practices that are subject to multiple competing priorities and unforeseen circumstances that must be made a higher priority). In a cluster stepped wedge trial, an important decision is the number of clusters randomized at each time step, given that optimal power is obtained when each cluster is randomized to the new intervention at its own randomization step[89]. Randomizing multiple clusters at each time point reduces the overall number of measurement time points and this significantly reduces power. Like classic cluster randomized trials, sample size calculations need to include the cluster design effect (the intraclass correlation coefficient)[90]. In addition, stepped wedge designs may not be suitable for testing all interventions, since those for which the full treatment effect is not expected to be realized for some time (until more than one time interval after the intervention is introduced) also significantly reduces power. It is desirable to ensure that each measurement time period is long enough so that the effect of the new intervention is fully realized before the next time period begins[89].

In summary, stepped wedge trials may be a useful way of evaluating intervention effectiveness during routine implementation. Two systematic reviews have shown that stepped wedge trial designs are increasingly used in the health research field to evaluate the implementation of interventions in routine practice[88,90], with the number of steps ranging from 2 to 36 steps and the period between steps ranging from 12 days to 1.5 years[90]. However, many stepped wedge designs published so far do not fulfill the methodological requirements of a controlled trial[88], and many do not ensure blinded assessment of key outcomes. Whilst the stepped wedge design has many advantages and their use has been advocated[88], the choice of a stepped wedge trial over a classic cluster RCT is not without controversy[94], and careful planning is needed to address the known challenges. To date, as far as we know, there are no completed, published stepped wedge randomized trials specifically addressing the implementation of best practice in OA, although there is a lack of consistency in reports clearly describing trial design as a stepped wedge[90].

An example of a trial currently in progress that is relevant to OA, rather than specifically focused on OA, is the cluster randomized, stepped wedge trial from Dreischulte *et al.*[95]. Their team is testing the effectiveness of a multi-faceted information technology-based intervention in reducing high-risk prescribing of nonsteroidal anti-inflammatory and anti-platelet medications in primary care (The DQIP study). The study involves 40 general practices (family practices) in Scotland, which will be randomized to one of 10 start dates (at four weekly intervals). Due to the nature of the complex intervention (involving a web-based informatics tool that provides weekly feedback at practice level, prompts the review of patients and summarizes each patients risk factors, as well as educational outreach visits to practices and payments for each patients reviewed), it is not possible to blind practices, the core research team or the data analyst. However, the outcome assessment is objective and completely automated. Outcomes during the DQIP intervention will be compared to care before its introduction. Their rationales for selection of the stepped wedge design are both logistical (they need to stagger the start of the new intervention at the different practices as they need to deliver educational outreach visits) and related to concerns about practices dropping out of the trial (a conventional two arm design would mean no financial incentives to practices in the usual care arm).

Adaptive trial designs and adaptive interventions

Adaptive trial designs allow for planned modifications of the study, based on accumulating data[96,97]. There are many options for adaptive trial designs[96,98], but the overarching goals are to improve trial efficiency, reduce sample size, and/or increase the likelihood of finding an effect if one does exist[99]. Adaptive clinical trial designs have been utilized for decades in pharmaceutical research and other efficacy and effectiveness trials[98], with little application to implementation research to date. However, these methods can certainly be applied to implementation research to improve efficiency of these studies. *Adaptive interventions* involve a sequence of individually tailored decision rules that specify whether, how, and when the intensity, type or delivery method of an intervention is altered for a given study participant[100]. For example, these interventions often involve increasing the intervention intensity in some way if participants fail to meet specified benchmarks for improvement or response[101]. Although we know of no completed adaptive interventions trials for OA treatment, some studies have evaluated adaptive interventions for physical activity and weight management (key behavioral strategies for OA) in other patient groups[102,103]. Sequential Multiple Assignment Randomized Trials (SMART) are one type of adaptive intervention, involving multiple stages at which participants are randomized to a set of treatment options[104–106]. Adaptive interventions have most commonly been applied to effectiveness trials[101] but have recently been utilized in implementation research[107]. In this context, adaptive interventions can allow clinical settings or patients not responding to an initial implementation strategy to receive an augmented or different type of intervention[107].

Process and formative evaluations in implementation trials

Implementation trials generally have as their goal improvements in quality of care through delivery of interventions designed to drive research evidence into *real-world* clinical practice. As such, the interventions being evaluated are often complex, with multiple interconnecting parts or components[108], must be tailored to local circumstances (available human and other resources)[54], and require substantial investment of resources to implement. As a result, a phased, non-linear[54,109] approach to evaluation, which incorporates considerable pilot work, and both *formative* (*outcome*) and *process* components to the evaluation, is recommended[34,37,110–112].

Formative or outcome evaluations measure the degree to which the research objectives have been achieved (did the intervention work?). As the key question being addressed in an implementation trial is whether or not the intervention is effective in everyday practice[113], *formative evaluation* of an implementation trial should incorporate validated measures of effectiveness and cost effectiveness that are relevant to patients, providers and other key stakeholders, such as policy makers, commissioners or payers of services, and administrators[114]. While a single primary outcome and a small number of secondary outcomes might be most straightforward, this approach may not make best use of the data or provide an adequate assessment of the success of an intervention across a range of domains[54]. As a result, *a priori* consideration of a range of outcome measures, including measures of unintended consequences, may be preferred. Regardless of which outcome measures are selected, each should be evaluated to ensure its feasibility, validity and responsiveness in the context of the proposed intervention prior to trial implementation.

For complex interventions, however, it is also – if not equally – important to ensure the intervention components are likely to succeed in achieving the desired outcomes, and once launched, to understand how the intervention works (e.g., what are the active ingredients and how are they exerting their effects?). Thus *process evaluations* are often used in pre-testing the components of the intervention to enable their refinement before the trial is launched[108]. For example, the success of many complex interventions depends on its acceptability to the participating clinicians, patients or other key stakeholders[108,115]. In the quality improvement arena, both barriers to and facilitators of clinician engagement in such initiatives have been elucidated[116–118]. Attention to ensuring identified barriers have been overcome and facilitators optimized is therefore key to address in implementation trial pilot work[119]. *Process evaluations* can also be nested within an implementation trial to assess the fidelity and quality of implementation, clarify potential causal mechanisms, and identify contextual factors associated with variation in outcomes[120,121]. Within the context of a multi-site implementation trial, where the same intervention may be implemented and received in different ways, process evaluation is particularly necessary. For example, process evaluation is useful in determining whether or not lack of an intervention effect was due to implementation failure or genuine ineffectiveness. Process evaluation can furthermore provide valuable insight into one or more of the following: why an intervention fails or has unexpected consequences, how a successful intervention works and how it can be optimized, differential

success of intervention components, contextual factors that affect an intervention, reach of the intervention, and the way effects vary in subgroups[122].

Mixed qualitative and quantitative methods are often required to perform a comprehensive process evaluation – to provide a detailed understanding of variations within and interactions between the components of the intervention[123]. Strategies that may be used to collect process level information include: semi-structured interviews where open ended questions regarding feelings, knowledge, opinions, experiences, perceptions are used and data recorded; focus groups; semi-structured and structured in-depth interviews using key informant or other community members; Delphi method using expert opinion and reiteration; observations from fieldwork descriptions of activities; and case studies[124]. Ideally, process evaluation is performed and analyzed before the formative trial outcome data in order to avoid bias in interpretation. While incorporating a qualitative component to the process evaluation will increase research costs, greater cost must be balanced against the potential for greater explanatory power and understanding of the generalizability findings and potential for sustainability of the intervention[125].

Limitations

Although we aimed to provide a comprehensive review of topics related to implementation research and its application to OA RCTs, a formal systematic review was not conducted due to the broad scope of the manuscript. There are other theoretical models applicable to implementation research, as well as detailed resources describing implementation research methods[34,126].

Facilitators and barriers of implementation strategies

Theoretical frameworks such as RE-AIM and KTA provide guidance on strategies to facilitate successful development and evaluation of interventions that are of value to stakeholders and can be adopted and sustained in clinical settings. However, there are common barriers to implementation studies, including: identifying and engaging all of the appropriate stakeholders and consumers, differing priorities among stakeholders, logistical challenges to weaving new interventions and processes into real-world clinical settings, competing priorities within clinical settings, and identifying resources and motivated leadership to sustain effective interventions or practices. Implementation research is a relatively young but rapidly emerging field; much is still unknown about the most effective strategies for implementation and knowledge translation, particularly when targeting policymakers and broad organizational culture[127,128]. However, there is some evidence that multifactorial interventions are more effective than single interventions[129]. In addition, confirming principles of the KTA framework, a systematic review showed that tailored interventions to overcome identified barriers to changing clinical practice are more effective than general dissemination of guidelines[130].

Summary principles for OA implementation trials

The following are key principles for OA implementation trials, generated by consensus of co-authors and based on general guidance for implementation research, best practices for development and evaluation of complex interventions, specific needs and considerations related to OA and its treatment:

1. Implementation studies are essential for testing whether results translate into real-world practice, and more of these types of RCTs are needed to inform which strategies support real change in practice for managing OA. Based on this need, it is recommended that all RCTs evaluating the effectiveness of OA treatments: (1) consider issues of generalizability in all aspects of the study design[131] and (2) whenever possible, incorporate implementation aims that will inform future uptake or delivery of effective interventions[25].
2. Development of OA implementation interventions should involve a rigorous developmental phase, including consideration of relevant theoretical frameworks and ensuring feasibility and acceptability to stakeholders[45,53,54].
3. OA implementation trials should utilize recruitment methods and participant inclusion/exclusion criteria that maximize generalizability. For example, patients should represent a range of general health states and resources to enact behavioral change, and healthcare professional participants should represent a broad range in terms of clinical experience and interest level in the clinical condition of OA.
4. OA implementation trials should involve strong analytic plans for evaluating both processes and outcomes[54].
5. Costs and cost effectiveness of implementation strategies should be considered as outcomes for OA implementation trials, and these should follow best practices for RCTs[132].
6. Complex interventions, often the focus of implementation trials, may work best if tailored to local circumstances[53,54]. Therefore OA implementation trials should consider interventions that allow some flexibility in the delivery approach.
7. Given the plethora of potential study designs from which to choose, teams studying implementation of best practice in OA need to clearly justify their choice, and attend to the challenges associated with their design.

Author contributions

All authors contributed to: conception and design, drafting the article, critical revision of the article, and final approval of the article. Kelli Allen (kdallen@email.unc.edu) takes responsibility for the integrity of the work as a whole.

Competing interest statement

The authors have no competing interests to declare.

Acknowledgement

The authors thank Timothy McAlindon, MD, MPH, who also served as a member of the work group for this project.

OARSI gratefully acknowledges support to defer in part the cost of printing of the Clinical Trial Recommendations from Abbvie, BioClinica, Boston Imaging Core Lab, and Flexion. The funding sources for printing had no role in the outcome of this manuscript.

References

1. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, *et al*. American College of Rheumatology 2012 recommendations for the use of non-pharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip and knee. Arthritis Care Res 2012;64:465–74.
2. Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, *et al*. OARSI recommendations for the management of hip and knee osteoarthritis Part III: changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage 2010;18:476–99.

3. Richmond J, Hunter DH, Irrgang JJ, Jones MH, Snyder-Mackler L, Van Durme D, *et al*. American Academy of Orthopaedic Surgeons clinical practice guideline on the treatment of osteoarthritis (OA) of the knee. J Bone Joint Surg 2010;92:990–3.
4. Do BT, Hootman JM, Helmick CG, Brady TJ. Monitoring healthy people 2010 arthritis management objectives: education and clinician counseling for weight loss and exercise. Ann Fam Med 2011;9:136–41.
5. Cisternas MG, Yelin E, Katz JN, Solomon DH, Wright EA, Losina E. Ambulatory visit utilization in a national, population-based sample of adults with osteoarthritis. Arthritis Rheum 2009;61:1694–703.
6. Song J, Hochberg MC, Chang RW, Hootman JM, Manheim LM, Lee J, *et al*. Racial and ethnic differences in physical activity guidelines attainment among people at high risk of or having knee osteoarthritis. Arthritis Care Res (Hoboken) 2013;65:195–202.
7. Centers for Disease C, Prevention. State-specific prevalence of walking among adults with arthritis – United States, 2011. MMWR Morb Mortal Wkly Rep 2013;62:331–4.
8. Hunter DJ, Neogi T, Hochberg MC. Quality of osteoarthritis management and the need for reform in the US. Arthritis Care Res (Hoboken) 2011;63:31–8.
9. Li LC, Sayre EC, Kopec JA, Esdaile JM, Bar S, Cibere J. Quality of nonpharmacological care in the community for people with knee and hip osteoarthritis. J Rheumatol 2011;38:2230–7.
10. Brand CA, Ackerman IN, Bohensky MA, Bennell KL. Chronic disease management: a review of current performance across quality of care domains and opportunities for improving osteoarthritis care. Rheum Dis Clin North Am 2013;39:123–43.
11. Glauser TA, Salinas GD, Roepke NL, Williamson JC, Reese A, Gutierrez G, *et al*. Management of mild-to-moderate osteoarthritis: a study of the primary care perspective. Postgrad Med 2011;123:126–34.
12. DeHaan MN, Guzman J, Bayley MT, Bell MJ. Knee osteoarthritis clinical practice guidelines – how are we doing? J Rheumatol 2007;34:2099–105.
13. Maserejian NN, Fischer MA, Trachtenberg FL, Yu J, Marceau LD, McKinlay JB, *et al*. Variations among primary care physicians in exercise advice, imaging, and analgesics for musculoskeletal pain: results from a factorial experiment. Arthritis Care Res (Hoboken) 2014;66:147–56.
14. Jordan KM, Sawyer S, Coakley P, Smith HE, Cooper C, Arden NK. The use of conventional and complementary treatments for knee osteoarthritis in the community. Rheumatology (Oxford) 2004;43:381–4.
15. Ganz DA, Chang JT, Roth CP, Guan M, Kamberg CJ, Niu F, *et al*. Quality of osteoarthritis care for community-dwelling adults. Arthritis Care Res 2006;55:241–7.
16. Bohensky MA, Sundararajan V, Andrianopoulos N, de Steiger RN, Bucknill A, Kondogiannis CM, *et al*. Trends in elective knee arthroscopies in a population-based cohort, 2000-2009. Med J Aust 2012;197:399–403.
17. Kim S, Bosque J, Meehan JP, Jamali A, Marder R. Increase in outpatient knee arthroscopy in the United States: a comparison of National Surveys of Ambulatory Surgery, 1996 and 2006. J Bone Joint Surg Am 2011;93:994–1000.
18. Hawker G, Guan J, Judge A, Dieppe P. Knee arthroscopy in England and Ontario: patterns of use, changes over time, and relationship to total knee replacement. J Bone Joint Surg Am 2008;90:2337–45.
19. Hawker GA, Wright JG, Glazier RH, Coyte PC, Harvey B, Williams JI, *et al*. The effect of education and income on need and willingness to undergo total joint arthroplasty. Arthritis Rheum 2002;46:3331–9.
20. Rahman MM, Kopec JA, Sayre EC, Greidanus NV, Aghajanian J, Anis AH, *et al*. Effect of sociodemographic factors on surgical consultations and hip or knee replacements among patients with osteoarthritis in British Columbia, Canada. J Rheumatol 2011;38:503–9.
21. Judge A, Welton NJ, Sandhu J, Ben-Shlomo Y. Equity in access to total joint replacement of the hip and knee in England: cross sectional study. BMJ 2010;341:c4092.
22. Steel N, Melzer D, Gardener E, McWilliams B. Need for and receipt of hip and knee replacement–a national population survey. Rheumatology (Oxford) 2006;45:1437–41.
23. Skinner J, Weinstein JN, Sporer SM, Wennberg JE. Racial, ethnic, and geographic disparities in rates of knee arthroplasty among medicare patients. N Engl J Med 2003;349:1350–9.
24. Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: the Chronic Osteoarthritis Management Initiative of the U.S. Bone and Joint Initiative. Semin Arthritis Rheum 2014;434:701–12.
25. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012;50:217–26.
26. Cochrane Effective Practice and Organisation of Care Group (EPOC). In: Reources for Review Authors, Vol. 2014. Oslo: Norwegian Knowledge Centre for the Health Services; 2013.
27. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999;89:1322–7.
28. Glasgow RE, Lichtenstein E, Marcus AC. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. Am J Public Health 2003;93:1261–7.
29. Kessler R, Glasgow RE. A proposal to speed translation of healthcare research into practice: dramatic change is needed. Am J Prev Med 2011;40:637–44.
30. Glasgow RE. What types of evidence are most needed to advance behavioral medicine? Ann Behav Med 2008;35:19–25.
31. Klesges LM, Estabrooks CA, Dzewaltowski DA, Bull BS, Glasgow RE. Beginning with the application in mind: designing and planning health behavior change interventions to enhance dissemination. Ann Behav Med 2005;29:66–75.
32. Feldstein A, Glasgow R. A practical, robust implemenation and sustainability model (PRISM) for integrating research findings into practice. Jt Comm J Qual Patient Saf 2008;34:228–73.
33. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, *et al*. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26:13–24.
34. Straus SE, Tetroe J, Grahah ID. Knowledge Translation in Health Care: Moving from Evidence to Practcie. Chichester: Wiley-Blackwell/BMJ; 2013.
35. Field B, Booth A, Ilott I, Gerrish K. Using the Knowledge to Action Framework in practice: a citation analysis and systematic review. Implement Sci 2014;9:172.
36. Straus SE, Tetroe JM, Graham ID. Knowledge translation is the use of knowledge in health care decision making. J Clin Epidemiol 2011;64:6–10.
37. Knowledge Translation Canada. Vol. 2015.
38. Tugwell PS, Santesso NA, O'Connor AM, Wilson AJ, Effective Consumer Investigative G. Knowledge translation for effective consumers. Phys Ther 2007;87:1728–38.

39. Sibley KM, Salbach NM. Applying knowledge translation theory to physical therapy research and practice in balance and Gait assessment: a case report. Phys Ther 2014. [Epub ahead of print].
40. Brachaniec M, Depaul V, Elliott M, Moore L, Sherwin P. Partnership in action: an innovative knowledge translation approach to improve outcomes for persons with fibromyalgia. Physiother Can 2009;61:123–32.
41. Kastner M, Straus SE. Application of the Knowledge-to-Action and Medical Research Council frameworks in the development of an osteoporosis clinical decision support tool. J Clin Epidemiol 2012;65:1163–70.
42. Sibley KM, Straus SE, Webster F, Jaglal SB. Moving balance and mobility evidence in to action: a primer in knowledge translation. Gait Posture 2011;33:527–31.
43. Brosseau L, Wells G, Brooks-Lineker S, Bennell K, Sherrington C, Briggs A, *et al*. Internet-based implementation of non-pharmacological interventions of the "people getting a grip on arthritis" educational program: an international online knowledge translation randomized controlled trial design protocol. JMIR Res Protoc 2015;4:e19.
44. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012;7:37.
45. French SD, Green SE, O'Connor DA, McKenzie JE, Francis JJ, Michie S, *et al*. Developing theory-informed behaviour change interventions to implement evidence into practice: a systematic approach using the Theoretical Domains Framework. Implement Sci 2012;7:38.
46. McEvoy R, Ballini L, Maltoni S, O'Donnell CA, Mair FS, Macfarlane A. A qualitative systematic review of studies using the normalization process theory to research implementation processes. Implement Sci 2014;9:2.
47. Finch TL, Mair FS, O'Donnell C, Murray E, May CR. From theory to 'measurement' in complex interventions: methodological lessons from the development of an e-health normalisation instrument. BMC Med Res Methodol 2012;12:69.
48. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. JAMA 2003;290:1624–32.
49. Allen KD, Bosworth HB, Brock DS, Chapman JG, Chatterjee R, Coffman CJ, *et al*. Patient and provider interventions for managing osteoarthritis in primary care: protocols for two randomized controlled trials. BMC Musculoskelet Disord 2012;13:60.
50. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, *et al*. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. CMAJ 2009;180:E47–57.
51. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, *et al*. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008;337:a2390.
52. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. J Chronic Dis 1967;20:637–48.
53. Medical Research Council G. A Framework for the Development and Evaluation of RCTs for Complex Interventions to Improve Health. London 2000.
54. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, *et al*. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;337:a1655.
55. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, *et al*. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001;134:663–94.
56. Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. Lancet 2002;359:696–700.
57. Minns Lowe CJ, Wilson MS, Sackley CM, Barker KL. Blind outcome assessment: the development and use of procedures to maintain and describe blinding in a pragmatic physiotherapy rehabilitation trial. Clin Rehabil 2011;25:264–74.
58. Gooch K, Marshall DA, Faris PD, Khong H, Wasylak T, Pearce T, *et al*. Comparative effectiveness of alternative clinical pathways for primary hip and knee joint replacement patients: a pragmatic randomized, controlled trial. Osteoarthritis Cartilage 2012;20:1086–94.
59. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, *et al*. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869.
60. Walsh N, Cramp F, Palmer S, Pollock J, Hampson L, Gooberman-Hill R, *et al*. Exercise and self-management for people with chronic knee, hip or lower back pain: a cluster randomised controlled trial of clinical and cost-effectiveness. Study protocol. Physiotherapy 2013;99:352–7.
61. Tejedor Varillas A, León Vázquez F, Lora Pablos D, Pérez Martín A, Vargas Negrín F, Gómez de la Cámara A, *et al*. Can an intervention on clinical inertia have an impact on the perception of pain, functionality and quality of life in patients with hip and/or knee osteoarthritis? Results from a cluster randomised tria. Aten Primaria 2012;44:65–72.
62. Ravaud P, Flipo RM, Boutron I, Roy C, Mahmoudi A, Giraudeau B, *et al*. ARTIST (osteoarthritis intervention standardized) study of standardised consultation versus usual care for patients with osteoarthritis of the knee in primary care in France: pragmatic randomised controlled trial. BMJ 2009;338:b421.
63. Victor CR, Triggs E, Ross F, Lord J, Axford JS. Lack of benefit of a primary care-based nurse-led education programme for people with osteoarthritis of the knee. Clin Rheumatol 2005;24:358–64.
64. Ravaud P, Giraudeau B, Logeart I, Larguier JS, Rolland D, Treves R, *et al*. Management of osteoarthritis (OA) with an unsupervised home based exercise programme and/or patient administered assessment tools. A cluster randomised controlled trial with a 2x2 factorial design. Ann Rheum Dis 2004;63:703–8.
65. Vitiello MV, Rybarczyk B, Von Korff M, Stepanski EJ. Cognitive behavioral therapy for insomnia improves sleep and decreases pain in older adults with co-morbid insomnia and osteoarthritis. J Clin Sleep Med 2009;5:355–62.
66. Tsai PF, Chang JY, Beck C, Kuo YF, Keefe FJ. A pilot cluster-randomized trial of a 20-week Tai Chi program in elders with cognitive impairment and osteoarthritic knee: effects on pain and other health outcomes. J Pain Symptom Manage 2013;45:660–9.
67. Marra CA, Cibere J, Grubisic M, Grindrod KA, Gastonguay L, Thomas JM, *et al*. Pharmacist-initiated intervention trial in osteoarthritis: a multidisciplinary intervention for knee osteoarthritis. Arthritis Care Res (Hoboken) 2012;64:1837–45.
68. Rosemann T, Joos S, Laux G, Gensichen J, Szecsenyi J. Case management of arthritis patients in primary care: a cluster-randomized controlled trial. Arthritis Rheum 2007;57:1390–7.
69. Hurley MV, Walsh NE, Mitchell HL, Pimm TJ, Patel A, Williamson E, *et al*. Clinical effectiveness of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain: a cluster randomized trial. Arthritis Care Res 2007;57:1211–9.

70. Reininga IH, Wagenmakers R, van den Akker-Scheek I, Stant AD, Groothoff JW, Bulstra SK, *et al*. Effectiveness of computer-navigated minimally invasive total hip surgery compared to conventional total hip arthroplasty: design of a randomized controlled trial. BMC Musculoskelet Disord 2007;8:4.
71. Veenhof C, Koke AJ, Dekker J, Oostendorp RA, Bijlsma JW, van Tulder MW, *et al*. Effectiveness of behavioral graded activity in patients with osteoarthritis of the hip and/or knee: a randomized clinical trial. Arthritis Rheum 2006;55:925–34.
72. Dziedzic KS, Healey EL, Porcheret M, Ong B, Main CJ, Jordan KP, *et al*. Implementing the NICE osteoarthritis guidelines: a mixed methods study and cluster randomised trial of a model osteoarthritis consultation in primary care – the Management of OsteoArthritis In Consultations (MOSAICS) study protocol. Implement Sci 2014;9:95.
73. Halloran ME, Longini Jr IM, Struchiner CJ. Design and interpretation of vaccine field studies. Epidemiol Rev 1999;21:73–88.
74. Donner A, Klar N. Pitfalls of and controversies in cluster randomization trials. Am J Public Health 2004;94:416–22.
75. Hutton JL. Are distinctive ethical principles required for cluster randomized controlled trials? Stat Med 2001;20:473–88.
76. Eldridge S, Ashby D, Bennett C, Wakelin M, Feder G. Internal and external validity of cluster randomised trials: systematic review of recent trials. BMJ 2008;336:876–80.
77. Taljaard M, Weijer C, Grimshaw JM, Eccles MP, Ottawa Ethics of Cluster Randomised Trials Consensus G. The Ottawa Statement on the ethical design and conduct of cluster randomised trials: precis for researchers and research ethics committees. BMJ 2013;346:f2838.
78. Eldridge S, Kerry S, Torgerson DJ. Bias in identifying and recruiting participants in cluster randomised trials: what can be done? BMJ 2009;339:b4006.
79. Ashcroft RE. Human research subjects, selection of. In: Chadwick F, Ed. Encyclopedia of Applied Ethics. San Diego, CA: Academic Press; 1998. 927–639.
80. Snowdon C, Elbourne D, Garcia J. Zelen randomization: attitudes of parents participating in a neonatal clinical trial. Control Clin Trials 1999;20:149–71.
81. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. BMJ 2003;327:785–9.
82. Farrin A, Russell I, Torgerson D, Underwood M, Team UBT. Differential recruitment in a cluster randomized trial in primary care: the experience of the UK back pain, exercise, active management and manipulation (UK BEAM) feasibility study. Clin Trials 2005;2:119–24.
83. Brealey S, Burton K, Coulton S, Farrin A, Garratt A, Harvey E, *et al*. UK Back pain Exercise and Manipulation (UK BEAM) trial–national randomised trial of physical treatments for back pain in primary care: objectives, design and interventions [ISRCTN32683578]. BMC Health Serv Res 2003;3:16.
84. Campbell MK, Piaggio G, Elbourne DR, Altman DG, Group C. Consort 2010 statement: extension to cluster randomised trials. BMJ 2012;345:e5661.
85. Eldridge SM, Ashby D, Feder GS, Rudnicka AR, Ukoumunne OC. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. Clin Trials 2004;1:80–90.
86. Murray DM. The Design and Analysis of Group Randomised Trials. Oxford: Oxford University Press; 1998.
87. Campbell MK. Analysis of cluster randomized trials in primary care: a practical approach. Fam Pract 2000;17:192–6.
88. Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. BMC Med Res Methodol 2006;6:54.
89. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials 2007;28:182–91.
90. Mdege ND, Man MS, Taylor Nee Brown CA, Torgerson DJ. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. J Clin Epidemiol 2011;64:936–48.
91. Sanson-Fisher RW, Bonevski B, Green LW, D'Este C. Limitations of the randomized controlled trial in evaluating population-based health interventions. Am J Prev Med 2007;33:155–61.
92. Hutson AD, Reid ME. The utility of partial cross-over designs in early phase randomized prevention trials. Control Clin Trials 2004;25:493–501.
93. Pearson D, Torgerson D, McDougall C, Bowles R. Parable of two agencies, one of which randomizes. Ann Am Acad Polit Soc Sci 2010;628:11–29.
94. Kotz D, Spigt M, Arts IC, Crutzen R, Viechtbauer W. Researchers should convince policy makers to perform a classic cluster randomized controlled trial instead of a stepped wedge design when an intervention is rolled out. J Clin Epidemiol 2012;65:1255–6.
95. Dreischulte T, Grant A, Donnan P, McCowan C, Davey P, Petrie D, *et al*. A cluster randomised stepped wedge trial to evaluate the effectiveness of a multifaceted information technology-based intervention in reducing high-risk prescribing of non-steroidal anti-inflammatory drugs and antiplatelets in primary medical care: the DQIP study protocol. Implement Sci 2012;7:24.
96. Kairalla JA, Coffey CS, Thomann MA, Muller KE. Adaptive trial designs: a review of barriers and opportunities. Trials 2012;13:145.
97. Brown CH, Ten Have TR, Jo B, Dagne G, Wyman PA, Muthen B, *et al*. Adaptive designs for randomized trials in public health. Annu Rev Public Health 2009;30:1–25.
98. Coffey CS, Kairalla JA. Adaptive clinical trials: progress and challenges. Drugs R D 2008;9:229–42.
99. Cook T, DeMets DL. Review of draft FDA adaptive design guidance. J Biopharm Stat 2010;20:1132–42.
100. Collins LM, Murphy SA, Bierman KL. A conceptual framework for adaptive preventive interventions. Prev Sci 2004;5:185–96.
101. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A "SMART" design for building individualized treatment sequences. Annu Rev Clin Psychol 2011;8:21–48.
102. Adams MA, Sallis JF, Norman GJ, Hovell MF, Hekler EB, Perata E. An adaptive physical activity intervention for overweight adults: a randomized controlled trial. PLoS One 2013;8:e82901.
103. Hall KS, Pieper CF, Edelman DE, Yancy Jr WS, Green JB, Lum H, *et al*. Lessons learned when innovations go awry: a baseline description of a behavioral trial-the Enhancing Fitness in Older Overweight Veterans with Impaired Fasting Glucose study. Transl Behav Med 2011;1:573–87.
104. Murphy SA. An experimental design for the development of adaptive treatment strategies. Stat Med 2005;24:1455–81.
105. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more

potent eHealth interventions. Am J Prev Med 2007;32:S112–8.
106. Nahum-Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, Fabiano GA, *et al*. Experimental design and primary data analysis methods for comparing adaptive interventions. Psychol Methods 2012;17:457–77.
107. Kilbourne AM, Abraham KM, Goodrich DE, Bowersox NW, Almirall D, Lai Z, *et al*. Cluster randomized adaptive implementation trial comparing a standard versus enhanced implementation intervention to improve uptake of an effective re-engagement program for patients with serious mental illness. Implement Sci 2013;8:136.
108. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, *et al*. Framework for design and evaluation of complex interventions to improve health. BMJ 2000;321:694–6.
109. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, *et al*. Designing and evaluating complex interventions to improve health care. BMJ 2007;334:455–9.
110. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008;27:379–87.
111. Hotopf M. The pragmatic randomized controlled trial. Adv Psychiatr Treat 2002;8:326–33.
112. Power R, Langhaug LF, Nyamurera T, Wilson D, Bassett MT, Cowan FM. Developing complex interventions for rigorous evaluation–a case study from rural Zimbabwe. Health Educ Res 2004;19:570–5.
113. Haynes B. Can it work? Does it work? Is it worth it? The testing of healthcare interventions is evolving. BMJ 1999;319:652–3.
114. Stiefel M, Nolan K. A Guide to Measuring the Triple Aim: Population Health, Experience of Care, and Per Capita Cost. . In: IHI Innovation Series White Paper. Cambridge, MA: Institute for Healthcare Improvement; 2012.
115. Reinertsen JL, Bisognano M, Pugh MD. Seven Leadership Leverage Points for Organization-level Improvement in Health Care. 2 edn. Cambridge, MA: Institutes for Health Care Improvement; 2007.
116. Lee TH. Turning doctors into leaders. Harv Bus Rev 2010;88:50–8.
117. Dickinson H, Ham C. Engaging Doctors in Leadership: Review of the Literature. Birmingham: Academy of Medical Royal Colleges; 2008.
118. Siriwardena AN. Engaging clinicians in quality improvement initiatives: art or science? Qual Prim Care 2009;17:303–5.
119. Reinertsen JL, Gosfield AG, Rupp W, Whittington JW. Engaging Physicians in a Shared Quality Agenda. Cambridge, MA: Institute for Healthcare Improvement; 2007.
120. Bradley F, Wiles R, Kinmonth AL, Mant D, Gantley M. Development and evaluation of complex interventions in health services research: case study of the Southampton heart integrated care project (SHIP). The SHIP Collaborative Group. BMJ 1999;318:711–5.
121. Oakley A, Strange V, Bonell C, Allen E, Stephenson J, Team RS. Process evaluation in randomised controlled trials of complex interventions. BMJ 2006;332:413–6.
122. Wight D, Obasi A. Unpacking the "black box": the importance of process data to explain outcomes. In: Stephenson J, Imrie J, Bonell C, Eds. Effective Sexual Health Interventions: Issues in Experimental Evaluation. Oxford: Oxford University PRess; 2002:151–66.
123. Oakley A. Design and analysis of social intervention studies in health research. In: Bowling A, Ebrahim S, Eds. Handbook of Health Research Methods. Maidenhead: Open University PRess; 2005:246–65.
124. World Health Organization. WHO Global Strategy on Diet, Physical Activity and Health: A Framework to Monitor and Evaluate Implementation. Geneva: World Health Organization; 2008.
125. Helitzer DL, Davis SM, Gittelsohn J, Going SB, Murray DM, Snyder P, *et al*. Process evaluation in a multisite, primary obesity-prevention trial in American Indian schoolchildren. Am J Clin Nutr 1999;69:816S–24S.
126. Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. Am J Prev Med 2012;43:337–50.
127. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement Sci 2012;7:50.
128. Parmelli E, Flodgren G, Beyer F, Baillie N, Schaafsma ME, Eccles MP. The effectiveness of strategies to change organisational culture to improve healthcare performance: a systematic review. Implement Sci 2011;6:33.
129. Boaz A, Baeza J, Fraser A, European Implementation Score Collaborative G. Effective implementation of research into practice: an overview of systematic reviews of the health literature. BMC Res Notes 2011;4:212.
130. Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S, *et al*. Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2010:CD005470.
131. Dzewaltowski DA, Estabrooks PA, Klesges LM, Bull S, Glasgow RE. Behavior change intervention research in community settings: how generalizable are the results? Health Promot Int 2004;19:235–45.
132. Hlatky MA, Owens DK, Sanders GD. Cost-effectiveness as an outcome in randomized clinical trials. Clin Trials 2006;3:543–51.