IMPLEMENTATION SCIENCE

METHODOLOGY Open Access

How pragmatic is it? Lessons learned using PRECIS and RE-AIM for determining pragmatic characteristics of research

Bridget Gaglio[1*], Siobhan M Phillips[2], Suzanne Heurtin-Roberts[2], Michael A Sanchez[2] and Russell E Glasgow[3]

Abstract

Background: The need for high-quality evidence that is applicable in real-world, routine settings continues to increase. Pragmatic trials are designed to evaluate the effectiveness of interventions in real-world settings, whereas explanatory trials aim to test whether an intervention works under optimal situations. There is a continuum between explanatory and pragmatic trials. Most trials have aspects of both, making it challenging to label and categorize a trial and to evaluate its potential for translation into practice.

Methods: We summarize our experience applying the Pragmatic-Explanatory Continuum Indicator Summary (PRECIS) combined with external validity items based on the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework to three studies to provide a more robust and comprehensive assessment of trial characteristics related to translation of research. We summarize lessons learned using domains from the combined frameworks for use in study planning, evaluating specific studies, and reviewing the literature and make recommendations for future use.

Results: A variety of coders can be trained to use the PRECIS and RE-AIM domains. These domains can also be used for diverse purposes, content areas, and study types, but are not without challenges. Both PRECIS and RE-AIM domains required modification in two of the three studies to evaluate and rate domains specific to study type. Lessons learned involved: dedicating enough time for training activities related to the domains; use of reviewers with a range of familiarity with specific study protocols; how to best adapt ratings that reflect complex study designs; and differences of opinion regarding the value of creating a composite score for these criteria.

Conclusions: Combining both frameworks can specifically help identify where and how a study is and is not pragmatic. Using both PRECIS and RE-AIM allows for standard reporting of key study characteristics related to pragmatism and translation. Such measures should be used more consistently to help plan more pragmatic studies, evaluate progress, increase transparency of reporting, and integrate literature to facilitate translation of research into practice and policy.

Keywords: Pragmatic trials, External validity, Research translation, PRECIS, RE-AIM

Background

Over the last several years, there has been a substantial movement toward practical, pragmatic implementation research that will translate into usable health-related policies, programs and practices [1–4]. Pragmatic research is conducted internationally in wide ranging settings [5–8]. Funding to support pragmatic research and evaluation is provided by major health institutions such as the National Institutes of Health in the United States (U.S.), the U.S. Department of Veterans Health Affairs, the Canadian Institutes for Health Research, and the National Health Service's National Institute for Health Research in the United Kingdom [9,10]. Pragmatic research is increasingly being conducted in networks of primary care practices, health maintenance organizations, and other research networks such as the Patient Centered Outcomes Research Institute patient-powered research networks and the clinical data research networks [11].

* Correspondence: bridget.gaglio@kp.org
[1]Mid-Atlantic Permanente Research Institute/Kaiser Permanente Mid-Atlantic States, 2101 E. Jefferson Street, Suite 300, Rockville, MD 20852, USA
Full list of author information is available at the end of the article

The differentiation of pragmatic from explanatory research can be traced to a seminal paper by Schwartz and Lellouch [12] wherein they define explanatory research as conducted under optimal circumstances to determine the 'efficacy' of an intervention while pragmatic research tests an intervention under usual conditions. This distinction is important because trials are frequently designed as explanatory investigations, when the researchers' intent is actually to answer the pragmatic question of effectiveness under usual or differing conditions. Inasmuch as trials are inadequately formulated for the type of research question asked, research outcomes are compromised and effort wasted [12]. The importance of pragmatic research has been given a major boost by the development of criteria and evaluation tools intended to increase transparency of research and results reporting and provide a means for practitioners and policy makers to assess local applicability of trial findings [13–15].

The 'Pragmatic-Explanatory Continuum Indicator Summary' or PRECIS framework was developed to assist trial designers to assess where a trial is positioned along the pragmatic to explanatory continuum [16]. The main purpose of PRECIS is to determine the degree to which study design decisions align with the trial's stated purpose, and thus was originally intended to be used at the design stage. The tool is comprised of 10 domains: participant eligibility criteria, experimental intervention flexibility, experimental intervention practitioner expertise, comparison intervention, comparison intervention practitioner expertise, follow-up intensity, primary trial outcome, participant compliance with prescribed intervention, practitioner adherence to study protocol, and analysis of primary outcome (see Table 1).

The original intent of the PRECIS framework was to inform trial designs by providing a visual display in the form of a hub and spoke diagram, where each of the 10 domains are represented by a line depicting the pragmatic-explanatory continuum. No numerical anchors were originally used. The endpoint closest to the hub represented a more explanatory study, whereas the endpoint furthest away from the hub represented a more pragmatic study [16]. However, modifications have been proposed and tested in variety of ways in an attempt to expand its utility to evaluate studies post completion, including use in systematic reviews [17–21]. Modifications have included quantifying the pragmatic-explanatory nature of a study by using numeric rating systems, where each domain is scored on a Likert-type scale. The original scale ranged from 0 to 4 where 0 represented an extremely pragmatic study and 4 was extremely explanatory. Over time the range most commonly used has been 1 to 5 (scales of 0 to 4 and 1 to 20 have also been used) [22]. Regardless of which scale is used, all have transposed the endpoints so the smaller number represents an extremely explanatory study and the larger number represents an extremely pragmatic study. Another modification was made to accommodate evaluating systematic reviews. Each study in a systematic review is scored individually on each of the 10 original PRECIS domains [20]. After the individual scoring, a 10-domain average for each individual trial can be calculated, as well as a single domain average across all trials included in the review and an overall combined average for the entire systematic review. Regardless of which version was used, all studies concluded that PRECIS was useful in designing trials and assessing the level of pragmatism of a trial or a body of evidence. However, PRECIS does not include domains to evaluate generalizability and applicability of a pragmatic trial to a specific context. Thus, additional domains are required.

The RE-AIM framework, which is an acronym for reach, effectiveness, adoption, implementation, and maintenance, was created out of the need for improved reporting on key issues related to robustness, translatability, and public health impact of health research [23,24]. RE-AIM was developed as a response to trends toward research conducted under optimal efficacy conditions instead of in real-world, complex settings [25] and is intended to be used at all stages of research from planning through evaluation and reporting, and across different types of research (*e.g.*, effectiveness, implementation, and dissemination trials) [26]. RE-AIM domains address issues focused on setting and participant representativeness, setting/site engagement with intervention, intervention adaptation during the study, program sustainability, and monetary/resource costs of an intervention. Over the past 14 years, RE-AIM has been applied to a wide range of conditions and study settings and has evolved to include additional items necessary for translation of research findings, such as use of qualitative methods and assessment of unanticipated consequences, both negative and positive (*e.g.* generalization effects). These domains address pragmatic and external validity issues not included within the PRECIS domains and are shown in Table 1. Originally, RE-AIM domains were not defined by a rating scale [23]. The first scale was modeled after our first use of the PRECIS rating scale that ranged from 0 to 4, where 0 represented an extremely pragmatic study and 4 was extremely explanatory. With subsequent uses of the RE-AIM domains, the scale has been changed to remain identical to the PRECIS scale where the smaller number represents an extremely explanatory study and the larger number represents an extremely pragmatic study.

The purpose of this article is to build on the work that has been done on the use and applicability of the PRECIS and RE-AIM frameworks by summarizing our experience applying these models to three studies that have combined both frameworks to provide a more robust and comprehensive assessment of issues related to

Table 1 Summary of the Pragmatic Explanatory Continuum Indicator Summary (PRECIS) and the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework criteria

Framework	Domain	Criteria as described in published literature
PRECIS [16]	Participant eligibility criteria	Extremely pragmatic studies would enroll all participants who have the condition of interest regardless of their responsiveness, past compliance, and co-morbidities.
		Extremely explanatory studies would use a variety of exclusion criteria to identify those individuals who would least likely respond to the intervention.
	Experimental intervention flexibility	Extremely pragmatic studies would leave the details of implementation of the experimental condition up to the practitioner.
		Extremely explanatory studies would have strict instructions for every intervention element (*e.g.*, timing of intervention delivery, tactics used, educational materials used).
	Experimental intervention practitioner expertise	Extremely pragmatic studies would be conducted by individuals from a variety of backgrounds and levels of experience and in a variety of settings.
		Extremely explanatory studies would only be conducted by credentialed or seasoned interventionists.
	Comparison intervention	Extremely pragmatic studies would use 'usual care' or available alternative interventions.
		Extremely explanatory studies would use a placebo condition as the comparison group.
	Comparison intervention practitioner expertise	Extremely pragmatic studies would have the comparison intervention delivered by individuals with full range of expertise levels, with only ordinary attention to training or experience.
		Extremely explanatory studies would have the expertise of the comparison condition standardized at a high level so as to be able to detect whatever comparative benefits the experimental intervention might have.
	Follow-up intensity	Extremely pragmatic studies would not have formal follow-up visits. Instead administrative databases would be used to assess outcomes.
		Extremely explanatory studies would have formal follow-up visits on a prescribed schedule and more extensive data collection than would occur in usual care or routine practice.
	Primary trial outcome	Extremely pragmatic studies would have a primary outcome that is objectively measured, has clinical significance, and is one that can be assessed under usual conditions.
		Extremely explanatory studies would have a primary outcome that is known to be a direct and immediate relation to the intervention. The outcome may also require specialized training or testing not normally used to determine outcome status.
	Participant compliance with prescribed intervention	Extremely pragmatic studies would have no measurement of compliance and no special strategies to try to improve it.
		Extremely explanatory studies would have study participants' compliance monitored closely and would have strategies to enhance compliance.
	Practitioner adherence to study protocol	Extremely pragmatic studies would have no measurement of practitioner adherence and no special strategies to try to improve it.
		Extremely explanatory studies would have close monitoring of how well the practitioners and study sites are adhering to the study protocol.
	Analysis of primary outcome	Extremely pragmatic studies would include all participants in the analyses regardless of compliance (*e.g.*, intent-to-treat analysis).
		Extremely explanatory studies would focus on analyses that allowed for estimating the maximum benefit of the intervention (*e.g.*, analysis restricted to those considered 'completers').
RE-AIM [23]	Reach	The absolute number, proportion, and representativeness of individuals who are willing to participate in a given initiative, intervention, or program. Reporting of exclusion criteria and percentage of potential participants excluded, Including the use of qualitative data to understand recruitment.
		Extremely pragmatic studies would include participants that are typical of those individuals with the specified condition (including hard-to-reach individuals).
		Extremely explanatory studies would include individuals who are not typical on most or all characteristics of those with the specified condition.
	Effectiveness	The impact of the intervention on outcomes. Measure of primary outcome, including potential negative effects, quality of life, and economic outcomes. Moderation analysis. Measure of short-term attrition and use of qualitative data to understand outcomes.
		Extremely pragmatic studies would have primary outcomes that are meaningful to patients and providers. Explicit discussion of efforts to prevent harm to participants; report on unintended harmful or beneficial consequences of the intervention.

Table 1 Summary of the Pragmatic Explanatory Continuum Indicator Summary (PRECIS) and the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework criteria *(Continued)*

	Extremely explanatory studies would have primary outcomes that are meaningful to investigators. No efforts to capture unintended harmful or beneficial consequences of the intervention; only what is necessary for IRB requirements.
Adoption	The absolute number, proportion, and representativeness of settings and the individuals within those settings who deliver the program who are willing to initiate a program. Reporting of setting and staff exclusion criteria and percentage excluded. Use of qualitative data to understand settling level adoption and staff participation.
	Extremely pragmatic studies would have few or no setting exclusions. Sites are either randomly selected or purposely selected for variation.
	Extremely explanatory studies would have many exclusion criteria for settings and/or would try to get the 'best' sites to participate.
Implementation	Fidelity to study/program protocol and adaptations made to intervention during study/program. Cost of intervention in terms of time and money. Consistency of implementation across staff, time, setting, and subgroups – focus is on process. Use of qualitative data to understand implementation.
	Extremely pragmatic studies would provide detailed reporting of modifications made and rationale. Explicit discussion of efforts to contain costs and to make the intervention feasible for low resource settings would also be included.
	Extremely explanatory studies would have no mention of modifications to protocols or measures. No effort to contain costs; uses state of the art resources and procedures.
Maintenance	The extent to which a program or policy becomes institutionalized or part of the routine organizational practices and policies. If and how the program/policy was adapted long-term. Some measure/discussion of alignment to organization mission or sustainability of business model. Use of qualitative data to understand setting level institutionalization.
	Within the RE-AIM framework, maintenance also applies at the individual level. At the individual level, maintenance has been defined as the long-term effects of a program on outcomes 6 or more months after the most recent intervention contact. Measure of long-term attrition and differential rates by patient characteristics and/or treatment condition. Use of qualitative data to understand long-term effects.
	Extremely pragmatic studies would focus on explicit plans for handing off intervention to setting/site for continuation of the program/intervention after the completion of the study.
	Extremely explanatory studies would not have a report of efforts to continue the intervention after completion of the study.

Note: PRECIS ratings were originally done on a hub and spoke visual diagram with no numerical anchors on the lines representing each domain along the pragmatic-explanatory continuum. The three exemplar studies used a numerical rating for each domain; one study used a scale of 1 to 5 (1 = explanatory and 5 = pragmatic) and the other two used a scale of 0 to 4 (where 0 = pragmatic and 4 = explanatory in one and 0 = explanatory and 4 = pragmatic in the other). As for the RE-AIM ratings, RE-AIM domains were originally not defined by a rating scale. The three studies used a numerical scale for each domain that was identical to the PRECIS rating scales for each study respectively.

translation of research. We begin by describing experiences of using both PRECIS and RE-AIM frameworks in three different studies. Thereafter, we summarize lessons learned using the combined criteria and make recommendations for future use. We conclude with a discussion on implications for the broader issue of designing and reporting results for studies intended to promote translation into policy and practice.

Methods

Description of studies to illustrate use of the frameworks

The following three studies illustrate our experiences applying both PRECIS and RE-AIM frameworks. These three studies were selected because they are the only studies to our knowledge that have combined both frameworks, we have access to the data, and they illustrate different applications (*e.g.*, planning, use to describe different interventions in a collaborative project and to conduct a literature review).

The Practice-Based Opportunities for Weight Reduction (POWER) Trials Collaborative Research Group included three individual studies funded by the National Heart Lung and Blood Institute (NHLBI) [19]. Although the studies did not share a common intervention protocol, all three tested a primary care-based intervention to reduce weight among obese primary care patients who had at least one other cardiovascular disease risk factor [27]. The POWER trials had common components to facilitate potential cross-site comparisons, but each protocol also incorporated distinct, trial-specific elements including different interventions and different secondary outcome measurements (see Table 2).

The second study was a systematic review of eHealth cancer prevention and control intervention trials [28].

Table 2 Use of the Pragmatic Explanatory Continuum Indicator Summary (PRECIS) and the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) Framework domains to evaluate three studies

	Practice-based Opportunities for Weight Reduction (POWER) trial [19]	Review of eHealth cancer prevention and control intervention trials [27]	My Own Health Report (MOHR) trial [30]
Setting	The POWER trials Collaborative Research Group tested weight loss interventions in primary care settings.	eHealth intervention trials that used patient facing technologies addressing all phases of the cancer care continuum (*i.e.*, prevention through end of life care).	Ongoing, delayed cluster randomized intervention trial in primary care settings.
Sample	Three individual studies (approximately 1,100 participants) aimed at reducing weight in primary care patients who were obese and had at least one other cardiovascular disease risk factor.	113 grouped studies (149 total papers) were reviewed.	One study (nine pairs of diverse primary care practices matched on practice type, practice ownership, geographic region, electronic health record infrastructure, and patient population served).
When study was conducted	2008 to 2011	Studies published between 1980 and mid-2010.	January 2013
When during the research process was the domains used (evaluation, planning, etc.)	Implementation of study protocols.	Reporting - evaluation of published, peer-reviewed literature.	Planning
Modifications made to domains	Individual composite scores were created for both sets of ratings - PRECIS and REAIM.	For studies in which eHealth intervention replaced practitioners with no personal or phone contact, 'not applicable' ratings were applied to relevant PRECIS domains on practitioner expertise and practitioner compliance to study protocol. For studies in which multiple interventions were compared, the most intensive intervention served as the experimental arm and least intensive intervention served as the control arm.	Added patient engagement items to RE-AIM domains.
Lessons learned from applying domains to type of study and stage of study	The utility of creating a composite score for each of the evaluation domains is questionable. Since results of the trials were not available at the time of applying the domains, one must wait to evaluate the utility of the domains in terms of the extent to which scores predict eventual success of programs and their adoption, implementation, and sustainability in real-world settings.	PRECIS and RE-AIM domains could be productively and reliably applied to a variety of eHealth studies. However, for studies in which eHealth interventions replaced practitioners with no personal or phone contact, deciding the best way to rate (or not rate) these domains was a challenge.	PRECIS domains prior to the start of the trial confirmed the design decisions aligned with the trial's stated pragmatic purpose.
Details of raters (number, who they were, how/why selected)	Nine reviewers total—six involved in one of the studies and three independent reviewers not associated with any part of the POWER studies.	Three PhD level and four masters level scientists were randomly paired for study reviews from a convenience sample.	Three PhD level scientists indirectly involved in the project were selected as convenience sample.
	Most had Ph.D. or M.D. degrees and at least moderate experience in clinical trials.	Three had previous experience using PRECIS.	Only one had previous experience using PRECIS.
Training (resources used, processes, etc.)	Read article on the PRECIS domains by Thorpe *et al.* [16] and reviewed the slide presentation on PRECIS by Sackett [33].	All reviewers read the article on PRECIS domains by Thorpe *et al.* [16] the RE-AIM original article by Glasgow, *et al.*, [22] and the slides on PRECIS by Sackett [33].	Read article on PRECIS domains by Thorpe *et al.* [16] and reviewed the slides by Sackett [33].
	RE-AIM training was integrated into PRECIS training following discussion of original PRECIS domains. No	Training sessions also served to develop consensus on all domains.	Conducted a one one-hour meeting to go over domains as a group and instructions for scoring.

Table 2 Use of the Pragmatic Explanatory Continuum Indicator Summary (PRECIS) and the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) Framework domains to evaluate three studies *(Continued)*

	background materials provided on RE-AIM.	The rating form was pilot-tested by all reviewers and refined based on the ratings of a sub-sample of four papers, not included in the study. After refinements and clarification of rating process, another two papers were assigned to the entire group.	
Evaluation process	Each reviewer read the centrally available protocol materials on each intervention.	Articles were identified and grouped into studies if more than one article per study. Reviewers were assigned by study grouped papers.	Study protocol was independently reviewed and scored by each reviewer and a consensus meeting was held to resolve any discrepancies
	Each reviewer read a background description of each project that appeared in Obesity and Weight Management [26].	After assignments, each reviewer conducted a brief web search of publications within PubMed to identify any relevant papers to the main study paper.	
	Any questions the reviewer had were answered by a contact person at each site, not involved in the ratings.	Each study was independently reviewed and scored by each reviewer and a consensus meeting was held to resolve any discrepancies among any 'not applicable' ratings.	
	Reviewers then independently rated each of the three projects on the PRECIS domains followed by rating on RE-AIM domains.		
Reliability procedures	PRECIS domains - The overall kappa inter-rater reliability on the composite PRECIS score was $r = 0.88$. The interclass correlation for individual items was 0.72 and for averages 0.96.	Each study was rated by two reviewers.	Percent agreement using both exact and within one-point measures
	RE-AIM domains- Inter-rater reliability on the composite scale was $r = 0.71$. The interclass correlations were 0.45 for single items and 0.83 for averages.	Weighted percent agreement scores were calculated for PRECIS and RE-AIM domains.	
Lessons learned from training on use of domains	Leave ample time for training. Several conference calls were held among reviewers to discuss and attempt to specify precisely what was meant by each dimension and to develop and refine the rating instruments and instructions.	Leave ample time for reviewers to become familiar with the evaluation domains. It may take multiple sessions to arrive at consensus on how to routinely apply the domains.	Domains can be used with relatively little training to evaluate a study protocol.
	Applications of criteria should be completed by independent reviewers not associated with studies being rated in addition to the team members due to observed differences in rating one's own project.		

For this review, eHealth interventions were defined as 'the use of emerging information and communication technology, especially the Internet, to improve or enable health and health care', [29] and included email, mobile phone text or applications, interactive voice response, automated and electronic programs, and computer tailored print but excluded telemedicine targeted solely at clinicians that did not have a patient or consumer facing interface. It included 113 studies across the cancer control continuum (*i.e.*, primary prevention, screening, treatment/disease management, survivorship, and end-of-life care) [30].

The third study is the My Own Health Report (MOHR) trial whose primary purpose was to study clinical implementation of and patient experience with the use of an automated health risk assessment and feedback system to help clinics focus on patient-centered care issues [31]. The MOHR trial used a paired, cluster randomized delayed intervention design with nine pairs of primary care clinics. The trial combined elements of pragmatic

trials, implementation science, systems science [32], and mixed methods approaches with practical outcome measures [33]. Research teams identified and selected matched clinics that were similar in type (*e.g.*, federally qualified health center, practice based research network, family practice, or internal medicine), and clinical characteristics including geographic region, approximate size and level of electronic health record integration. One clinic in each pair was randomized to early implementation while the second clinic was assigned to the delayed implementation condition.

Training on use of the domains

The three evaluations were conducted during different phases of the research process. For the POWER study, evaluation occurred during the implementation phase of the project. The eHealth evaluation was conducted after study completion as the review consisted solely of published literature, and the MOHR evaluation was conducted in the planning phase. A variety of reviewers were used in the three different evaluation exercises as described in Table 2. In the POWER trial, reviewers familiar and not familiar with the research protocols being evaluated were used. In eHealth, one reviewer was the lead investigator for one of the included studies. However, he was not assigned to review the study. None of the other reviewers were associated with any of the published works included in the review. In the MOHR trial, individuals indirectly associated with the study were used as reviewers. In all three cases, individuals were highly educated and trained in the research process as described below and in Table 2.

While the training process for the reviewers varied across each evaluation, all began in a similar fashion with reviewers studying the original PRECIS article [16] and the PowerPoint presentation by Dr. Sackett [34], and having two or more group meetings to discuss application of the rating criteria. The POWER study was our first use of the PRECIS framework. After review and discussions on applying the PRECIS domains to the POWER protocols, it was evident that additional domains were necessary to capture key contextual factors for translation. Thus the additional domains from the RE-AIM framework were added. Reviewers then re-assessed each of the protocols with the additional RE-AIM domains. The eHealth study had reviewers not familiar with the RE-AIM framework read the original RE-AIM article [23] in addition to reviewing the PRECIS training materials. Multiple training sessions were held to develop consensus on both frameworks among raters on all domains. The rating form that included both sets of evaluation criteria was piloted and refined based on the ratings of a subsample of four papers by all reviewers. After refinements and clarification of the rating process, all reviewers evaluated two additional papers to pilot the revised criteria. The MOHR study conducted a one-hour training session to review the criteria as a group and instructions for using the criteria.

Use of the domains

All three projects were rated on a 5-point Likert-type scale for the PRECIS and RE-AIM domains. The POWER study used the original 0 – 4 scale, as described in the Sackett presentation, where 0 was extremely pragmatic and 4 was extremely explanatory. However, the eHealth study used a 1 – 5 rating scale, as described in Koppenaal, *et al.* [20], and the MOHR study used a 0 – 4 scale, both such that the lower score, the more explanatory the trial and the higher the score, the more pragmatic the trial. In addition, the POWER and eHealth studies created composite scores for both PRECIS and RE-AIM domains.

In the POWER study, reviewers independently rated each of the three protocols on all PRECIS and RE-AIM domains using a paper rating form. In the eHealth study, two reviewers were randomly assigned to each study and reviewers rated approximately 38 studies each. All rating information for each study was collected via a web-based form in Survey Monkey. In the MOHR trial, reviewers rated the study protocol using a paper rating form.

Different approaches to inter-rater reliability were used because the three different studies had vastly different designs, strategies for allocating reviewers, number of reviewers, and number of studies rated per reviewer. Therefore, the approaches to assessing inter-rater reliability differed as well. For POWER, intraclass correlation coefficients were calculated for individual items and an overall kappa was calculated for each of the composite scores [19]. In the eHealth review, weighted percent agreement scores for PRECIS and RE-AIM domains were calculated [28]. For the MOHR study, percent agreement score for each PRECIS and RE-AIM domain was calculated using a standard of exact agreement [31].

Results

Experiences using PRECIS and RE-AIM to evaluate three different studies

The POWER trial was our first experience using PRECIS and the first numerical rating using RE-AIM domains. Although the PRECIS article examples and the presentation were useful background, there were several issues that were unclear to some reviewers, and we found it necessary to add explicit anchors for the ratings and to rate and discuss example studies not part of the formal evaluation. We also identified one person from each of the three POWER research centers very familiar with that center's protocol, not a reviewer (*e.g.*, a program manager) who was available to answer any questions and clarify issues that were unclear to reviewers from the study protocols.

The review of eHealth cancer prevention and control intervention trials was the first published article using PRECIS and RE-AIM to evaluate eHealth intervention (EHI) studies. Several unanticipated issues were unique to applying such domains to EHI studies. For studies in which EHI replaced practitioners with no personal or phone contact, 'not applicable' ratings were applied to relevant PRECIS domains on practitioner expertise and practitioner compliance to study protocol. Any discrepancies in 'not applicable' ratings between reviewers were identified and discussed for consensus during the data cleaning process. Additionally, reviewers had to discuss and agree upon assignment of experimental and control interventions for studies in which multiple interventions were compared. For these studies, the most intensive intervention served as the experimental arm and the least intensive intervention served as the control arm. We also found that few studies reported on factors related to cost and setting representativeness relevant to the RE-AIM domains and therefore, could not rate such aspects of the individual studies.

Because the MOHR study only used three reviewers for one protocol that they were already familiar with, there were far fewer issues in terms of training and rating. However, the reviewers did feel that the RE-AIM domains did not capture one factor they felt was important to generalizability, patient engagement. As such, this domain was added and rated. The reviewers tended to rate some aspects of the protocol highly with regard to pragmatism and generalizability. However, this was not consistent across specific domains and consensus discussions seemed to resolve any bias towards these responses.

Table 3 Average ratings (standard deviation) on PRECIS and RE-AIM domains: scores by study

Domains	POWER mean (SD)	eHealth mean (SD)	MOHR mean (SD)
Participant Eligibility Criteria	1.8 (0.6)	3.2 (0.7)	3.0 (0)
Experimental Intervention Flexibility	1.6 (0.9)	2.8 (0.9)	4.0 (0)
Practitioner Expertise (Experimental)	1.6 (0.9)	2.7 (0.9)	3.3 (0.6)
Comparison Intervention Flexibility	2.9 (1.2)	2.9 (1.2)	4.0 (0)
Practitioner Expertise (Comparison)	2.5 (1.3)	2.8 (1.0)	4.0 (0)
Follow-up Intensity	1.4 (1.1)	3.2 (0.9)	3.0 (0)
Primary Trial Outcome	2.4 (0.7)	2.9 (0.7)	3.3 (0.6)
Participant Compliance	1.2 (0.8)	3.2 (0.8)	2.7 (0.6)
Practitioner Adherence to Protocol	1.7 (0.9)	3.6 (1.0)	3.3 (0.6)
Analysis of Primary Outcome	2.0 (1.4)	3.5 (0.8)	3.3 (0.6)
Participant Representativeness	2.3 (0.8)	2.5 (0.7)	3.3 (0.6)
Setting Representativeness	2.1 (1.3)	2.8 (1.0)	3.0 (0)
Adaptation/Change	1.4 (1.0)	1.5 (0.6)	3.7 (0.6)
Sustainability	1.4 (1.0)	1.6 (0.7)	2.0 (0)
Costs/Feasibility of Treatment	2.0 (1.2)	1.6 (0.7)	4.0 (0)

Note: PRECIS: Pragmatic Explanatory Continuum Indicator Summary. RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance Framework. POWER Trials = Practice-Based Opportunities for Weight Reduction Trials Collaborative Research Group; eHealth Reviews = Systematic Review of eHealth Cancer Prevention and Control Interventions; MOHR = My Own Health Report. To make ratings comparable, the ratings were converted so that for all studies, scoring continuum: 1 = Explanatory, 5 = Pragmatic.

Lessons learned using PRECIS and RE-AIM frameworks

These three diverse applications illustrate that both the PRECIS and RE-AIM frameworks can be used for diverse purposes and across diverse content areas and types of studies. The following lessons can be taken from this experience. First, although the domains can be reliably coded by a variety of research staff after a short training activity, time should be dedicated to discussions about precise definitions for each domain and practice using the criteria (see Table 2). Second, reviewers in all studies found that the RE-AIM domains in combination with the PRECIS domains addressed important additional information related to pragmatic research. Both sets of domains can reveal meaningful differences across studies and across domains within a study. The most consistent and largest differences across studies were that studies were less pragmatic on the RE-AIM domains than on the original PRECIS domains (see Table 3). In particular, adaption, sustainability, and costs were seldom reported. It is both sobering and ironic that these types of issues are precisely the ones about which stakeholders most need information to consider adoption and replication of an intervention program [3]. Third, two of the three evaluations had reviewers who were directly or somewhat directly related to the study being evaluated. It was observed that reviewers directly involved with a study tended to rate their own study as more pragmatic than others. Having reviewers who are both familiar and unfamiliar with an intervention or program could help minimize, or control for, this finding. Fourth, given the nature of eHealth interventions, reliance on technology as the intervention delivery mechanism, the role of the practitioner (*i.e.*, practitioner expertise and adherence in PRECIS) was not applicable to many of the self-administered intervention studies. The impact of not scoring these two PRECIS domains is unclear and thus warrants further discussion on how to best incorporate or properly rate domains when used with eHealth and other automated intervention studies.

Fifth, given the PRECIS and RE-AIM domains focus on trials that have explicit experimental and control arms, reviewers had to designate study arms as experimental or control in research that compared three or more interventions. This is likely to present similar challenges to research

that studies multiple arms, such as comparative effectiveness research, adaptive design interventions, or multicomponent intervention trials. Sixth, there was difference of opinion regarding the value of calculating a summary score for both PRECIS and RE-AIM domains. Calculating such a score can be helpful but also potentially misleading. The summary score can give a sense as to where a study falls on the pragmatic-explanatory continuum as whole, but it masks the diversity of the individual domains. For example, two studies could have an identical overall PRECIS summary score. However, one study might have been much more explanatory in terms of eligibility criteria for trial participants and the other trial much more pragmatic on this domain. It is recommended that when overall summary scores are used, individual domains should also be reported to identify how results on different domains contribute to the overall score and to be able to assess how each domain aligns with the purpose of the study.

Discussion

An increasing number of programs and studies claim to be pragmatic. Use of both PRECIS and RE-AIM frameworks can be used to demonstrate specifically where and how an individual study, or group of studies, is and is not pragmatic. Comparing ratings of domains within the same study allows for understanding the pragmatic versus explanatory design elements of the trial. Whereas comparing domain ratings across trials allows clinicians, policy makers, and study reviewers to compare across studies to make meaningful judgments about which intervention has generalizability and applicability to their population(s) of interest and the level of reasonable effectiveness that can be expected in different contextual settings versus those in explanatory trials.

Several evaluation frameworks have been developed to facilitate translation of research findings. However, many are designed solely for evaluation [35]. Combining both PRECIS and RE-AIM allows for standard reporting of *both* development and evaluation over the life course of a study. In the planning phase of a study, PRECIS allows for assessing the match between the trial design and the research and RE-AIM can be used to provide greater detail relative to some PRECIS domains (*e.g.*, description of eligibility criteria and calculation of reach), and also to address other issues not in PRECIS important to potential adopting settings (*e.g.*, costs required, representativeness of settings). RE-AIM can be used across the entire span of the study to understand the *why* behind success or failure of a study by describing the context in which the study occurred [36]. PRECIS can be used periodically throughout the study and at study conclusion to assess how adaptations and changes made over the course of the study impact the design and whether the end result still aligned with the original purpose of the study, respectively.

There is considerable benefit to using both frameworks to assess key components necessary for designing and reporting results for studies intended to promote translation of research into practice and policy. However, there are still many questions that need to be explored as use of both frameworks increases. First, what is the best rating scale to use? Is a 5-point Likert scale or some other scale the best way to evaluate a study or should one solely use a diagram without defined end-points? Is there value to using a scale to assess each PRECIS and RE-AIM domain or is a visual diagram sufficient? If a visual diagram is sufficient, is the PRECIS 'spoke and hub' diagram effective for also displaying the RE-AIM domains?

Second, use of some PRECIS domains to rate some health services studies is currently problematic. For example, in the eHealth review, there were studies that evaluated automated interventions without involvement of practitioners, with no personal or phone contact. 'Not applicable' ratings were applied to the PRECIS domains on practitioner expertise and practitioner compliance to study protocol (see Additional file 1). Is this the right way to apply the domains or should it be given a score? Moreover, usual care comparison conditions could be viewed as either explanatory or pragmatic depending on the lens of the evaluator. For example, how would participant compliance be rated when a health educator meets with a patient regarding self-management of diabetes and the educator encourages the individual to problem solve concerning self-monitoring their blood sugar levels and/or exercise more frequently to reach their health goals? This is usual care so could be viewed as extremely pragmatic in nature. However, it could also be viewed as being explanatory as it is encouraging patients to be more compliant. Thus, use of RE-AIM in addition to PRECIS can help complete, or at minimum, provide additional information to help understand why PRECIS domains might be viewed as not relevant or interpreted differently by two different evaluators.

Third, who should be a reviewer? There are pros and cons to including reviewers who are intimately familiar with a project versus those who are completely independent. Although not investigated in these three studies, these domains could be used by both researchers and stakeholders including patients and practitioners to help collaboratively design pragmatic studies. Additional studies are needed to determine if the finding that those familiar with a study rate it as more pragmatic is a generalizable phenomenon, as we observed in the ratings of the MOHR study. If so, this would imply that familiarity with a study should be balanced across studies to prevent potential bias.

Strengths and limitations

This evaluation only included three studies, and replication, especially in different content areas and types of

settings, is needed. Other researchers are invited, especially those involved in team science [37] and community engaged projects to use the PRECIS and RE-AIM frameworks to increase collaboration and transparency, as well as for program planning and adaptations. Also, since the RE-AIM domains were developed to supplement the PRECIS domains for each of the three applications reported, the specific RE-AIM domains varied slightly across studies. The PRECIS domains are due to be revised later in 2014 [22], and at that time, it may be possible to also arrive at a common, standard set of accompanying RE-AIM domains, assuming they are still needed to supplement PRECIS. Strengths of the paper include the consistency of results and general usefulness of these rating tools across three different content areas, different phases of the research enterprise, and by different types of reviewers.

Conclusion

The importance of pragmatic trials and dissemination and implementation research to improve health and health care delivery in the U.S. is gaining increased attention [38,39]. Reporting on pragmatic rating criteria such as the PRECIS and RE-AIM scales can increase transparency and help reviewers and potential adoption settings make more informed judgments about programs and their applicability to different settings and under different conditions. However, because pragmatic research focuses on real-world applications of interventions, understanding the context in which it occurred is critical. Understanding whether a study design aligns with one's research question in terms of being pragmatic versus explanatory should not stand alone without an understanding of how the context of participants, setting, and processes involved affected the results. We encourage those planning and evaluating health research interventions to use and report on PRECIS and RE-AIM domains, and to contribute to their refinement.

Additional file

Additional file 1: Appendix PRECIS and RE-AIM domains rating form – all 3 studies combined.

Competing interests

The authors declare they have no competing interests.

Authors' contributions

BG and REG analyzed the data from the POWER study, MS and BG analyzed the data from the eHealth study, and SP and SHR analyzed the data from the MOHR study. All authors drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Funding Sources: Partially supported by National Heart, Lung, Blood Institute (NHLBI) – Grant #5 U01 HL087071-01, U01 HL087085-01, and U01 HL 087072–01 and the National Cancer Institute (NCI) - Grant # CA124401, CA140959, CA163526, and CA154549. The opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or views of the National Cancer Institute.

Author details

[1]Mid-Atlantic Permanente Research Institute/Kaiser Permanente Mid-Atlantic States, 2101 E. Jefferson Street, Suite 300, Rockville, MD 20852, USA. [2]Division of Cancer Control and Population Sciences/National Cancer Institute, 9609 Medical Center Drive, 3E212, Rockville, MD 20850, USA. [3]Colorado Health Outcomes Program and Department of Family Medicine, University of Colorado School of Medicine, 13199 E. Montview Blvd, Suite 300, Mail Stop F443, Room 323, Aurora, CO 80045, USA.

Received: 26 September 2013 Accepted: 8 July 2014
Published: 28 August 2014

References

1. Glasgow RE: **What does it mean to be pragmatic? Pragmatic methods, measures, and models to facilitate research translation.** *Health Educ Behav* 2013, **40:**257–265.
2. Lobb R, Colditz GA: **Implementation science and its application to population health.** *Annu Rev Public Health* 2013, **34:**235–251.
3. 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:**1624–1632.
4. Zwarenstein M, Treweek S: **What kind of randomized trials do we need?** *CMAJ* 2009, **180:**998–1000.
5. Rycroft-Malone J, Seers K, Crichton N, Chandler J, Hawkes CA, Allen C, Bullock I, Strunin L: **A pragmatic cluster randomised trial evaluating three implementation interventions.** *Implement Sci* 2012, **7:**80.
6. Schmidt B, Wenitong M, Esterman A, Hoy W, Segal L, Taylor S, Preece C, Sticpewich A, McDermott R: **Getting better at chronic care in remote communities: study protocol for a pragmatic cluster randomised controlled of community based management.** *BMC Public Health* 2012, **12:**1–8.
7. Shi GX, Liu CZ, Wang LP: **Application of pragmatic randomized controlled trial's design in clinical research of acupuncture.** *Zhongguo Zhong Xi Yi Jie He Za Zhi* 2010, **30:**193–196.
8. Voigt-Radloff S, Graff M, Leonhart R, Schornstein K, Vernooij-Dassen M, Olde-Rikkert M, Huell M: **WHEDA study: effectiveness of occupational therapy at home for older people with dementia and their caregivers–the design of a pragmatic randomised controlled trial evaluating a Dutch programme in seven German centres.** *BMC Geriatr* 2009, **9:**44.
9. Heisler M, Hofer TP, Schmittdiel JA, Selby JV, Klamerus ML, Bosworth HB, Bermann M, Kerr EA: **Improving blood pressure control through a clinical pharmacist outreach program in patients with diabetes mellitus in 2 high-performing health systems: the adherence and intensification of medications cluster randomized, controlled pragmatic trial.** *Circulation* 2012, **125:**2863–2872.
10. Mays GP, Scutchfield FD: **Advancing the science of delivery: public health services and systems research.** *J Public Health Manag Pract* 2012, **18:**481–484.
11. Mold JW, Peterson KA: **Primary care practice-based research networks: working at the interface between research and quality improvement.** *Ann Fam Med* 2005, **3**(Suppl 1):S12–20.
12. Schwartz D, Lellouch J: **Explanatory and pragmatic attitudes in therapeutical trials.** *J Clin Epidemiol* 2009, **62:**499–505.
13. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF: **Improving the quality of reporting of randomized controlled trials. The CONSORT statement.** *JAMA* 1996, **276:**637–639.
14. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D, CONSORT group, Pragmatic Trials in Healthcare (Practihc) group: **Improving the reporting of pragmatic trials: an extension of the CONSORT statement.** *BMJ* 2008, **337:**a2390.
15. Schulz KF, Altman DG, Moher D, CONSORT group: **CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.** *J Clin Epidemiol* 2010, **63:**834–840.
16. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, Tunis S, Bergel E, Harvey I, Magid DJ, Chalkidou K: **A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers.** *J Clin Epidemiol* 2009, **62:**464–475.

17. Selby P, Brosky G, Oh PI, Raymond V, Ranger S: **How pragmatic or explanatory is the randomized, controlled trial? The application and enhancement of the PRECIS tool to the evaluation of a smoking cessation trial.** *BMC Med Res Methodol* 2012, **12:**1–13.
18. Riddle DL, Johnson RE, Jensen MP, Keefe FJ, Kroenke K, Bair MJ, Ang DC: **The Pragmatic-Explanatory Continuum Indicator Summary (PRECIS) instrument was useful for refining a randomized trial design: experiences from an investigative team.** *J Clin Epidemiol* 2010, **63:**1271–1275.
19. Glasgow RE, Gaglio B, Bennett G, Jerome GJ, Yeh HC, Sarwer DB, Appel L, Colditz G, Wadden TA, Wells B: **Applying the PRECIS criteria to describe three effectiveness trials of weight loss in obese patients with comorbid conditions.** *Health Serv Res* 2012, **47**(3 pt 1):1051–1067.
20. Koppenaal T, Linmans J, Knottnerus A, Spigt M: **Pragmatic vs. explanatory: an adaptation of the PRECIS tool helps to judge the applicability of systematic reviews for daily practice.** *J Clin Epidemiol* 2011, **64:**1095–1101.
21. Tosh G, Soares-Weiser K, Adams CE: **Pragmatic vs explanatory trials: the pragmascope tool to help measure differences in protocols of mental health randomized controlled trials.** *Dialogues Clin Neurosci* 2011, **13:**209–215.
22. Loudon K, Zwarenstein M, Sullivan F, Donnan P, Treweek S: **Making clinical trials more relevant: Improving and validating the PRECIS tool for matching trial design decisions to trial purpose.** *Trials* 2013, **14:**115.
23. 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:**1322–1327.
24. Glasgow RE, Linnan LA: **Evaluation of theory-based interventions.** In *Health Behavior and Health Education: Theory, Research, and Practice.* 4th edition. Edited by Glanz K, Rimer BK, Viswanath K. San Francisco: Jossey-Bass; 2008:487–508.
25. Kessler R, Glasgow RE: **A proposal to speed translation of healthcare research into practice: dramatic change is needed.** *Am J Prev Med* 2011, **40:**637–644.
26. Gaglio B, Shoup JA, Glasgow RE: **The RE-AIM framework: a systematic review of use over time.** *Am J Public Health* 2013, **103:**e38–46.
27. Wells B: **Weight loss in obese adults with cardiovascular risk factors: three randomized control trials to assess interventions in clinical practice.** *Obesity and Weight Management* 2009, **5:**207–209 [http://online.liebertpub.com/doi/pdfplus/10.1089/obe.2009.0504]
28. Sanchez MA, Rabin BA, Gaglio B, Henton M, Elzarrad MK, Purcell P, Glasgow RE: **A systematic review of eHealth cancer prevention and control interventions: new technology, same methods and designs?** *Transl Behav Med* 2013, **3:**392–401.
29. Eng TR: *The eHealth Landscape: A Terrain Map of Emerging Information and Communication Technologies in Health and Health Care.* Princeton: Robert Wood Johnson Foundation; 2001.
30. Rabin BA, Glasgow RE, Kerner JF, Klump MP, Brownson RC: **Dissemination and implementation research on community-based cancer prevention: a systematic review.** *Am J Prev Med* 2010, **38:**443–456.
31. Krist AH, Glenn BA, Glasgow RE, Balasubramanian BA, Chambers DA, Fernandez ME, Heurtin-Roberts S, Kessler R, Ory MG, Phillips SM, Ritzwoller DP, Roby DH, Rodriguez HP, Sabo RT, Sheinfeld Gorin SN, Stange KC, MOHR Study Group: **Designing a valid randomized pragmatic primary care implementation trial: the my own health report (MOHR) project.** *Implement Sci* 2013, **8:**73.
32. Sterman JD: **Learning from evidence in a complex world.** *Am J Public Health* 2006, **96:**505–514.
33. Glasgow RE, Riley WT: **Pragmatic measures: what they are and why we need them.** *Am J Prev Med* 2013, **45:**237–243.
34. Sackett D: **PrECIs (Pragmatic-Explanatory Continuum Indicators).** 2009. Accessd: February 13, 2012; [http://www.support-collaboration.org/precis.pdf]
35. Gaglio B, Glasgow RE: **Evaluation approaches for dissemination and implementation research.** In *Dissemination and Implementation Research in Health.* Edited by Brownson RC, Colditz G, Proctor E. New York: Oxford University Press; 2012:327–356.
36. Kessler RS, Purcell P, Benkeser R, Glasgow RE, Klesges LM, Peek CJ: **What does it mean to 'employ' the RE-AIM model?** *Eval Health Prof* 2013, **36:**44–66.
37. Bennett LM, Gadlin H: **Collaboration and team science: from theory to practice.** *J Investig Med* 2012, **60:**768–775.
38. National Institutes of Health: *Pragmatic Research in Healthcare Settings to Improve Diabetes Prevention and Care (R18). Funding announcement PAR-13-366, 2013.* Bethesda, MD: National Institutes of Health. Accessed date: February 18, 2014.
39. Patient Centered Outcomes Research Institute: **Pragmatic Clinical Studies and Large Simple Trials to Evaluate Patient-Centered Outcomes.** 2014, [http://www.pcori.org/funding-opportunities/funding-announcements/pragmatic-clinical-studies-and-large-simple-trials-to-evaluate-patient-centered-outcomes/] Accessed date: February 18, 2014.

doi:10.1186/s13012-014-0096-x
Cite this article as: Gaglio *et al.*: How pragmatic is it? Lessons learned using PRECIS and RE-AIM for determining pragmatic characteristics of research. *Implementation Science* 2014 **9**:96.