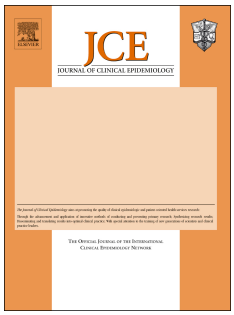
Accepted refereed manuscript of:

Forbes G, Loudon K, Treweek S, Taylor S & Eldridge S (2017) Understanding the applicability of results from primary care trials: lessons learned from applying PRECIS-2, *Journal of Clinical Epidemiology*, 90, pp. 119-126..

DOI: 10.1016/j.jclinepi.2017.06.007

# Accepted Manuscript




Understanding the applicability of results from primary care trials: lessons learned from applying PRECIS-2

Gordon Forbes, Kirsty Loudon, Shaun Treweek, Stephanie JC. Taylor, Sandra Eldridge

PII: S0895-4356(16)30249-9

DOI: 10.1016/j.jclinepi.2017.06.007

Reference: JCE 9421

To appear in: *Journal of Clinical Epidemiology*

Received Date: 22 July 2016

Revised Date: 15 February 2017

Accepted Date: 12 June 2017

Please cite this article as: Forbes G, Loudon K, Treweek S, Taylor SJ, Eldridge S, Understanding the applicability of results from primary care trials: lessons learned from applying PRECIS-2, *Journal of Clinical Epidemiology* (2017), doi: 10.1016/j.jclinepi.2017.06.007.

# Understanding the applicability of results from primary care trials: lessons learned from applying PRECIS-2

Authors: Gordon Forbes[1*], Kirsty Loudon[2], Shaun Treweek[3], Stephanie JC Taylor[1], Sandra Eldridge[1]

[1]Queen Mary, University of London [2]University of Stirling [3]University of Aberdeen

*Corresponding author

# Abstract

## Objective

To compare two approaches for trial teams to apply PRECIS-2 to pragmatic trials: independent scoring and scoring following a group discussion.

## Study Design and Setting

We recruited multidisciplinary teams who were conducting or had conducted trials in primary care in collaboration with the Pragmatic Clinical Trials Unit, Queen Mary University of London. Each team carried out two rounds of scoring on the 9 PRECIS-2 domains: first independently using an online version of PRECIS-2 and secondly following a discussion.

## Results

Seven teams took part in the study. Prior to the discussion within-team agreement in scores was generally poor and not all raters were able to score all domains; agreement improved following the discussion. The PRECIS-2 wheels suggested that the trials were pragmatic, though some domains were more pragmatic than others.

## Conclusion

PRECIS-2 can facilitate information exchange within trial teams. To apply PRECIS-2 successfully we recommend a discussion between those with detailed understanding of: what usual care is for the intervention, the trial's design including operational and technical aspects, and the PRECIS-2 domains. For some cluster randomised trials greater insight may be gained by plotting two PRECIS-2 wheels, one at the individual participant level and one at the cluster level.


**Key Words:** Randomized controlled trials; Clinical trial methodology; Pragmatic trial; primary care; Trial design.

## What is new?

- PRECIS-2 offers a new tool for trial teams to exchange information and knowledge about trial design.
- To score each PRECIS-2 domain successfully understanding is required of what usual care would be for the intervention, what is planned for the trial, and the interpretation of each domain in the context of the trial; a discussion amongst a multidisciplinary group may be required to score each domain.
- The trials included in this study were broadly pragmatic, areas where they were commonly less pragmatic were recruitment, and organisation (the level of resources used to deliver the intervention).
- When applying PRECIS-2 to cluster randomised trials greater insight may sometimes be gained by plotting two wheels, one at the level of the cluster and another at the level of individual participants

# Background

Many randomised controlled trials (RCTs) are carried out with the aim of informing health professionals, patients and policy makers about whether or not an intervention should be adopted in practice. To facilitate this it is recommended that trials taking a pragmatic design approach are carried out, testing the intervention under conditions as similar as possible to the conditions that would pertain if the intervention was rolled out in routine care (1-6). It may be that the intervention itself alters or replaces some aspects of usual care, but the principle remains that if the purpose of the trial is to directly inform clinical practice, aside from the intervention itself, other aspects of care should be as they usually would be in the real world. For brevity we refer to these types of conditions as "usual care". Recently there has been a growing interest in the benefits from more pragmatic designs. This has led to the publication of literature on how to design and conduct pragmatic trials (7-11), the latest of these developments, PRECIS-2 allows researchers to plot a graph illustrating whether their design is more or less pragmatic across a number of domains.

PRECIS-2 is an updated version of PRECIS (9) with significant changes including revisions to the domains, the addition of a Likert scoring scale and a website which can be used to support use of the tool (https://precis-2.org). PRECIS-2 (Figure 1) has nine domains covering different aspects of a trial: *eligibility, recruitment, setting, organisation, flexibility of delivery, flexibility of adherence*, *follow-up*, *primary outcome* and *primary analysis* (12). To apply PRECIS-2 each domain is scored from one to five - a score of one indicates an explanatory design with highly controlled or ideal conditions for the intervention and a score of five indicates a very pragmatic design, replicating usual care conditions for that domain. The tool was developed to be used at the design stage so that if the tool highlights that a trial design does not match the investigators' aims, they may choose to modify the design or reasons for the design may become more transparent. The tool can, however, also be used retrospectively as part as a critical appraisal of the generalisability of results from a trial or to illustrate a trial design to those using results.

Investigators who used the original PRECIS either discussed PRECIS wheels as a group to come to some consensus (13-18) or used independent scoring by different team members (19-21). Challenges in applying the tool were reported both by investigators who discussed PRECIS wheels and those that did not hold a discussion. The approach we used to apply PRECIS-2 was informed by this work (13-21). In the earlier studies applying the original PRECIS, the tool was applied both in the design phase, as the authors intended PRECIS to be used, and also to trials which were already completed. Likewise, in work described here we consider trials at all stages in design.

Our primary aim in this work was to inform investigators how best to use the tool regardless of whether it was to be used for design or retrospectively. We were particularly interested in doing so for trials at the pragmatic end of the pragmatic-explanatory spectrum for which PRECIS-2 was designed, and in exploring the difference between using the tool with and without discussion between group members.

**>> insert Figure 1: A blank PRECIS-2 Wheel*.<<**

*Reproduced with permission from (12)

# Methods

We invited trial teams carrying out trials in collaboration with the Pragmatic Clinical Trials Unit (PCTU) at Queen Mary University of London to partake in this study. To identify trials at the pragmatic end of the pragmatic-explanatory spectrum we focused on trials in primary care (generally considered more pragmatic than trials in other settings (22)). . We invited all trial teams that had recently completed their trial, were in the process of running their trial or were working towards submitting new proposals for funding.

Challenges encountered applying the original PRECIS tool came from a lack of information or difference in interpretation across three areas: what is happening or is being planned for a trial (15); what usual care would be if the intervention were to be implemented in practice (16); and how to interpret the PRECIS domains in the context of the trial (18). Researchers also found a bias towards raters scoring their own trial as more pragmatic (18). To overcome these problems we selected different raters including the trial chief investigator, who had both clinical expertise of usual practice and the trial design, and the trial manager and statistician, who had detailed information about the technical and operational aspects of the trial design, conduct, and analysis. In addition we sought two independent views of the design: a member of the relevant trial steering committee plus one of the authors of this study (GF). GF had understanding of the PRECIS-2 domains, gained from discussions with two of the authors of PRECIS-2 (KL, ST). For trials which were in the design stage not all these positions had been appointed so we included instead co-applicants on the trial grants with appropriate experience.

For each trial we asked the raters to independently score the trial using PRECIS-2 on two separate occasions: firstly pre-discussion using the online tool (www.PRECIS-2.org) and the information contained on the PRECIS-2 website and, secondly, post-discussion scoring following a group discussion of raters involved in the relevant trial. This discussion lasted up to two hours. PRECIS-2 scores were plotted using the median of the scores for all of the raters for each trial. The group discussion was based on pre-discussion scores and structured around the three key areas identified from work on PRECIS: the interpretation of each PRECIS-2 domain, what usual care was for the trial and what was planned or had happened in the trial in relation to each domain; the discussion aimed achieve consensus on these three areas. However, the discussion would move on if there were areas of disagreement that could not be addressed within the time available. The post-discussion scores were made independently at the end of the meeting.

We report the median of the pre-discussion and post-discussion scores for each domain and each trial as well as plotting PRECIS-2 wheels displaying the median pre-discussion and post-discussion scores. To assess the level of agreement within trial teams from the two sets of scores (pre-discussion and post-discussion), we compare the range of the scores for each domain across raters for each trial. For each domain we report the median and interquartile range (IQR) of the ranges.

Ethical approval for this study was granted by the Queen Mary Research Ethics Committee (reference QMREC1360d). The members of the trials teams involved in this study provided informed consent.

**>> insert Figure 2: PRECIS-2 wheels for trials included in the study. The solid shapes indicate the median post discussion scores. Dashed lines show the median pre-discussion scores. <<**

## Results

We invited 10 trial teams working on trials supported by the PCTU to participate in the study and seven agreed to take part. The three trials that declined to take part all did so due to lack of availability of the trial team during the study period. The seven trials are described in table 1. All seven trials aimed to answer the question of whether the intervention of interest would work in practice and the chief investigators considered their designs to be pragmatic.

The PRECIS-2 wheels (figure 2) indicate that trials being carried out by the PCTU are generally pragmatic. Domains for which trials were most often less pragmatic were *Recruitment* where four trials had post-discussion scores of three or less and *Organisation* where four trials had scores of three or less. All of the trials with less pragmatic scores for *Recruitment* were individually randomised, the cluster randomised trials were more pragmatic in their recruitment.

We produced eight PRECIS-2 wheels (figure 2) for the seven trials included in the study. For STOP, one of the cluster randomised trials, during the discussion the trial team thought that it would help them understand their design more if they produced two sets of scores, one set at the individual participant level and one at the cluster level for the pharmacy/pharmacist. For this trial, not all the domains were applicable at the individual level and this is reflected in the PRECIS-2 wheel by the domains with no score showing. This decision was consistent with the purpose of PRECIS-2 and description of its use by those who developed it (12), but was not deemed necessary by the rating teams for the other two cluster randomised trials.

Small differences between the pre-discussion scores and post-discussion score were observed. Almost all the differences between the pre-discussion scores and post-discussion scores were one point or less (figure 2, table 2). Thus for the majority of trials, design discussion led, at most, to a refinement in the extent a domain was considered pragmatic or explanatory rather than a change from a trial being rated as more pragmatic to more explanatory or visa versa. The one exception to this was the STOP trial which was at an earlier stage of development than the other trials. In this trial larger changes between the pre and post discussion scores were observed. This was in part due to greater understanding of how planned design related to the PRECIS-2 domains gained from plotting PRECIS-2 wheels at the individual and cluster level, and in part due to some aspects of the design not being clear to all the raters pre-discussion.

The median range of the pre-discussion scores was two or greater for five out of nine domains: *Recruitment* (median range 3), O*rganisation* (median range 2), F*lexibility of delivery* (median range 2), F*lexibility of adherence* (median range 2), and *Follow-up* (median range 2). This indicates poor agreement sometimes to the extent that raters disagreed about whether the trial was essentially pragmatic or essentially explanatory. Agreement in the post-discussion scores was good, for all domains the median range of scores was one or less: raters disagreed over the extent to which a domain was pragmatic or explanatory rather than whether the domain was pragmatic or explanatory.

Not all raters were able to score every domain using the online tool either due to insufficient information in the trial documents or due to insufficient knowledge of usual care. This was particularly an issue for *Flexibility of adherence* (21% of raters unable to score), *Flexibility of delivery* (14%), and *Organisation* (11%) (table 3).

**>>insert Table 1: Description of trials included in study<<**

**>> insert Table 2: Median pre-discussion scores and post-discussion PRECIS-2 scores for each trial<<**

*Only one set of pre-discussion scores were made for STOP, combining both the cluster and patient levels of the trial.

**>>insert Table 3: Median and interquartile range of the range of pre and post discussion scores <<**

## Discussion

We produced PRECIS-2 wheels for seven pragmatic trials of interventions in primary care. Most domains for each trial were scored by the teams as pragmatic indicating that the designs were appropriate to answer questions about the effectiveness of the interventions. Two domains, *Recruitment* and *Organisation* stood out as being rated as less pragmatic across four of the seven trials. This indicated that steps to recruit in these trials could impact on the applicability of their results and that resources over and above those required for the intervention itself might be needed to ensure successful roll out of interventions.

Two of the trials included in the study were at the design stage and there was the potential for changes to their design following the application of PRECIS-2. For one of these trials, TANDEM, the trial team made no changes following the application of the tool; the trial team were experienced in designing pragmatic trials and had already discussed the issues raised by applying PRECIS-2 prior to using the tool. The team commented that PRECIS-2 covered important areas of a trial's design that researchers should be aware of but that experienced teams may not need to go through the process of applying it fully. For the second trial, STOP, the design was yet to be finalised at the time of writing and the team found that PRECIS-2 helped clarify their thinking about their design. Of the completed trials, the researchers involved commented that the tool helped them think more about generalizability and had they applied it at the design stages could have pre-empted feedback they received from research funders on submitted applications.

The group discussions lead to a refinement rather than large change in median PRECIS-2 scores for each trial with the exception of one trial. Prior to the discussion agreement in scores, as measured by the range in scores for each domain, was generally poor and some raters were unable to rate all domains. The domains *Recruitment*, *Flexibility of adherence*, *Flexibility of delivery* and *Organisation* posed the greatest challenges to the raters when scoring the trials independently. Following the discussion the agreement in scores improved and raters were able to score all domains. The challenges encountered by investigators applying the original PRECIS tool (9): lack of information or understanding about the trial, usual care, or the PRECIS-2 domains, were mostly overcome. It was often the case that not one person in the trial team would have full information to score all domains and good understanding came only through sharing information. This suggests that greater attention may need to be given to the reporting of operational aspects of trials if PRECIS-2 is to be applied by external groups. Updating the SPIRIT (26) and CONSORT (10, 27) reporting guidelines for trial protocols and RCTs, to include more focus on how interventions are delivered and the resources and expertise used, could help assessment of whether the results of trials are applicable to a particular setting.

We identified one further challenge when applying PRECIS-2 to pragmatic trials in primary care. When an intervention significantly changed the way in which patients were treated it made comparisons with usual care difficult; especially for the *Organisation* domain, as it was not always clear what level of expertise or resourcing would be available to deliver the intervention if it were to be implemented in primary care post-trial. For example, in the COPERS trial, the intervention was delivered partly by lay people specifically recruited to deliver the intervention. This made judging the level of expertise and resources that would be available to deliver the intervention in usual care difficult as it was unclear what resources would be made available in primary care practice to recruit and train these individuals were the intervention to be rolled out following the trial.

At the time of writing there has been one other published account of using PRECIS-2 in which Johnson et al (28) applied PRECIS-2 to five pragmatic, cluster randomised trials in health care systems research set in the USA producing one PRECIS-2 wheel for each trial. There are similarities with our findings in that there were challenges applying PRECIS-2 to interventions which changed usual care in a significant way and that not all the information required to apply PRECIS-2 was available in the trial documents. A notable difference in findings was that Johnson et al encountered more difficulties applying PRECIS-2 possibly because these authors did not report holding a discussion before scoring. Johnson et al also did not have access to the full PRECIS-2 publication at the time of scoring (12) and had no involvement from the authors of PRECIS-2 in clarifying the definitions of the PRECIS-2 domains. Some of the trials in Johnson et al (28) may also have benefited from producing two PRECIS-2 wheels, one at the cluster level and one at the individual level.

The results of this study are based on a sample of trials undertaken in a clinical trials unit that specialises in pragmatic trials so different findings may arise from a unit with less of an expertise in pragmatic trials. In particular there may be greater disagreement in scores made independently by researchers with less experience designing pragmatic trials as there could be greater scope for differences in the interpretation of the PRECIS-2 domains or what it means for a trial to be pragmatic. The focus of this study was pragmatic trials carried out in primary care; different challenges may present when applying PRECIS-2 to trials with more explanatory designs or in other healthcare settings.

## Conclusions

Discussing PRECIS-2 wheels facilitated an exchange of information between different members of the trial team. It also highlighted two areas, recruitment and the level of resources used to deliver the intervention, where design decisions may be impacting the applicability of results of trials from the PCTU, and it focused discussions around generalisability. When considering cluster randomised trials or other more complex trial designs greater insight may be gained from plotting more than one PRECIS-2 wheel. To apply PRECIS-2 successfully we recommend holding a discussion between a group of people who between them have knowledge of what usual care is for the intervention, details of the trial's design including operational and technical aspects and an understanding of the PRECIS-2 domains.

## Funding statement

Some of the work described was part of Chief Scientist Office (CSO) grant CZH/4/773 and UK Medical Research Council and the University of Dundee through the provision of a PhD stipend for Kirsty Loudon, the University of Stirling and from the Health Services Research Unit at the University of Aberdeen, Nursing Midwifery Allied Health Professionals Research Unit at the University of Stirling which are both core funded by the CSO of the Scottish Government Health Directories.

Stephanie Taylor was (in part) supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) North Thames at Bart's Health NHS Trust. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

## References


1. Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. Trials. 2009;10:37.
2. Chalkidou K, Tunis S, Whicher D, Fowler R, Zwarenstein M. The role for pragmatic randomized controlled trials (pRCTs) in comparative effectiveness research. Clinical trials (London, England). 2012;9(4):436-46.
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 : the journal of the American Medical Association. 2003;290(12):1624-32.
4. Glasgow RE, Chambers D. Developing Robust, Sustainable, Implementation Systems Using Rigorous, Rapid and Relevant Science. Cts-Clin Transl Sci. 2012;5(1):48-55.
5. Zwarenstein M, Treweek S. What kind of randomised trials do patients and clinicians need? Evidence-based medicine. 2009;14(4):101-3.
6. Ford I, Norrie J. Pragmatic Trials. The New England journal of medicine. 2016;375(5):454-63.
7. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet. 2005;365(9453):82-93.
8. Gartlehner G, Hansen RA, Nissman D, Lohr KN, Carey TS. A simple and valid tool distinguished efficacy from effectiveness studies. Journal of clinical epidemiology. 2006;59(10):1040-8.
9. 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. Journal of clinical epidemiology. 2009;62(5):464-75.
10. 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 (Clinical research ed). 2008;337:a2390.
11. Johnson KE, Tachibana C, Coronado GD, Dember LM, Glasgow RE, Huang SS, et al. A guide to research partnerships for pragmatic clinical trials. BMJ (Clinical research ed). 2014;349:g6826.
12. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ (Clinical research ed). 2015;350:h2147.
13. 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 medical research methodology. 2012;12:101.
14. Riddle DL, Johnson RE, Jensen MP, Keefe FJ, Kroenke K, Bair MJ, et al. The Pragmatic-Explanatory Continuum Indicator Summary (PRECIS) instrument was useful for refining a randomized trial design: experiences from an investigative team. Journal of clinical epidemiology. 2010;63(11):1271-5.
15. Koppenaal T, Linmans J, Knottnerus JA, Spigt M. Pragmatic vs. explanatory: an adaptation of the PRECIS tool helps to judge the applicability of systematic reviews for daily practice. Journal of clinical epidemiology. 2011;64(10):1095-101.
16. Bratton DJ, Nunn AJ, Wojnarowska F, Kirtschig G, Sandell A, Williams HC. The value of the pragmatic-explanatory continuum indicator summary wheel in an ongoing study: the bullous pemphigoid steroids and tetracyclines study. Trials. 2012;13:50.
17. Witt CM, Manheimer E, Hammerschlag R, Ludtke R, Lao L, Tunis SR, et al. How well do randomized trials inform decision making: systematic review using comparative effectiveness research measures on acupuncture for back pain. PloS one. 2012;7(2):e32399.
18. Glasgow RE, Gaglio B, Bennett G, Jerome GJ, Yeh HC, Sarwer DB, et al. Applying the PRECIS criteria to describe three effectiveness trials of weight loss in obese patients with comorbid conditions. Health services research. 2012;47(3 Pt 1):1051-67.
19. Krist AH, Glenn BA, Glasgow RE, Balasubramanian BA, Chambers DA, Fernandez ME, et al. Designing a valid randomized pragmatic primary care implementation trial: the my own health report (MOHR) project. Implementation science : IS. 2013;8:73.

20. 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 in clinical neuroscience. 2011;13(2):209-15.
21. Harden SM, Burke SM, Haile AM, Estabrooks PA. Generalizing the Findings From Group Dynamics-Based Physical Activity Research to Practice Settings: What Do We know? Evaluation & the health professions. 2013.
22. Eldridge S. Pragmatic trials in primary health care: what, when and how? Family practice. 2010;27(6):591-2.
23. Carnes D, Taylor SJ, Homer K, Eldridge S, Bremner S, Pincus T, et al. Effectiveness and cost-effectiveness of a novel, group self-management course for adults with chronic musculoskeletal pain: study protocol for a multicentre, randomised controlled trial (COPERS). BMJ open. 2013;3(1).
24. Leber W, McMullen H, Anderson J, Marlin N, Santos AC, Bremner S, et al. Promotion of rapid testing for HIV in primary care (RHIVA2): a cluster-randomised controlled trial. The lancet HIV. 2015;2(6):e229-35.
25. Nwokoro C, Pandya H, Turner S, Eldridge S, Griffiths CJ, Vulliamy T, et al. Intermittent montelukast in children aged 10 months to 5 years with wheeze (WAIT trial): a multicentre, randomised, placebo-controlled trial. The Lancet Respiratory medicine. 2014;2(10):796-803.
26. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. Annals of internal medicine. 2013;158(3):200-7.
27. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. PLoS Med. 2010;7(3):e1000251.
28. Johnson KE, Neta G, Dember LM, Coronado GD, Suls J, Chambers DA, et al. Use of PRECIS ratings in the National Institutes of Health (NIH) Health Care Systems Research Collaboratory. Trials. 2016;17(1):32.

**Table 1: Description of trials included in study**

| Short Title | Full Title | Intervention & Setting | Unit of Randomisation | Status at time of study | No. in rating group |
|---|---|---|---|---|---|
| COPERS (23) | Effectiveness and cost-effectiveness of a novel, group self-management course for adults with chronic musculoskeletal pain | Group self-management course for chronic pain patients. Groups facilitated by a healthcare professional and lay person with experience of chronic pain. Trial set in local medical or community venues in urban and rural areas of the UK. | Individual | Complete | 5 (CI, trial manager, TSC member, trial statistician, GF). |
| HepFree | Chronic viral hepatitis in first and second generation immigrants from 'at risk' countries. A controlled randomised cross sectional cluster trial to assess the impact of identifying, screening and treating immigrants with viral hepatitis | Two interventions tested: 1) targeted screening for viral hepatitis for first and second generation immigrants from at risk countries 2) involvement of community therapy following diagnosis. Trial carried out across two cities in the UK with large immigrant populations (London & Bradford) | GP Surgery | Recruiting | 5 ((CI, trial manager, TSC member, trial statistician, GF). |
| Rhiva 2 (24) | Promotion of rapid testing for HIV in primary care: a cluster randomised trial | Intervention involved an education programme promoting rapid HIV testing to adults newly registering at GP surgeries. Trial carried out in primary care settings in London, UK. | GP Surgery | Complete | 4 (CI, trial manager, trial statistician, GF).. |
| STOP | Optimising pharmacist-based treatment from smoking cessation | Educational intervention package delivered to pharmacy staff. Pharmacies recruited to trial from those offering NHS smoking cessation services in two London boroughs. | Pharmacy | Protocol writing | 4 (CI, co-applicant, trial statistician, GF). |
| SWAP | A peer-support weight action programme to supplement brief advice in general practice | Group health behaviour modification intervention providing participants with the tools to lose weight and maintain a healthy lifestyle including pedometer use. Groups delivered by two trained advisors in community settings in London, UK. | Individual | Recruiting | 5 (CI, trial manager, TSC member, trial statistician, GF). |
| Tandem | Tailored intervention for anxiety and depression Management in chronic obstructive pulmonary disease. | Psychological intervention based on cognitive behavioural approach to proceed routine pulmonary rehabilitation. Delivered 1-1 by respiratory nurses or allied health professionals. Intervention delivered in participant's home or convenient NHS facility. Trial run from sites across the UK. | Individual | Submitted for funding | 4 (3x co-applicant, GF). |

| | | | | | |
|---|---|---|---|---|---|
| WAIT (25) | Intermittent montelukast in children aged 10 months to 5 years with wheeze: a multicentre, randomised, placebo-controlled trial | Intermittent Montelukast given by parents at every viral cold or wheezing episode. Patients recruited via primary and secondary care sites across the UK and intervention delivered at participant's home. | Individual | Complete | 5 (CI, trial manager, TSC member, trial statistician, GF). |

**Table 2: Median pre-discussion scores and post-discussion PRECIS-2 scores for each trial**

| | COPERS | | HepFree | | Rhiva 2 | | STOP Cluster | | STOP Patient | | SWAP | | TANDEM | | WAIT | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pre | Post | Pre | Post | Pre | Post | Pre* | Post | Pre* | Post | Pre | Post | Pre | Post | Pre | Post |
| Eligibility | 4.5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 3 | 4 | 4 | 4 | 4 |
| Recruitment | 4 | 3 | 4 | 4 | 4.5 | 5 | 3 | 4.3 | 3 | 5 | 3.5 | 3 | 3.5 | 2.5 | 3 | 3 |
| Setting | 3 | 4 | 4.5 | 4 | 5 | 4 | 4.5 | 4.5 | 4.5 | 5 | 5 | 4 | 5 | 5 | 4 | 5 |
| Organisation | 2.5 | 2 | 4 | 3 | 3 | 4 | 2.5 | 4 | 2.5 | 4.5 | 4 | 3 | 2 | 2 | 5 | 5 |
| Flexibility of delivery | 3.5 | 4 | 3 | 4 | 4 | 4 | 2 | 4 | 2 | | 4 | 4 | 2 | 2 | 4 | 4.5 |
| Flexibility of adherence | 5 | 5 | 5 | 4 | 4 | 4 | 1 | 5 | 1 | 5 | 4.5 | 4 | 5 | 3.5 | 4 | 3 |
| Follow up | 4 | 4 | 4.5 | 4 | 4.5 | 5 | 2 | 5 | 2 | 3 | 3 | 2 | 4 | 4 | 2 | 1 |
| Primary outcome | 4 | 4 | 4 | 5 | 4.5 | 5 | 5 | 3.5 | 5 | | 4.5 | 4 | 5 | 5 | 5 | 5 |
| Primary analysis | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | | 5 | 5 | 5 | 5 | 5 | 5 |

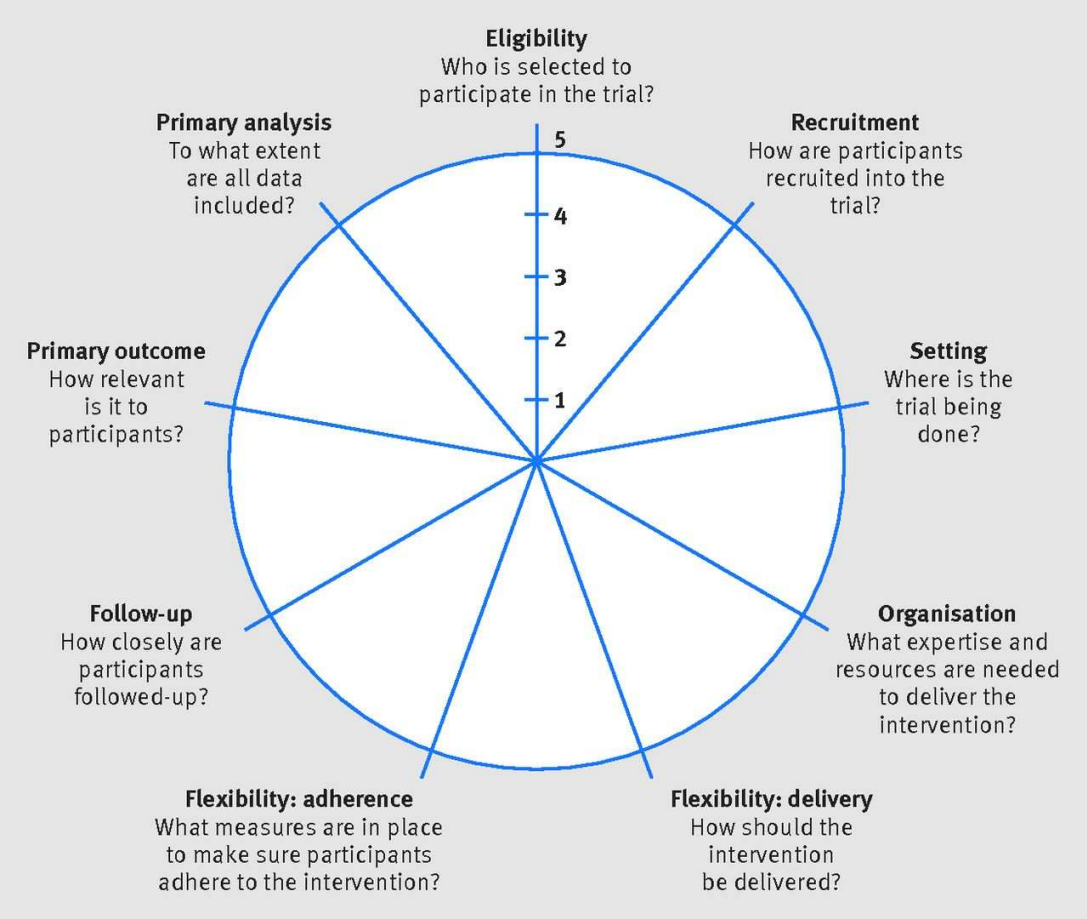
*Only one set of pre-discussion scores were made for STOP, combining both the cluster and patient levels of the trial.

**Table 3: Median and interquartile range of the range of pre and post discussion scores**

| Domain | Range of pre-discussion scores, | Range of post-discussion scores, | Raters unable to score domain pre-discussion, |
|---|---|---|---|

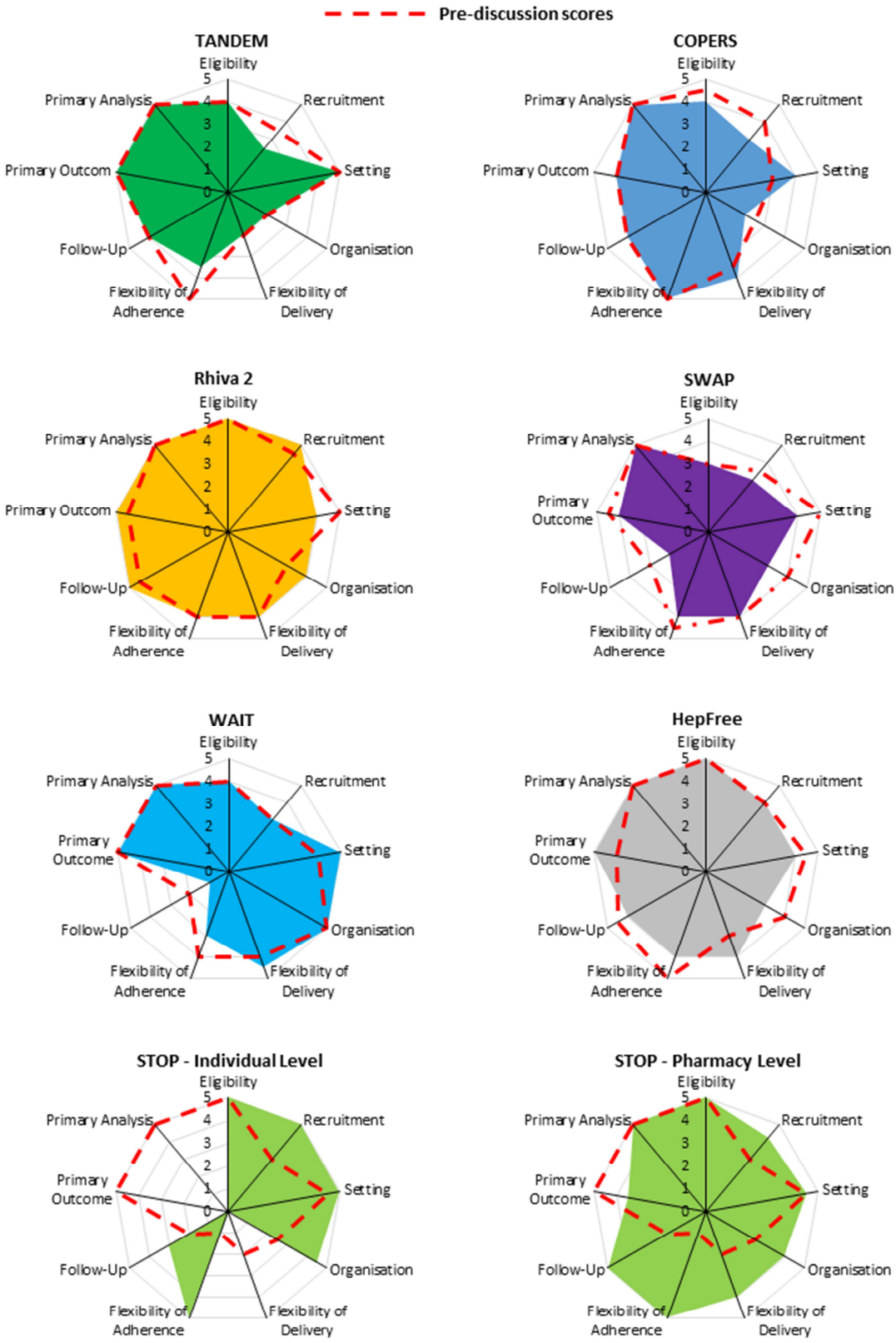
| | median (interquartile range) | median (interquartile range) | n (%) |
|---|---|---|---|
| Eligibility | 1 (0, 2) | 0 (0, .5) | 1 (4%) |
| Recruitment | 3 (2, 4) | 1 (0, 1) | 0 (0%) |
| Setting | 1 (0, 2) | 1 (.5, 1) | 1 (4%) |
| Organisation | 2 (2, 2) | 1 (1, 1) | 3 (11%) |
| Flexibility of delivery | 2 (1, 4) | 1 (0, 1) | 4 (14%) |
| Flexibility of adherence | 2 (0, 3) | 0.5 (0, 1.5) | 6 (21%) |
| Follow up | 2 (1, 3) | 0 (0, 1.5) | 0 (0%) |
| Primary outcome | 1 (0, 2) | 1 (0, 2) | 0 (0%) |
| Primary analysis | 0 (0, 1) | 0 (0, 0) | 1 (4%) |

**Figure 1: A blank PRECIS-2 Wheel*.**




*Reproduced with permission from (12)

**Figure 2: PRECIS-2 wheels for trials included in the study. The solid shapes indicate the median post discussion scores. Dashed lines show the median pre-discussion scores.**


Pre-discussion scores
TANDEM
COPERS
Rhiva 2
SWAP
WAIT
HepFree
STOP - Individual Level
STOP - Pharmacy Level
Eligibility
Recruitment
Setting
Organisation
Flexibility of Delivery
Flexibility of Adherence
Follow-Up
Primary Outcom
Primary Outcome
Primary Analysis
0
1
2
3
4
5

# Applying PRECIS-2 to pragmatic trials in primary care.

Authors: Gordon Forbes[1*], Kirsty Loudon[2], Shaun Treweek[3], Stephanie JC Taylor[1], Sandra Eldridge[1]

[1]Queen Mary, University of London [2]University of Stirling [3]University of Aberdeen

## Conflict of interest declaration

The authors of this study have no conflicts of interest to declare

## What is new?

- PRECIS-2 offers a new tool for trial teams to exchange information and knowledge about trial design.
- To score each PRECIS-2 domain successfully understanding is required of what usual care would be for the intervention, what is planned for the trial, and the interpretation of each domain in the context of the trial; a discussion amongst a multidisciplinary group may be required to score each domain.
- The trials included in this study were broadly pragmatic, areas where they were commonly less pragmatic were recruitment, and organisation (the level of resources used to deliver the intervention).
- When applying PRECIS-2 to cluster randomised trials greater insight may sometimes be gained by plotting two wheels, one at the level of the cluster and another at the level of individual participants