ibtn international behavioural trials network

DFID Department for International Development

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL NCRM National Centre for Research Methods

UNIVERSITY OF OXFORD

Professor of Evidence-Based Social Intervention and Policy Evaluation

Paul Montgomery

15 May 2016

This Talk

Overview of 3 Projects

- SPIRIT
- CONSORT–SPI
- GRADE–CI

Applied Behavioral/Social Scientists Live in Exciting Times

- UK What Works Network
  - Create, share, and use high-quality evidence on policy programs and practices which combined receive public spending of more than £200 billion
  - Public health, social care, education, crime reduction, and economic growth
- US Social and Behavioral Science Team
  - Assists federal agencies in applying behavioural science insights to policies and operations
  - Improve public welfare, programme outcomes, and cost effectiveness

Obama's Executive Order in Sept 2015

The White House

Office of the Press Secretary

For Immediate Release — September 15, 2015

Executive Order -- Using Behavioral Science Insights to Better Serve the American People

EXECUTIVE ORDER

- - - - - - -

USING BEHAVIORAL SCIENCE INSIGHTS TO

BETTER SERVE THE AMERICAN PEOPLE

A growing body of evidence demonstrates that behavioral science insights -- research findings from fields such as behavioral economics and psychology about how people make decisions and act on them -- can be used to design government policies to better serve the American people.

https://www.whitehouse.gov/the-press-office/2015/09/15/executive-order-using-behavioral-science-insights-better-serve-american

Most Published Research May be False

- We have problems reproducing psychological science (Open Science Collaboration, 2015)
  - Replications of 100 experimental and correlational studies
  - 97% of original studies vs 36% of replications had significant results
  - Mean effect size was half the magnitude of the mean effect size of the original effects
  - “There is still more work to do to verify whether we know what we think we know”

ibtn international behavioural trials network

DFID Department for International Development

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL NCRM National Centre for Research Methods

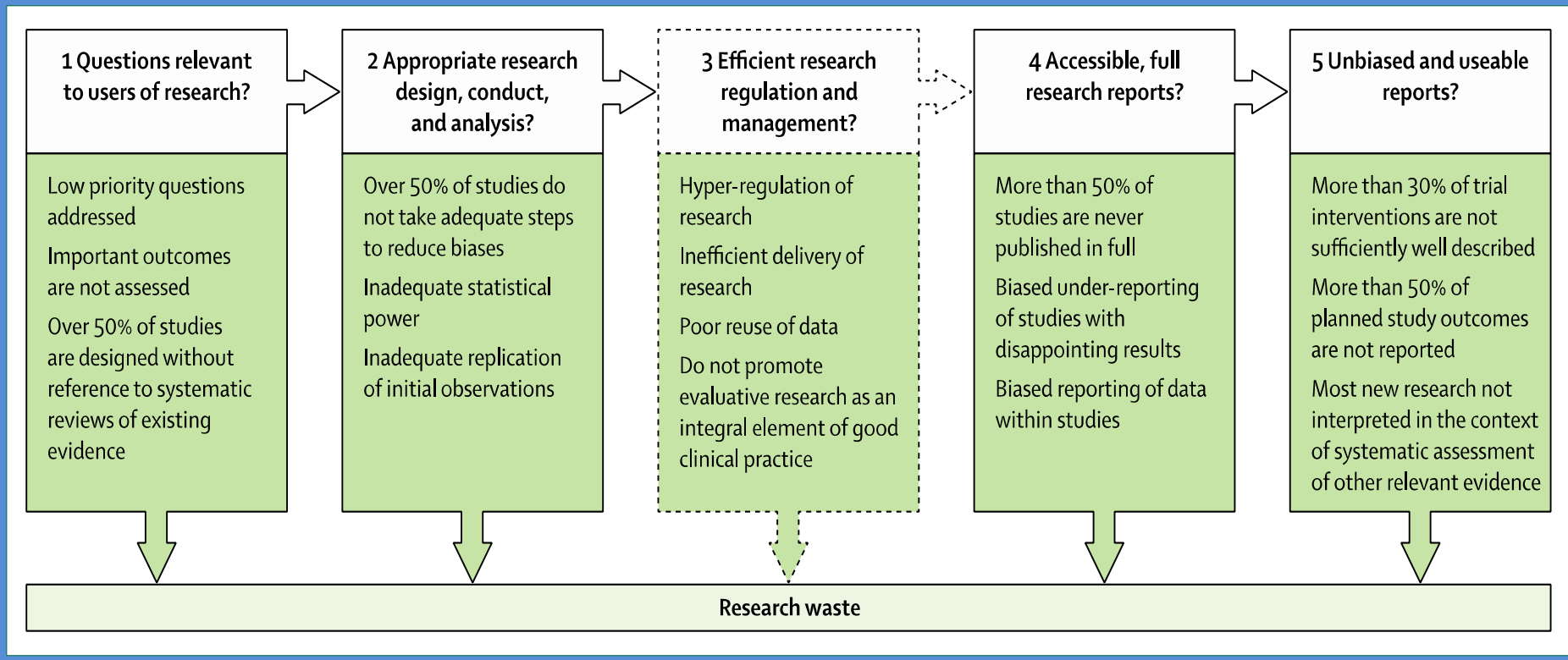
UNIVERSITY OF OXFORD

Why Most Published Research Findings are False

- A research finding is less likely to be true (Ioannidis 2005):
  - when the studies are smaller
  - when effect sizes are smaller
  - when there is a greater number and lesser pre-selection of tested relationships
  - where there is greater flexibility in designs, definitions, outcomes, and analytical modes
  - when there is greater financial and other interest and prejudice
  - when more teams are involved in chase of statistical significance

85% of Biomedical Research Funding ($210 Billion) Is Being Avoidably Wasted

Several stages of research production may lead to waste (Moher 2015)



How to Make More Published Research True

- Some research practices that may help increase the proportion of true research findings (Ioannidis 2014):
  - Large-scale collaborative research
  - Adoption of replication culture and reproducibility practices
  - Registration (studies, protocols, analysis codes, datasets, raw data)
  - Sharing (data, protocols, materials, software, and other tools)
  - Containment of conflicted sponsors and authors
  - More appropriate statistical methods
  - Standardization of definitions and analyses
  - More stringent thresholds for claiming discoveries or "successes"
  - Improvement of study design standards
  - Improvements in peer review, reporting, and dissemination of research
  - Better training of scientific workforce in methods and statistical literacy

Research Transparency Can Increase Value, Reduce Waste

- *The Lancet* REWARD (REduce research Waste And Reward Diligence) Campaign
- Center for Open Science to improve openness and integrity of scientific practices
  - Open Science Framework for transparent, cloud-based management of scientific projects
  - Transparency and Openness Promotion (TOP) Guidelines (Nosek 2015)
- Berkeley Initiative for Transparency in the Social Sciences (BITSS)
- Data Access and Research Transparency (DA-RT) Statement for social scientists
- Meta-Research Innovation Center at Stanford (METRICS)
- Laura and John Arnold Foundation (LJAF) Research Integrity Grants (over $80 million since 2012)
- **Enhancing the Quality and Transparency of Health Research (EQUATOR) Network reporting guidelines**

Standard Protocol Items: Recommendations for Intervention Trials
The SPIRIT Statement

http://www.spirit-statement.org/spirit-statement/

Definition of a Trial Protocol

A document that provides sufficient detail to enable understanding of the background, rationale, objectives, study population, interventions, methods, statistical analyses, ethical considerations, dissemination plans and administration of the trial

- **A cohesive document**
- **Provides appropriate context and narrative of the trial elements**
- **Enables replication of an intervention**

The SPIRIT Statement (2013)

- Guidance for the minimum *protocol content* of an intervention trial
- Promotes transparency and a full description of what is planned
- Does not prescribe how to design or conduct a trial
- 33-item checklist

ibtn international behavioural trials network

DFID Department for International Development

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL NCRM National Centre for Research Methods

UNIVERSITY OF OXFORD

The SPIRIT Statement (2013)

RESEARCH AND REPORTING METHODS | Annals of Internal Medicine

SPIRIT 2013 Statement: Definir Clinical Trials

An-Wen Chan, MD, DPhil; Jennifer M. Tetzlaff, MSc; Dougla
Karmela Krleža-Jerić, MD, DSc; Asbjørn Hróbjartsson, PhD; H
Caroline J. Doré, BSc; Wendy R. Parulekar, MD; William S.M.
Harold C. Sox, MD; Frank W. Rockhold, PhD; Drummond Rer

The protocol of a clinical trial serves as the foundation for planning, conduct, reporting, and appraisal. However, trial pro and existing protocol guidelines vary greatly in content and c This article describes the systematic development and scc SPIRIT (Standard Protocol Items: Recommendations for Int tional Trials) 2013, a guideline for the minimum content of a trial protocol.

The 33-item SPIRIT checklist applies to protocols for all trials and focuses on content rather than format. The ch recommends a full description of what is planned; it do prescribe how to design or conduct a trial. By providing gu

SPIRIT

STANDARD PROTOCOL ITEMS: RECOMMENDATIONS FOR INTERVENTIONAL TRIALS

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description
Administrative information		
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry
	2b	All items from the World Health Organization Trial Registration Data Set
Protocol version	3	Date and version identifier
Funding	4	Sources and types of financial, material, and other support
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors
	5b	Name and contact information for the trial sponsor
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)

Example Template of Recommended Content for the Enrolment and Interventions

	STUDY PERIOD							
	Enrolment	Allocation	Post-allocation					Close-out
TIMEPOINT**	$-t_1$	0	t_1	t_2	t_3	t_4	etc.	t_x
ENROLMENT:								
Eligibility screen	X							
Informed consent	X							
[List other procedures]	X							
Allocation		X						
INTERVENTIONS:								
[Intervention A]			Schedule of enrolment, interventions, and assessments					
[Intervention B]			X		X			
[List other study groups]								

See SPIRIT 2013 Explanation and Elaboration for examples from protocols

Main Publications

- Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin J, Doré C, Parulekar W, Summerskill W, Groves T, Schulz K, Sox H, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. Ann Intern Med 2013;158:200-207.

- Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin J, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D. SPIRIT 2013 Explanation and Elaboration: Guidance for protocols of clinical trials. BMJ 2013;346:e7586

- A comment about SPIRIT 2013 has also been published in the Lancet: **http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2812%2962160-6/fulltext?rss=yes**

A New Reporting Guideline for Trials of Social and Psychological Interventions: CONSORT-SPI

CONSORT
TRANSPARENT REPORTING of TRIALS

http://www.spi.ox.ac.uk/research/site/consort-spi/home.html

CONSORT Initiative: Goals

- Emphasising the importance of research transparency
- Highlighting the need to use reporting guidelines for all future research
- Promoting use of research transparency tools to colleagues and grantees

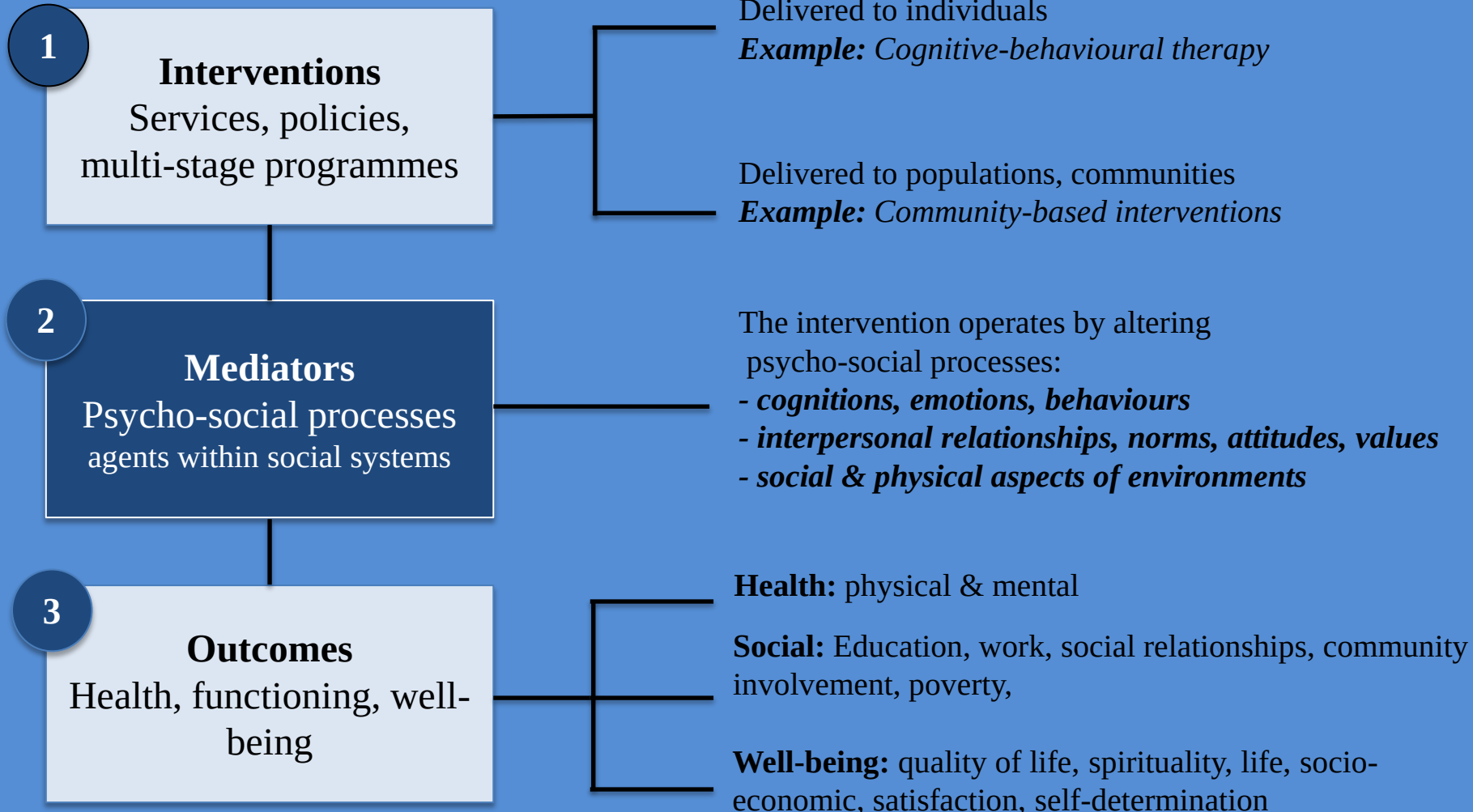
CONSORT-SPI: Objectives

- Social and psychological intervention RCTs
- Reporting Guidelines & CONSORT
- Developing CONSORT-SPI
- The CONSORT-SPI Checklist

What is an “Intervention”?

- The ***action*** of intervening, “stepping in” or interfering in any affair, so as to affect its ***course or issue***
  (Oxford English Dictionary)
- The ***act*** or fact of becoming involved intentionally
  (Cambridge English Dictionary)
- The ***act*** or … a method of interfering with the ***outcome or course*** especially of a condition or process
  (Merriam-Webster Dictionary)
- An ***action*** that aims to bring about identifiable ***outcomes***
  (Rychetnik et al., 2004; A glossary for evidence-based public health. JECH; 2004;58:538-545)
- Intentional change strategies (delivered at different levels)
  (Fraser M et al., Intervention Research. 2009. Oxford University Press)

Psychosocial Interventions



Adapted from: Grant et al. (2014). Development of CONSORT-SPI.
England et al. (2015). Psycho-social interventions for mental health and substance use disorders. IOM (Institute of Medicine).

Social Practices, Programs, and Policies are Interventions

- **"Practices":** the materials and activities through which better quality of life is enabled (e.g., coaching, mentoring, parenting, peer interactions, teaching)
  - Practices involve direct interaction with participants (though not necessarily in person)
- **"Programs":** coordinated sets of activities designed to achieve specific aims
  - Getting To Outcomes© (GTO) and ECHO© (Extension for Community Healthcare Outcomes)
- **"Policies":** broader initiatives intended to promote success through the allocation of resources or regulation of activities
  - Policies may be located at the federal, state, local, or organizational level

Adapted from William T Grant Foundation "Research Grants Application Guide"

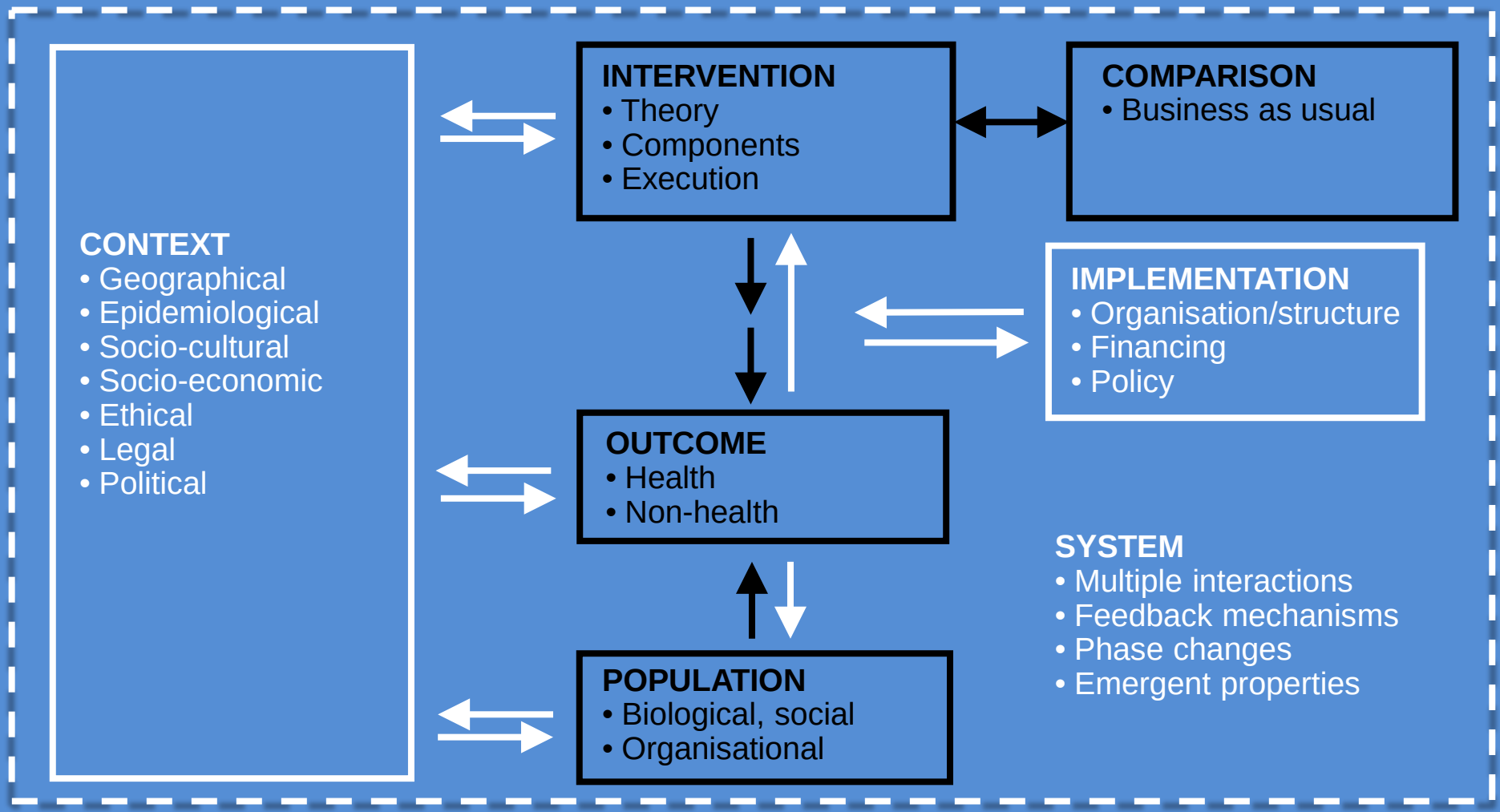
Complex Interventions: UK MRC Framework

- Number of and *interactions between components* within the experimental and control interventions
- Number and *difficulty of behaviours* required by those delivering or receiving the interventions
- Number of *groups or organisational levels targeted* by the intervention
- Degree of *flexibility or tailoring of the intervention* permitted

Craig P. et al. (2008). Developing and evaluating complex interventions. New guidance. Medical Research Council

Interventions in Complex Adaptive Systems

(slide taken from Eva Rehfuess)



Multi-Systemic Therapy

- Intensive intervention for chronic juvenile offenders
- Therapists, caseworkers, psychologists, psychiatrists
- Work with individual, family, peers, and neighbourhood
- Settings: home, school, community
- Services may focus on cognition and behaviour change, communication skills, parenting skills, family relations, peer relations, school performance, or social networks
- Tailored to the specific needs of the youth and family

Good RCT Reporting Includes…

- Participant and setting characteristics
- Interventions and their implementation
- Outcome assessment
- Theories informing the study
- Trial design

The Problem: Poor Reporting

Reporting Guidelines

- Minimum set of items on article content
- Reflect issues related to bias
- Based on evidence and consensus

ibtn international behavioural trials network

DFID Department for International Development

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL | NCRM National Centre for Research Methods

UNIVERSITY OF OXFORD

equator network

Enhancing the QUAlity and Transparency Of health Research

Visit the EQUATOR Spanish Website

The resource centre for good reporting of health research studies

Library for health research reporting

The Library contains a comprehensive searchable database of reporting guidelines and also links to other resources relevant to research reporting.

Search for reporting guidelines

Visit the library for more resources

Key reporting guidelines

CONSORT	Full Record | Checklist | Flow Diagram
STARD	Full Record | Checklist | Flow Diagram
STROBE	Full Record | Checklist
PRISMA	Full Record | Checklist | Flow Diagram
COREQ	Full Record
ENTREQ	Full Record
SQUIRE	Full Record | Checklist
CHEERS	Full Record

Toolkits

The EQUATOR Network works to improve the reliability and value of medical research literature by promoting transparent and accurate

EQUATOR highlights

9/08/2013 - EQUATOR Network at the Peer Review Congress 2013 in Chicago

EQUATOR will be present at the Seventh International Congress on Peer Review and Biomedical Publication, 8-10 September 2013. We are

News

The New ICMJE Recommendations
29/08/2013

Better Reporting of Scientific Studies: Why It Matters

Various stakeholders can benefit from the adoption of reporting guidelines

- Researchers: study design and final report
- Editors and peer-reviewers: improve manuscripts
- Research funders: improve submissions and utility of funded projects
- Policy-makers and practitioners: promoting RGs could lead to publications they can use
- Faculty: education and training of next generation of researchers

ibtn international behavioural trials network

DFID Department for International Development

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL NCRM National Centre for Research Methods

UNIVERSITY OF OXFORD

Our Case: The CONSORT STATEMENT

RESEARCH METHODS & REPORTING

CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials

Kenneth F Schulz,[1] Dou

EDITORIAL by Antes
RESEARCH, p 697

[1]Family Health International, Research Triangle Park, NC 27709, USA
[2]Centre for Statistics in Medicine, University of Oxford, Wolfson College, Oxford
[3]Ottawa Methods Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada
Correspondence to: K F Schulz kschulz@fhi.org

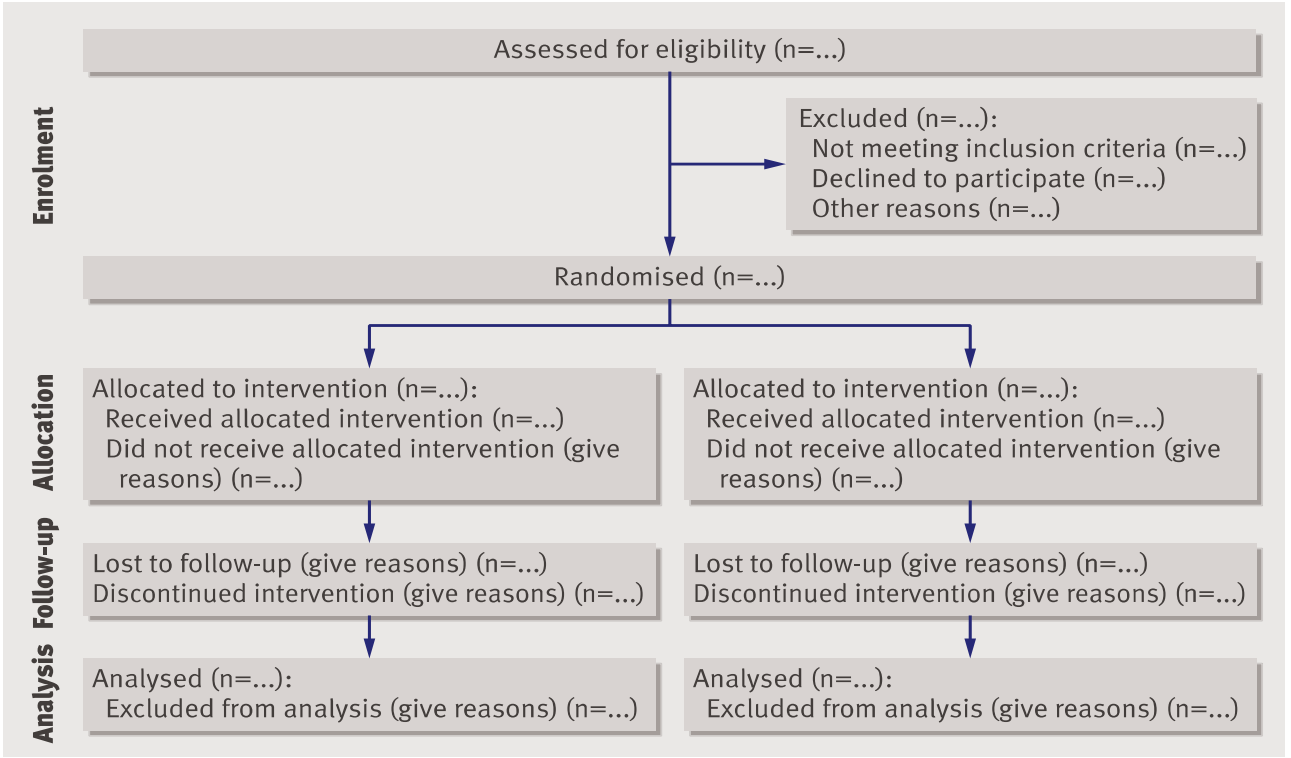
Accepted: 9 December 2009

Cite this as: *BMJ* 2010;340:c332
doi: 10.1136/bmj.c332

The CONSORT state
to improve the repor
controlled trials. **Ken**
colleagues describe
CONSORT 2010, wh
guideline based on r
evidence and accum

Randomised controlled tri
conducted, and reported, r
uating healthcare interven
can yield biased results if th
assess a trial accurately, re
complete, clear, and transp
ology and findings. Unfor
frequently fail because aut
to provide lucid and comp
information.[2-4]

That lack of adequate rep
the original CONSORT (Cor
Trials) statement in 1996[5]
While those statements in
some randomised controll
remain inadequate.[2] Furth



Flow diagram of the progress through the phases of a parallel randomised trial of two groups (that is, enrolment, intervention allocation, follow-up, and data analysis)

ibtn
international
behavioural
trials network

DFID Department for International Development

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL NCRM National Centre for Research Methods

UNIVERSITY OF OXFORD

CONSORT Extensions

RESEARCH ME

Consort 2010 statement
randomised tri

The Consolidated Standard
the reporting of randomisec
reporting of parallel group r
further update in 2010. A sep
in 2008. In earlier papers w
statement for the reporting o
guidance, based on the 201(
for the reporting of abstract

Marion K Campbell *director*
healthcare evaluation[2], Dou

[1]Health Services Research Unit, Univers
and Tropical Medicine, London, UK; [3]Cer

CONSORT for reporting randomised trials in journal and conference abstracts

Published Online
January 22, 2008
DOI:10.1016/S0140-6736(07)61835-2

In 2006, Arthur Amman, President of Global Strategies for HIV Prevention, made a disquieting remark: "I recently met a physician from southern Africa, engaged in perinatal HIV prevention, whose primary access to information was abstracts posted on the internet. Based on a single abstract, they had altered their perinatal HIV prevention program from an effective therapy to one with lesser efficacy. Had they read the full text article they would have undoubtedly realized that the study results were based on short-term follow-up, a small pivotal group, incomplete data, and unlikely to be applicable to their country situation. Their decision to alter treatment based solely on the abstract's conclusions may have

Yet a study that examined 35 journals' instructions for authors found that only 4% of the text was devoted to the content or format of the abstract.[6] When key details about a trial are lacking, it is difficult to assess the validity of the results and their applicability.

In collaboration with members of the CONSORT Group, we have extended the current CONSORT Statement to develop a checklist of essential items which authors should include when reporting the main results of a randomised trial in a journal or conference abstract. We recognise that many journals have developed their own structure for reporting abstracts. Our intention is not to suggest changes to these formats, but to recommend

Annals of Internal Medicine | ACADEMIA AND CLINIC

Extending the CONSORT Statement to Randomized Trials of Nonpharmacologic Treatment: Explanation and Elaboration

Isabelle Boutron, MD, PhD; David Moher, PhD; Douglas G. Altman, DSc; Kenneth F. Schulz, PhD, MBA; and Philippe Ravaud, MD, PhD, for the CONSORT Group*

Adequate reporting of randomized, controlled trials (RCTs) is necessary to allow accurate critical appraisal of the validity and applicability of the results. The CONSORT (Consolidated Standards of Reporting Trials) Statement, a 22-item checklist and flow diagram, is intended to address this problem by improving the reporting of RCTs. However, some specific issues that apply to trials of nonpharmacologic treatments (for example, surgery, technical interventions, devices, rehabilitation, psychotherapy, and behavioral intervention) are not specifically addressed in the CONSORT Statement. Furthermore, considerable evidence suggests that the reporting of nonpharmacologic trials still needs improvement. Therefore, the CONSORT group developed an extension of the CONSORT Statement for trials assessing nonpharmacologic treatments. A consensus meeting of 33 experts was organized in Paris, France, in February 2006, to develop an extension of the CONSORT Statement for trials of nonpharmacologic treatments. The participants extended 11 items from the CONSORT Statement, added 1 item, and developed a modified flow diagram.

To allow adequate understanding and implementation of the CONSORT extension, the CONSORT group developed this elaboration and explanation document from a review of the literature to provide examples of adequate reporting. This extension, in conjunction with the main CONSORT Statement and other CONSORT extensions, should help to improve the reporting of RCTs performed in this field.

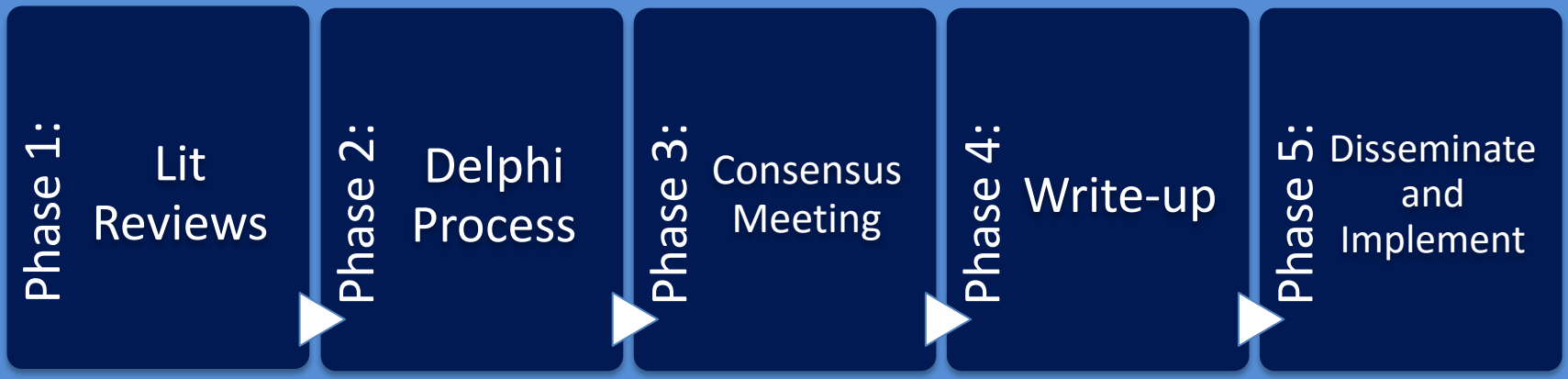
Ann Intern Med. 2008;148:295-309. www.annals.org

For author affiliations, see end of text.

*For contributors to the CONSORT Extension for Nonpharmacologic Treatment Interventions, see the **Appendix** (available at www.annals.org).

CONSORT-SPI Project

- Official CONSORT Extension
- Rigorous consensus development
- Multi-pronged dissemination strategy



Project Executive

- Paul Montgomery, University of Oxford
- Evan Mayo-Wilson, Johns Hopkins University
- Sean Grant, University of Oxford
- Geraldine Macdonald, Queen's University Belfast
- Sally Hopewell, University of Oxford
- Susan Michie, University College London
- David Moher, Ottawa Health Research Institute

International Advisory Group

- J Lawrence Aber
- Chris Bonell
- David Clark
- Frances Gardner
- Steve Hollon
- Jim McCambridge
- Laurence Moore
- Mark Petticrew
- Steve Pilling
- Lawrence Sherman
- James Thomas
- Elizabeth Waters
- David Weisburd
- Jo Yaffe

Phase 1:
Largest Review Ever on Topic

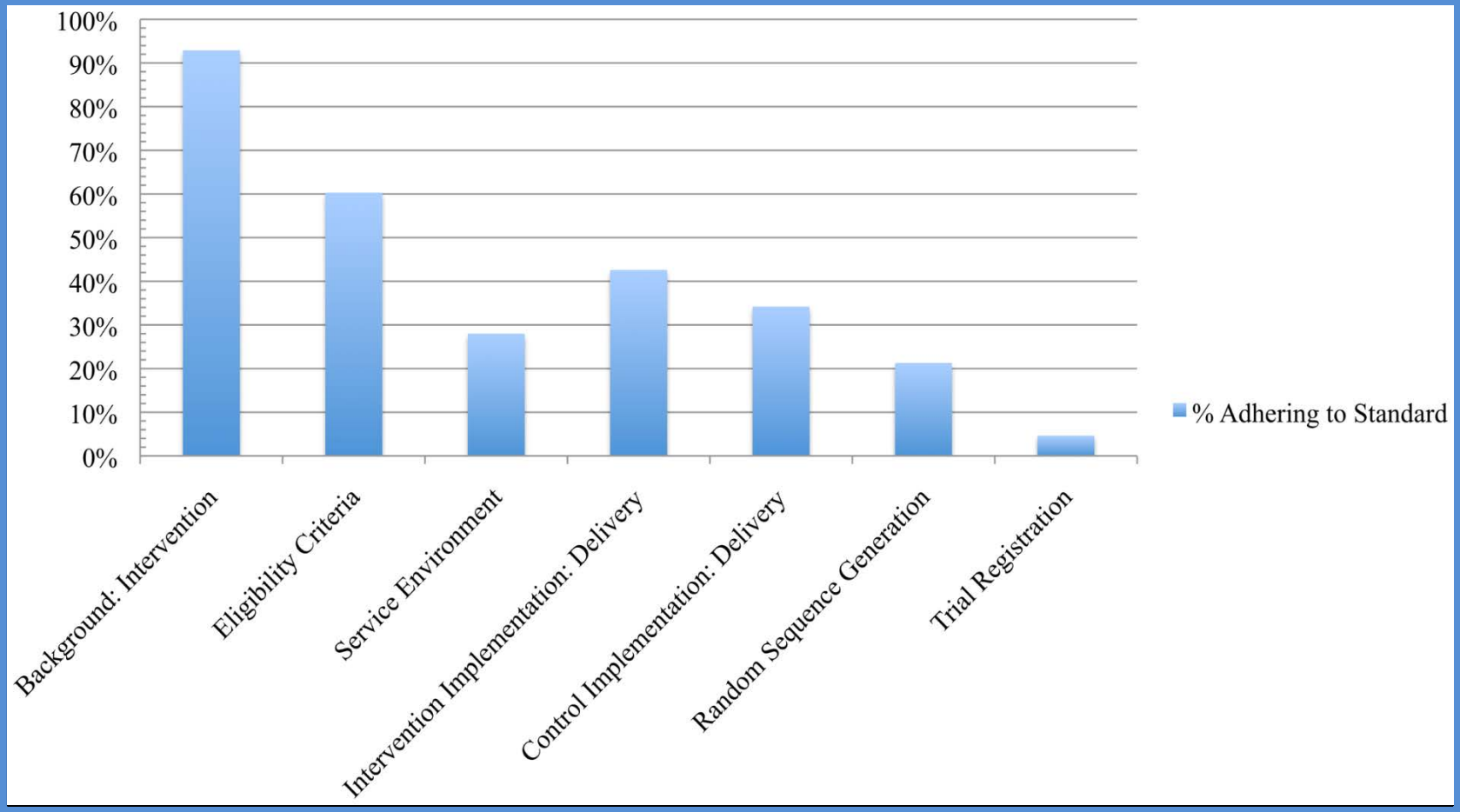
- 19 reporting guidelines with 147 reporting standards
  - 6 developed by CONSORT Group
  - 6 for biomedical trials in
  - 7 for social and behavioural sciences (public health, education, psychology, criminal justice, substance use, occupational therapy, behavioural change)
- 40 journals publishing 239 RCTs in 2010
  - Clinical Psychology (99 RCTs), Crime & Justice (31), Education (89), Social Work (20)

Phase 1:
Largest Review Ever on Topic

- Social/behavioural science guidelines developed/disseminated with less rigour
- 89 new/modified reporting standards compared to CONSORT guidelines
- 239 RCTs report <50% of standards on average

Ref: Grant et al (2013). PLoS One, 8(5), e65442

Average Compliance of RCTs with Key Reporting Standards



Ref: Grant et al (2013). PLoS One, 8(5), e65442

Phase 2:
Largest RG Delphi Process

- N = 384 (32 countries total)
  - 355 (92%) identified as an academic or researcher
  - 110 (29%) as practitioners/providers of social and psychological interventions
  - 132 (34%) as journal editors
  - 47 (12%) holding positions funding research
  - 36 (9%) involved in policy-making
  - 21 (6%) as recipients of interventions

Phase 2:
Largest RG Delphi Process

- 58 items recommended for inclusion
  - All but 1 of CONSORT 2010 checklist items (registration)
- Substantive qualitative feedback for consensus meeting and E&E

Phase 3: Consensus Meeting

- 31 participants from Delphi process
- 9 extended CONSORT 2010 items
  - 14 “sub-items” in total
- Other “Delphi” items discussed in E&E

New/Adapted Items

- Intervention theory of change
- Eligibility criteria for settings and providers
- Intervention/comparator delivery and uptake
- Intervention materials (e.g., manual, website)
- How missing data were handled
- Number approached, screened, and eligible

New/Adapted Items

- Socioeconomic baseline variables
- Availability of trial data
- Other potential interests than funder
- Involvement of the intervention developer
- Other stakeholder involvement
- Incentives offered as part of the trial

Phase 4: Write-Up

- Official guideline extension
  - *Draft checklist in appendix of this PPT
- Tailored E&E documents to disciplines
  - Rationale for each item
  - Examples of good reporting

Phase 5: Dissemination

- Simultaneous co-publication
- Journal endorsement and adherence
- Presentations at conferences/meetings
- Editorials and newsletters
- Training and education

Dissemination To Date

- 13 Publications: *JAMA, BMJ, Lancet, JCPP, Implementation Science, RSWP, J Exp Crim, BJP, AJPH, Trials, BJSW, Addiction, PLoS One*
- 15+ Presentations: Royal Society of Medicine, EQUATOR/LANCET, Cochrane and Campbell Colloquia, BERA, SPR, APPAM, SSWR, SREE, ASC, Global Implementation Conference
- Other Output: Cross-Whitehall Trial Advice Panel, influence on US Institute of Medicine Framework and Society for Prevention Research 2015 Standards of Evidence, Berkeley Initiative for Transparency in the Social Sciences, MRC Advisory Board (Process Evaluation Guidance), TOP Guidelines, International Behavioural Trials Network

A New Evidence Grading System for Complex Interventions
GRADE-CI

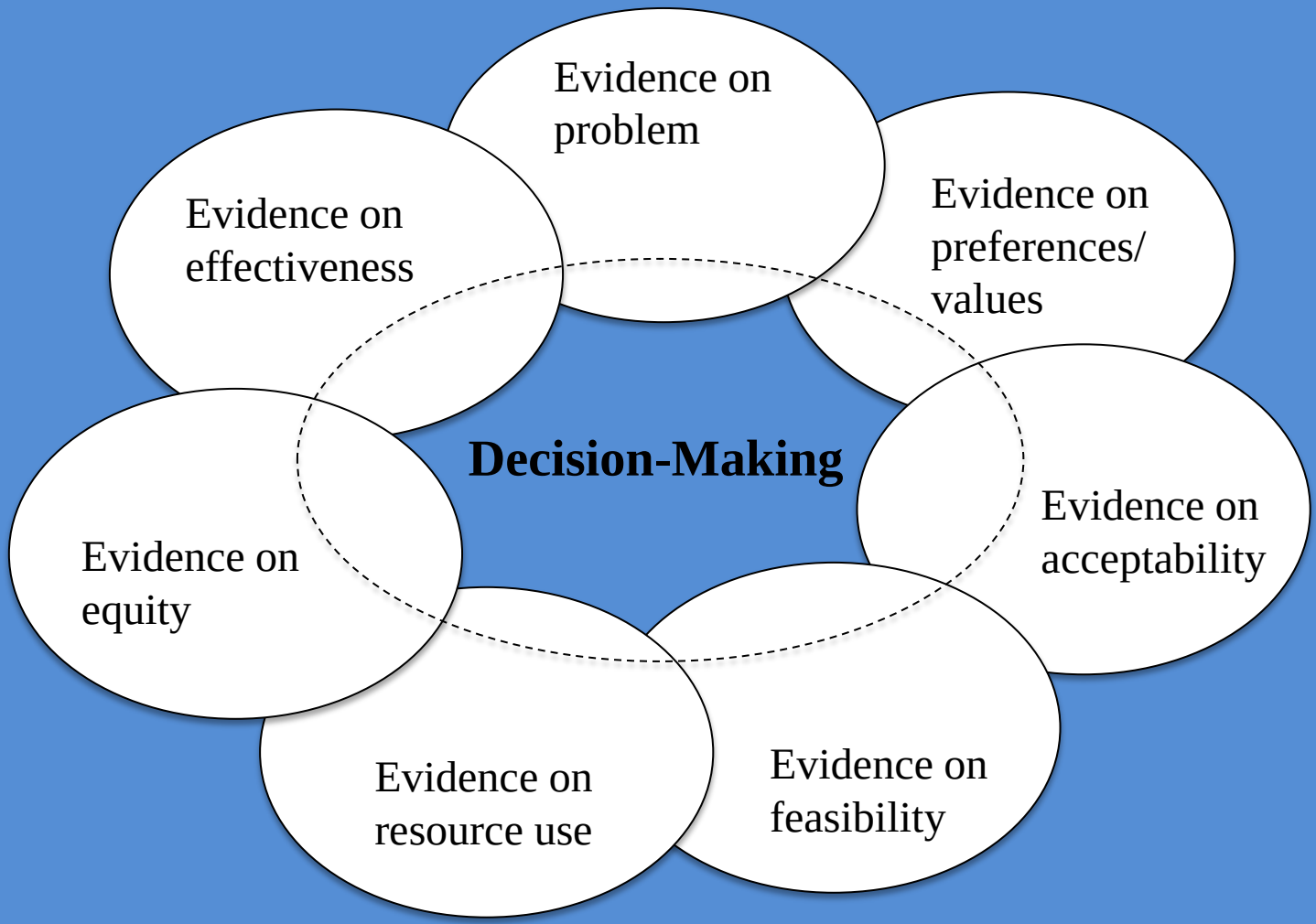
https://www.spi.ox.ac.uk/research/details/grade-extension-for-complex-social-inter.html

The GRADE Approach

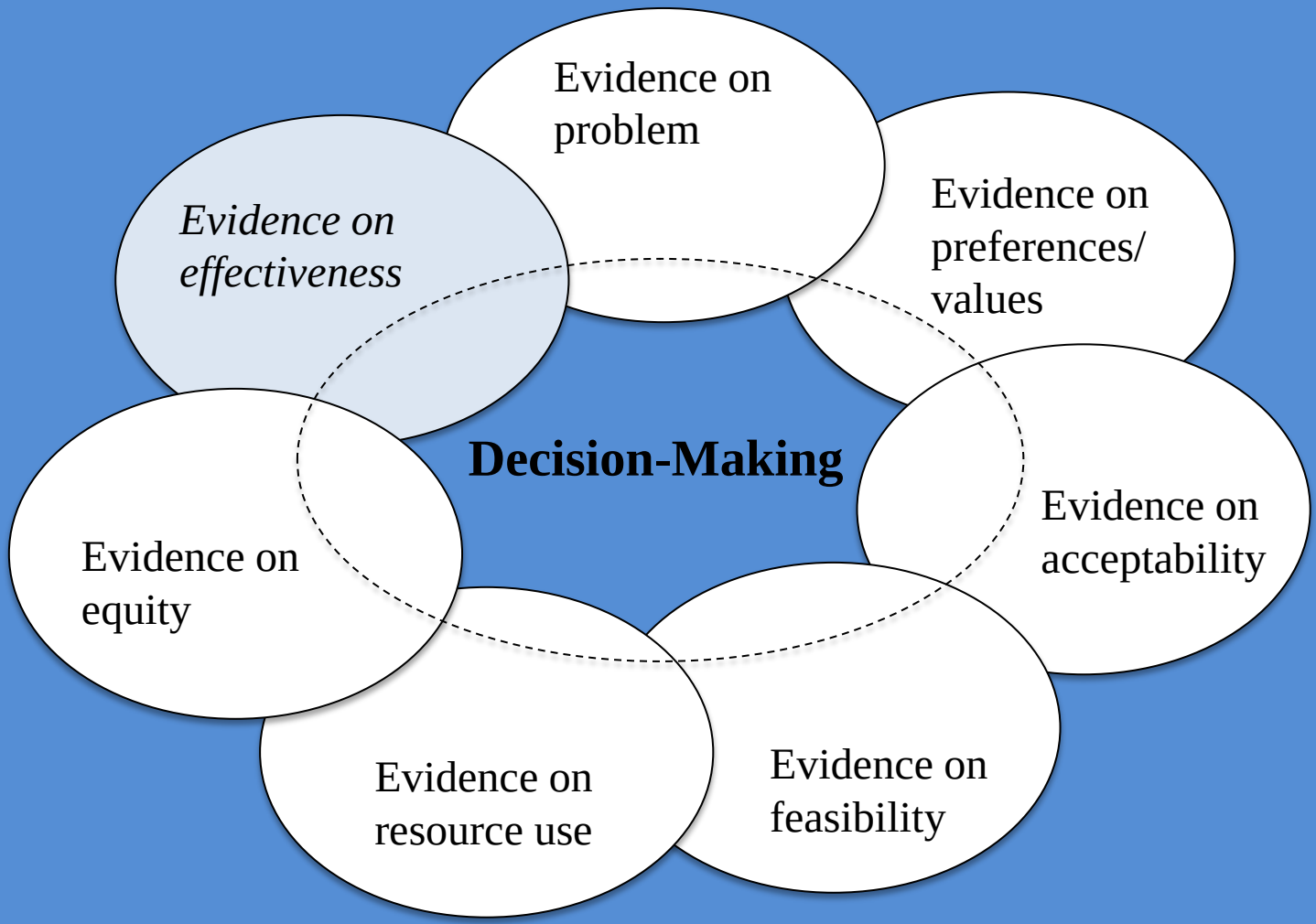
The GRADE approach offers a transparent and structured process for developing and presenting (effectiveness) evidence summaries for systematic reviews and for carrying out steps involved in developing recommendations: It specifies and approach to:

- Framing questions for systematic reviews and guidelines
- Choosing outcomes of interest and rating their importance
- Assessing and rating the quality of a body of evidence
- Incorporating effectiveness evidence with other considerations to arrive at recommendations (DECIDE)

DECIDE criteria to support informed decisions based on evidence



DECIDE criteria to support informed decisions based on evidence



The GRADE Methodology and Process

1. **Question formulation (PICO)** — *Comparative effectiveness*
2. **Search & retrieval of relevant studies**
3. **Evidence synthesis** — *Meta-analysis of RCTs; narrative summary*
4. **Rating the quality of evidence for each outcome**
   RCTs – high, Observational – Low

Downgrading	Upgrading
- Risk of bias	- Large effect
- Inconsistency	- Dose-response
- Imprecision	- All plausible residual confounding
- Indirectness	
- Publication bias	

   Quality of evidence is defined as the extent of our confidence that the estimates of the effect are correct
5. **Rating the overall quality of evidence**
6. **Grading recommendations for practice** — **DECIDE**

- Problem priority	- Acceptability
- Benefits/harms	- Preferences/values
- Quality of evidence	- Resource use
- Feasibility	- Equity

Definitions of the GRADE quality of evidence ratings

Level	GRADE definition	GRADE/DECIDE definition
High	We are very confident that the true effect lies close to that of the estimate of the effect	The research provides a very good indication of the likely effect. The likelihood that the effect will be substantially different is low
Moderate	We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of effect, but there is a possibility that it is substantially different	The research provides a good indication of the likely effect. The likelihood that the effect will be substantially different moderate
Low	Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect	The research provides some indication of the likely effect. However, the likelihood that it will be substantially different is high
Very Low	We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect	The research does not provide a reliable indication of the likely effect. The likelihood that the effect will be substantially different is very high

Guyatt et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011(64);12:1283-1293

Systematic Review/ Guideline Questions

Current Practice

What is the effectiveness of intervention A compared to intervention B for a specific problem in a specific population/setting

- Should present a clear statement of review's objectives
- Should be specific (PICOS)
- Should be relevant and address the needs of different potential stakeholder audiences

Example: Do saving promotion interventions (I) compared to no saving promotion interventions (C) reduce household poverty (O) in sub-Saharan Africa (S)?

Higgins JPT & Green S (2011). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0

Beyond internal validity
Context-specific effectiveness

Moving from "what works" to "what happens"

"The proper agenda for the next generation of treatment effectiveness research, for both primary and meta-analytic studies, is investigation into which treatment variants are most effective, the mediating causal processes through which they work, and the characteristics of recipients, providers, and settings that most influence their results"

Lipsey and Wilson, 1993: p. 1201

ibtn international behavioural trials network

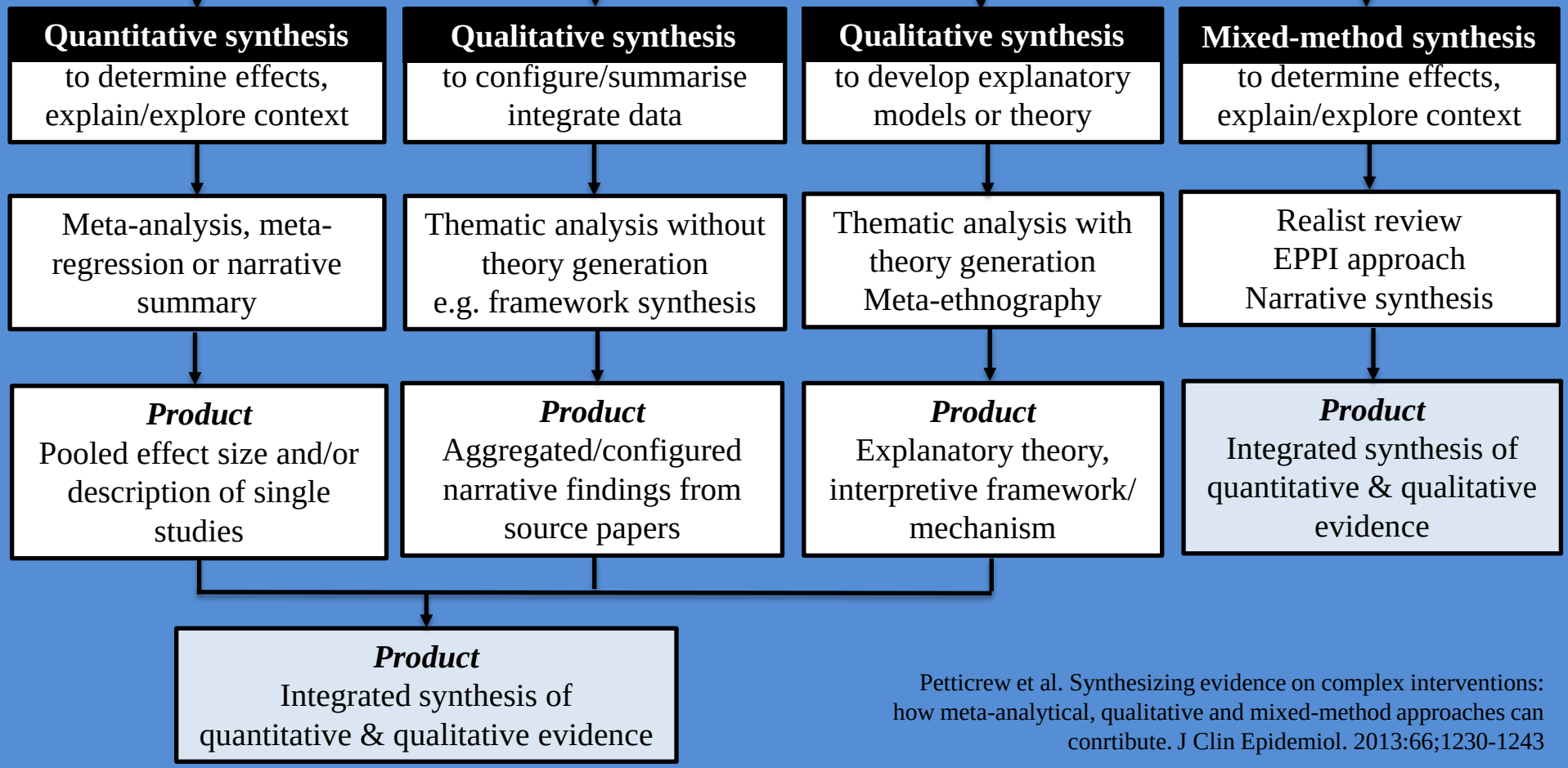
DFID Department for International Development

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL NCRM National Centre for Research Methods

UNIVERSITY OF OXFORD

Using a combination of evidence synthesis methods

(J Clin Epidemiol series)



Petticrew et al. Synthesizing evidence on complex interventions: how meta-analytical, qualitative and mixed-method approaches can conrtibute. J Clin Epidemiol. 2013:66;1230-1243

Challenges of using GRADE in social interventions

1. GRADE terminology and definitions

- **Inappropriate use of terminology**
  Example: use of terms, such as patients and clinicians
- **Irrelevant definition and meaning of quality/confidence**
  Concern: the effects are critically influenced by modes of delivery and contextual factors
  Alternative definition: "confidence that the effect is meaningful across a range of plausible implementation contexts"
- **Inappropriate interpretations of the levels of evidence quality**
  misinterpretations of "low quality evidence" by policymakers?

Rehfuess & Akl (2013)

Challenges of using GRADE in social interventions

2. Evidence base and rigour hierarchy

- **Scarcity of RCTs to address effectiveness questions**
  Concern: GRADE is inflexible when RCTs are not feasible (rigour versus feasibility)
- **Non-randomised studies versus other observational studies**
  Concern: GRADE doesn't differentiate between designs less prone to bias (e.g. ITS) and other observational studies
  Alternative: the selected designs enter the assessment as "moderate"
- **Selection of an appropriate body of evidence**
  Concern: how to prioritise between one large RCT and many Non-RCTs conducted in different contexts?

Harder et al. (2015); Movsisyan et al. (2015); Rehfuess & Akl (2013)

Challenges of using GRADE in social interventions

3. Specific criteria

- **Interpretation of Inconsistency**

  Concern: how to interpret heterogeneity for multi-component interventions when either lumping or splitting?

- **Judgment of Indirectness**

  Concern (1): how to judge about the degree of indirectness for multi-component interventions when either lumping or splitting?

  Concern (2): how to prioritise between many outcomes (short-term versus long-term) and outcome measures, and what are the implications of this for indirectness?

- **Risk of bias assessment**

  Concern (1): downgrading evidence for lack of blinding, when impossible to blind (rigour versus feasibility)

  Concern (2): study designs used for these interventions do not have risk of bias tools for consistent use, which complicates the GRADE assessment (e.g. NRS, SSED, etc.)

Harder et al. (2015); McPheeters et al. (2015); Movsisyan et al. (2015); Rehfuess & Akl (2013)

Challenges of using GRADE in social interventions

4. Making the best use of available evidence

- **Use of non-epidemiological evidence**

  Concern (1): how to incorporate evidence on implementation & context to facilitate context-specific effectiveness assessment in GRADE?

  Alternative (1): Using non-epidemiological evidence not as a separate low quality evidence, but to augment the credibility of epidemiological evidence (e.g. a causal-chain approach)

- **Insufficient possibilities for upgrading observational evidence**

  Alternative (1): upgrade for *consistency* across study designs, settings, research groups

  Alternative (2): upgrade for *analogy* from "parallel evidence", such as evidence from related population groups, interventions

Harder et al. (2015); Movsisyan et al. (2015); Rehfuess & Akl (2013)

GRADE Extension for Complex Social Interventions?

1. Question formulation (what happens?) — *Context-specific effectiveness?*
2. Search & retrieval of relevant studies
3. Evidence synthesis — *A spectrum of evidence synthesis methods?*
4. Rating the quality of evidence for each outcome
   RCTs – high, Observational – Low?

   Downgrading?
   - Risk of bias
   - Inconsistency
   - Imprecision
   - Indirectness
   - Publication bias

   Upgrading?
   - Large effect
   - Dose-response
   - All plausible residual confounding

   Confidence that the effect is meaningful across a range of implementation contexts?
5. Rating the overall quality of evidence
6. **Grading recommendations for practice**
   - Problem priority
   - Benefits/harms
   - Quality of evidence
   - Feasibility
   - Acceptability
   - Preferences/values
   - Resource use
   - Equity

DECIDE

Project Executive

Revise aspects of the GRADE methodology to enable the best use of available evidence to inform decision-making on the effectiveness of complex social interventions

- revise GRADE terminology & definitions
- reconsider the evidence hierarchy within GRADE
- rethink the criteria for rating the quality of evidence

Project Protocol

Start Date: 01.01. 2016 — **End Date: 30.06.2018**

STEERING & COORDINATION

PHASE 1
Project Launch

- Finalise the team
- Build collaboration
- Systematic review

PHASE 2
Online Expert Panel

- Identify participants
- Conduct the panel
- Data analysis

PHASE 3
Consensus Meeting

- Pre-meeting
- Host the meeting
- Data analysis

PHASE 4
Write-up & Testing

- Draft documents
- Feedback & revise
- Finalise documents

DISSEMINATION

Start Date: 01.01. 2016 — **End Date: 30.06.2018**

Project Executive

- Dr Erik von Elm – Institut Universitaire de Médecine Sociale et Preventive (IUMSP), Lausanne, Switzerland
- Dr Eva Rehfuess – Institute of Medical Informatics, Biometry and Epidemiology Ludwig-Maximilians – University, Munich, Germany
- Prof Geraldine Macdonald – University of Bristol, Bristol, UK
- Dr Jane Dennis – Research Synthesis Ltd, Bristol, UK
- Prof Paul Montgomery – Centre for Evidence-Based Social Intervention, University of Oxford
- Dr Sean Grant – RAND Corporation, Santa Monica, USA
- Dr Susan Norris – Guideline Review Committee Secretariat, WHO

International Steering Committee

- Gordon Guyatt
- Holger Schunemann
- Peter Tugwell
- Ian Shemilt
- Stephanie Chang
- Andrew Booth
- Philip Davies
- Birte Snilstveit
- Matthew Morton
- Mark Petticrew
- Steven Hollon
- Bonnie Spring
- Frances Gardner
- Julia Littell
- James Thomas
- Sandra Wilson
- Manual Eisner

Thank you!

- Please email with questions/comments: paul.montgomery@spi.ox.ac.uk
- Visit our websites: http://tinyurl.com/CONSORT-study

  https://www.spi.ox.ac.uk/research/details/grade-extension-for-complex-social-inter.html

Appendix

- CONSORT-SPI Checklist and Flow Diagram

CONSORT-SPI Checklist

Title and Abstract

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
1a	Identification as a randomised trial in the title §	
1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts) §	

§ Indicates that an extension item for cluster trials exists

CONSORT-SPI Checklist

Introduction: Background and objectives

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
2a	Scientific background and explanation of rationale §	
2b	Specific objectives or hypotheses §	If pre-specified, how the intervention was hypothesised to work

CONSORT-SPI Checklist

Methods: Trial Design

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
3a	Description of trial design (such as parallel, factorial) including allocation ratio §	If the unit of random assignment is not the individual, please refer to CONSORT for Cluster Randomised Trials
3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	

CONSORT-SPI Checklist

Methods: Participants

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
4a	Eligibility criteria for participants §	When applicable, eligibility criteria for settings and those delivering the interventions
4b	Settings and locations where the data were collected	

CONSORT-SPI Checklist

Methods: Interventions

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered §	Extent to which interventions were delivered and taken up as planned, including what they actually involved
		*Where other informational materials about delivering the intervention can be accessed
		When applicable, how intervention providers were assigned to each group

***Indicates item might move to another section**

CONSORT-SPI Checklist

Methods: Outcomes

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed §	
6b	Any changes to trial outcomes after the trial commenced, with reasons	

CONSORT-SPI Checklist

Methods: Sample Size

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
7a	How sample size was determined §	
7b	When applicable, explanation of any interim analyses and stopping guidelines	

CONSORT-SPI Checklist

Methods: Randomisation

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
8a	Method used to generate the random allocation sequence	
8b	Type of randomization; details of any restriction (such as blocking and block size) §	
9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned §	
10	Where applicable, who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions §	

CONSORT-SPI Checklist

Methods: Awareness of Assignment

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
11a	Who was aware after assignment to interventions (for example, participants, providers, those assessing outcomes), and how any masking was done	
11b	If relevant, description of the similarity of interventions	

CONSORT-SPI Checklist

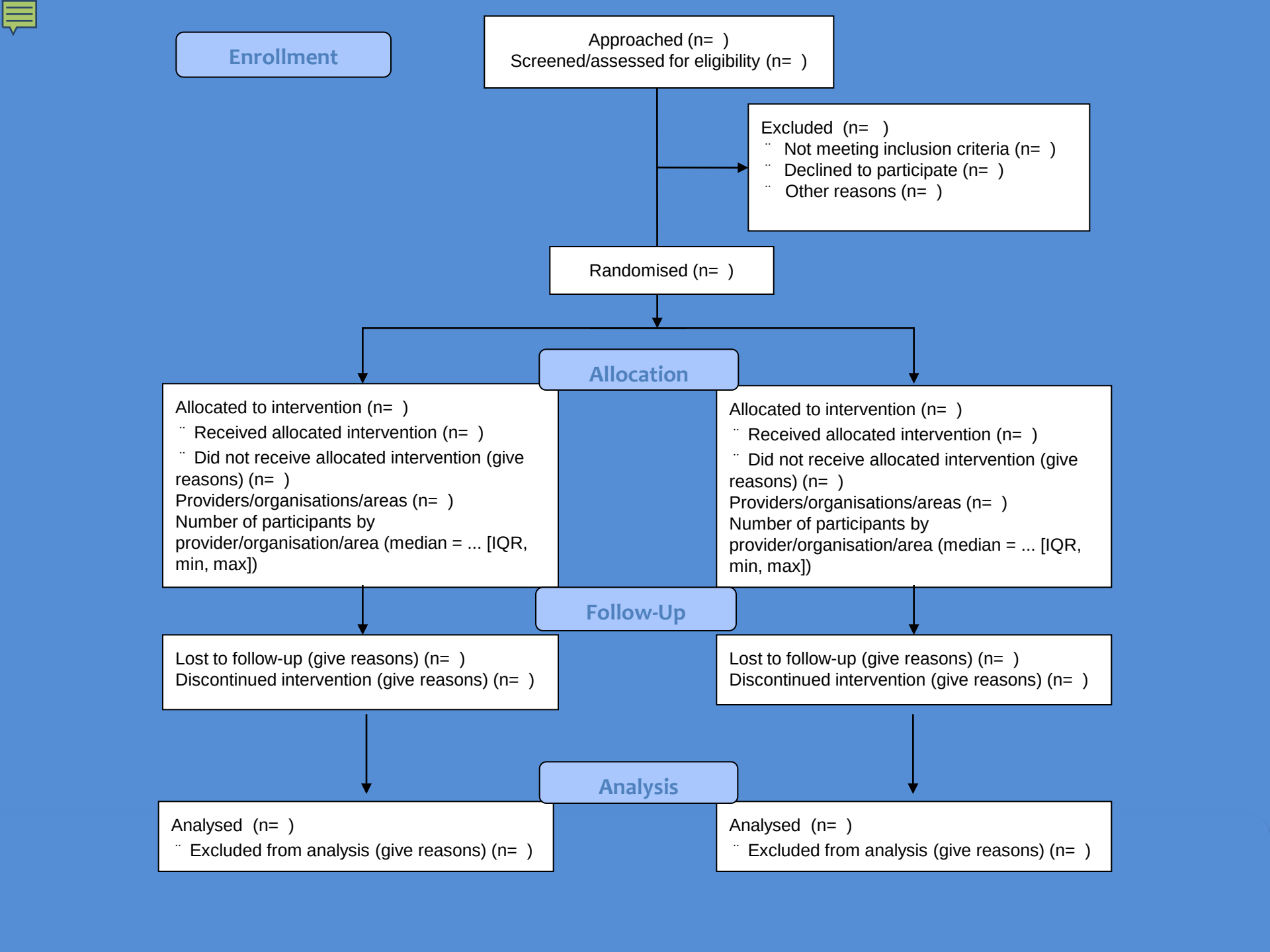
Methods: Analytical Methods

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
12a	Statistical methods used to compare groups for primary and secondary outcomes §	How missing data were handled (e.g., complete case analysis, simple imputation, multiple imputation), with details of any imputation method
12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	

CONSORT-SPI Checklist

Results: Participant Flow

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
13a	For each group, the numbers randomly assigned, received intended treatment, and analysed for the primary outcome §	Where possible, the number approached, screened, and eligible prior to random assignment, with reasons for dropout
13b	For each group, losses and exclusions after randomization, together with reasons §	


Enrollment
Approached (n=)
Screened/assessed for eligibility (n=)
Excluded (n=)
¨ Not meeting inclusion criteria (n=)
¨ Declined to participate (n=)
¨ Other reasons (n=)
Randomised (n=)
Allocation
Allocated to intervention (n=)
¨ Received allocated intervention (n=)
¨ Did not receive allocated intervention (give reasons) (n=)
Providers/organisations/areas (n=)
Number of participants by provider/organisation/area (median = ... [IQR, min, max])
Allocated to intervention (n=)
¨ Received allocated intervention (n=)
¨ Did not receive allocated intervention (give reasons) (n=)
Providers/organisations/areas (n=)
Number of participants by provider/organisation/area (median = ... [IQR, min, max])
Follow-Up
Lost to follow-up (give reasons) (n=)
Discontinued intervention (give reasons) (n=)
Lost to follow-up (give reasons) (n=)
Discontinued intervention (give reasons) (n=)
Analysis
Analysed (n=)
¨ Excluded from analysis (give reasons) (n=)
Analysed (n=)
¨ Excluded from analysis (give reasons) (n=)

CONSORT-SPI Checklist

Results: Recruitment

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
14a	Dates defining the periods of recruitment and follow-up	
14b	Why the trial ended or was stopped	

CONSORT-SPI Checklist

Results: Baseline Data and Numbers

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
15	A table showing baseline characteristics for each group §	Including socioeconomic variables where applicable
16	For each group, number included in each analysis and whether the analysis was by original assigned groups §	

CONSORT-SPI Checklist

Results: Outcomes and Estimation

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval) §	*Indicate availability of trial data
17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	
18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	
19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	

CONSORT-SPI Checklist

Discussion

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	
21	Generalisability (external validity, applicability) of the trial findings §	
22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	

CONSORT-SPI Checklist

Important Information

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
23	Registration number and name of trial registry	
24	Where the full trial protocol can be accessed, if available	
25	Sources of funding and other support, role of funders	Declaration of any other potential interests

CONSORT-SPI Checklist

Stakeholder Involvement

Item #	Standard CONSORT Description	Extension for CONSORT-SPI
New Item		*Any involvement of the intervention developer in the design, conduct, analysis, and reporting of the trial
		*Other stakeholder involvement in trial design, conduct, and/or analyses
		*Incentives offered as part of the trial