PAFS Pilot and Feasibility Studies Consort Extension

Pilot and feasibility studies

NIHR workshop

April 21st

London

McMaster University

University of Aberdeen

The University Of Sheffield.

University of Oxford

Lancaster University

Introduction

Sandra Eldridge and Claire Coleman

Facilitators

- Sandra Eldridge
- Claire Coleman
- Gill Lancaster
- Mike Campbell
- Christine Bond

Sessions

Introduction

The need for guidelines for reporting & conduct

Definitions and objectives

Study design and sample size

Analysis and progression criteria

Reporting

Closing remarks

Why this topic?
Why us?

Developing reporting guidelines for pilot and feasibility studies

Sandra Eldridge
Mike Campbell
Gill Lancaster
Lehana Thabane
Christine Bond
Sally Hopewell
Claire Coleman

Background

- Poor reporting
- Poor design
- Difficulty getting funded and published
- Annual meeting of Society for Academic Primary Care - July 2011
- UK Medical Research Council (MRC) guidance on developing complex interventions
- UK National Institute of Health Research (NIHR), interest & definitions

Journal of Evaluation in Clinical Practice, 10, 2, 307–312

Design and analysis of pilot studies: recommendations for good practice

Gillian A. Lancaster MSc PhD,[1] Susanna Dodd MSc[2] and Paula R. Williamson PhD[3]
[1]Lecturer in Medical Statistics, [2]Research Assistant in Medical Statistics [3]Senior Lecturer in Medical Statistics, Department of Mathematical Sciences, University of Liverpool, Liverpool, UK

Correspondence Abstract

Arain *et al. BMC Medical Research Methodology* 2010, **10**:67
http://www.biomedcentral.com/1471-2288/10/67

BMC Medical Research Methodology

CORRESPONDENCE Open Access

What is a pilot or feasibility study? A review of current practice and editorial policy

Mubashir Arain[1], Michael J Campbell*[1], Cindy L Cooper[1] and Gillian A Lancaster[2]

Thabane *et al. BMC Medical Research Methodology* 2010, **10**:1
http://www.biomedcentral.com/1471-2288/10/1

BMC Medical Research Methodology

COMMENTARY Open Access

A tutorial on pilot studies: the what, why and how

Lehana Thabane[1,2*], Jinhui Ma[1,2], Rong Chu[1,2], Ji Cheng[1,2], Afisi Ismaila[1,3], Lorena P Rios[1,2], Reid Robson[3], Marroon Thabane[1,4], Lora Giangregorio[5], Charles H Goldsmith[1,2]

Project scope

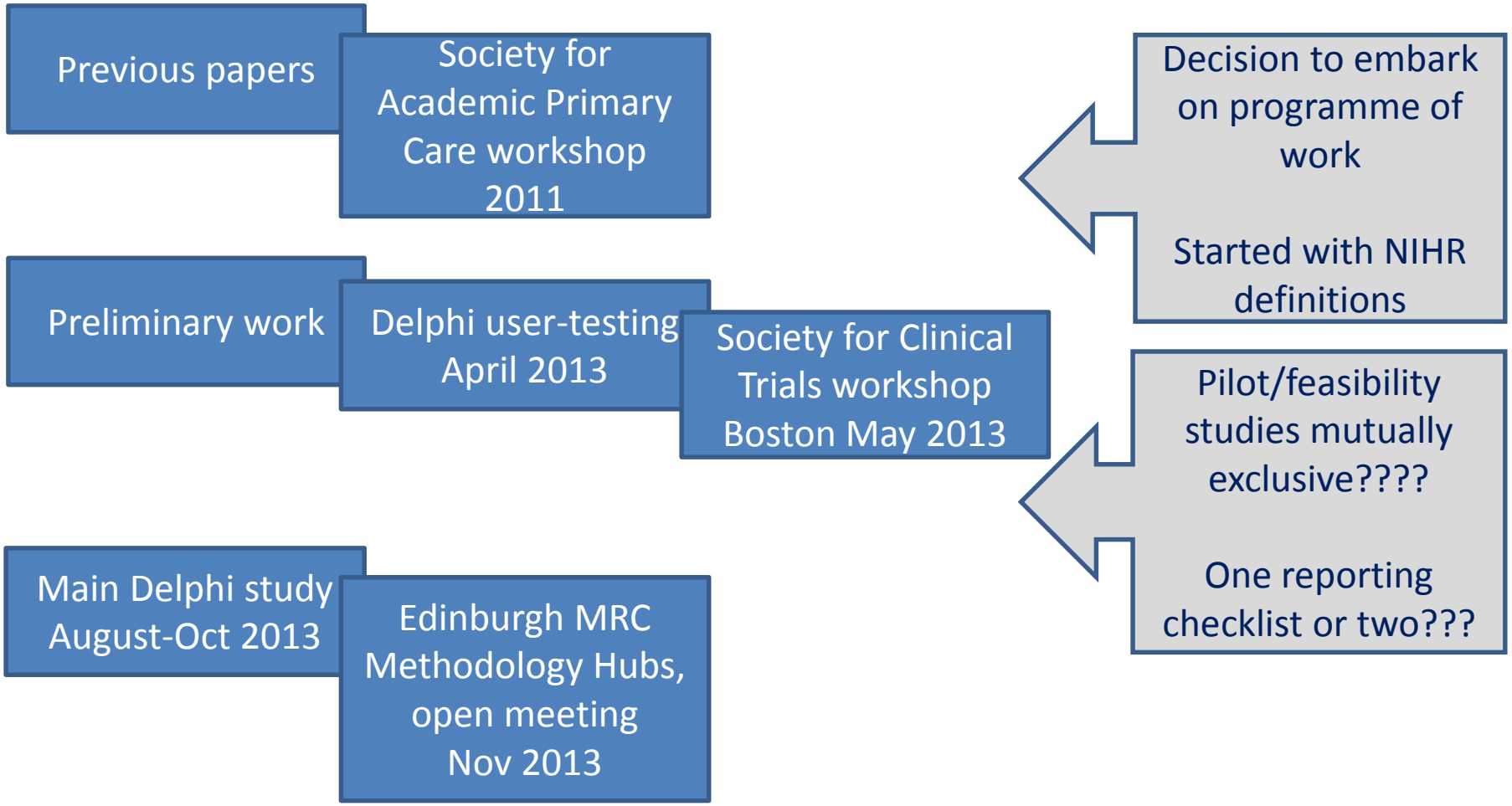
Reporting guidelines

Framework for understanding pilot and feasibility studies


Previous papers
Society for Academic Primary Care workshop 2011
Preliminary work
Delphi user-testing April 2013
Society for Clinical Trials workshop Boston May 2013
Decision to embark on programme of work
Started with NIHR definitions

User-testing of DELPHI (n=13)

- *"….. study was both feasibility and pilot study"*
- "*No longer sure knows the difference*"
- "*Well nobody uses the definitions so it doesn't seem to matter, also there are many more terms used*"
- *"The definitions are taken from the funders so how can you change them?"*


Previous papers
Society for Academic Primary Care workshop 2011
Preliminary work
Delphi user-testing April 2013
Society for Clinical Trials workshop Boston May 2013
Main Delphi study August-Oct 2013
Edinburgh MRC Methodology Hubs, open meeting Nov 2013
Decision to embark on programme of work
Started with NIHR definitions
Pilot/feasibility studies mutually exclusive????
One reporting checklist or two???

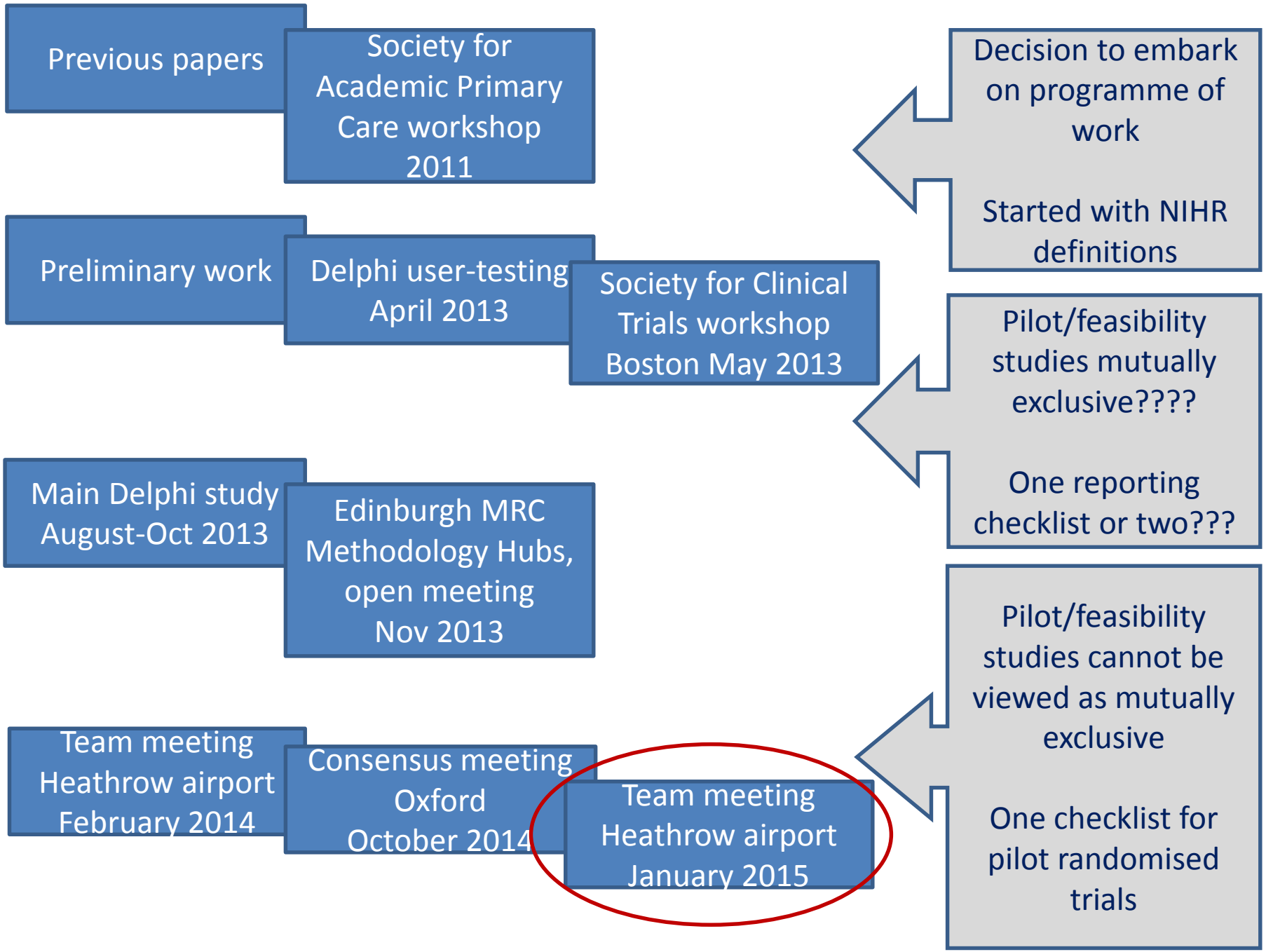
Results from main DELPHI

- Very strong consensus on items in checklist both for pilot and feasibility studies
- BUT still substantial disagreement about definitions of pilot and feasibility studies and their separation
- For example:

"The terms 'feasibility' and 'pilot' are not mutually exclusive. They are used interchangeably in the literature and it would be confusing to try and separate them out into two artificial sets of definitions"

Edinburgh open meeting

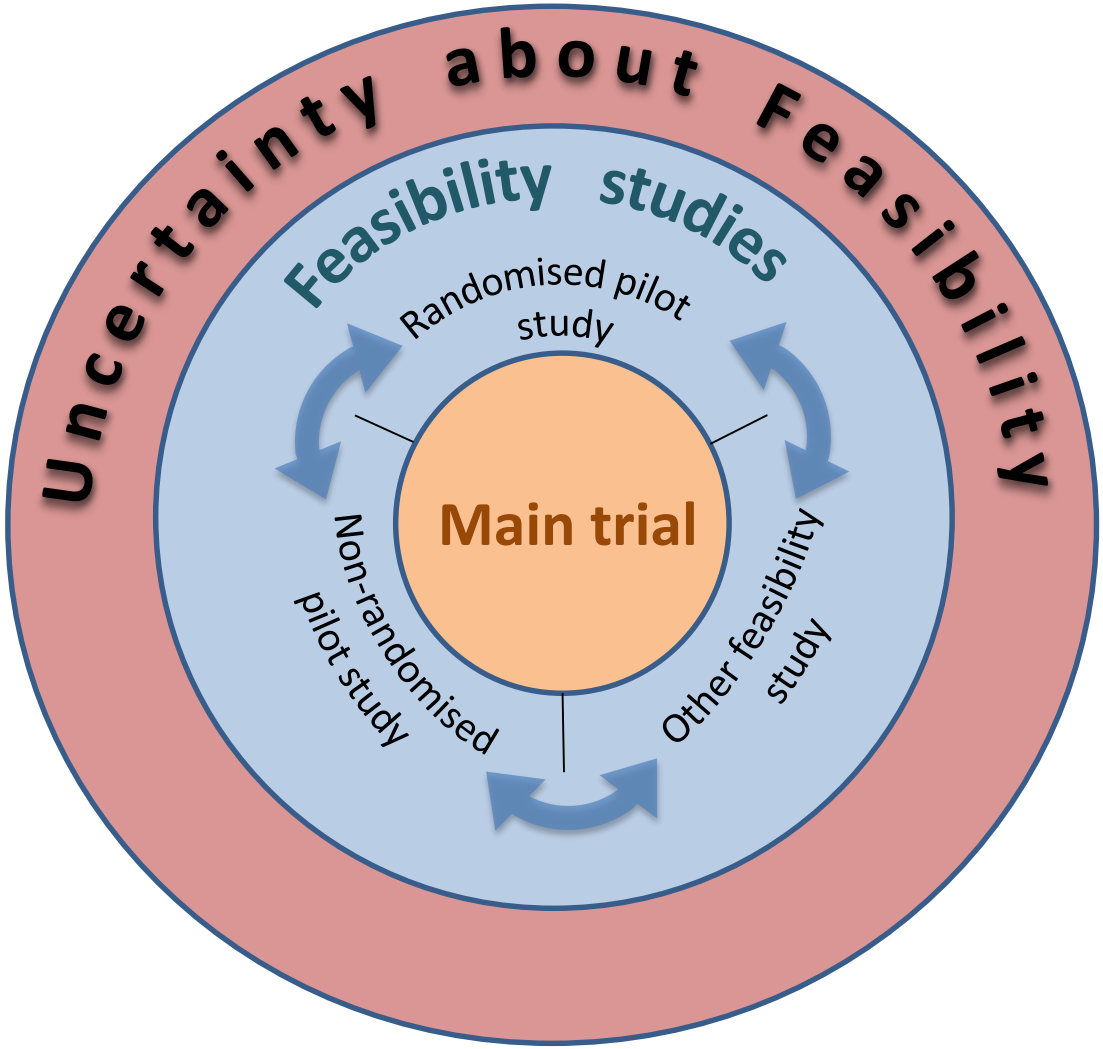
- Four propositions for definitions
- 14/15 participants preferred propositions which implied
  - Feasibility and pilot studies not mutually exclusive
  - Develop only one checklist


Previous papers
Society for Academic Primary Care workshop 2011
Decision to embark on programme of work
Started with NIHR definitions
Preliminary work
Delphi user-testing April 2013
Society for Clinical Trials workshop Boston May 2013
Pilot/feasibility studies mutually exclusive????
One reporting checklist or two???
Main Delphi study August-Oct 2013
Edinburgh MRC Methodology Hubs, open meeting Nov 2013
Team meeting Heathrow airport February 2014
Consensus meeting Oxford October 2014
Team meeting Heathrow airport January 2015
Pilot/feasibility studies cannot be viewed as mutually exclusive
One checklist for pilot randomised trials

Current stage

- Finalising checklist from Consensus meeting for CONSORT extension for randomised pilot studies (more in session 4)
- Finalising presentation of framework relating to definitions (more in session 1)
- Agreement that need further reporting guidance but not further checklists

Framework (more in session 1)



What we are covering today

- Studies conducted in preparation for a future trial designed to measure the effect of an intervention
- That ask about the feasibility of the future trial and whether we should proceed with that future trial

Your examples

The need for guidelines for reporting and conduct

Sandra Eldridge

Introduction

- Poor reporting
- Poor design
- Difficulty getting funded and published
- Annual meeting of Society for Academic Primary Care - July 2011
- UK Medical Research Council (MRC) guidance on developing complex interventions
- UK National Institute of Health Research (NIHR), interest & definitions

Journal of Evaluation in Clinical Practice, 10, 2, 307–312

Design and analysis of pilot studies: recommendations for good practice

Gillian A. Lancaster MSc PhD,[1] Susanna Dodd MSc[2] and Paula R. Williamson PhD[3]
[1]Lecturer in Medical Statistics, [2]Research Assistant in Medical Statistics [3]Senior Lecturer in Medical Statistics, Department of Mathematical Sciences, University of Liverpool, Liverpool, UK

Correspondence Abstract

Arain *et al. BMC Medical Research Methodology* 2010, **10**:67
http://www.biomedcentral.com/1471-2288/10/67

BMC Medical Research Methodology

CORRESPONDENCE Open Access

What is a pilot or feasibility study? A review of current practice and editorial policy

Mubashir Arain[1], Michael J Campbell*[1], Cindy L Cooper[1] and Gillian A Lancaster[2]

Thabane *et al. BMC Medical Research Methodology* 2010, **10**:1
http://www.biomedcentral.com/1471-2288/10/1

BMC Medical Research Methodology

COMMENTARY Open Access

A tutorial on pilot studies: the what, why and how

Lehana Thabane[1,2*], Jinhui Ma[1,2], Rong Chu[1,2], Ji Cheng[1,2], Afisi Ismaila[1,3], Lorena P Rios[1,2], Reid Robson[3], Marroon Thabane[1,4], Lora Giangregorio[5], Charles H Goldsmith[1,2]

Journal Survey

Arain et al wrote to seven journal editors, Lancet, BMJ, JAMA, NEJM, Brit J Cancer, BR J Surgey, Brit J Obs & Gynae (those selected by Gill earlier)

Journal responses

Mostly reported that pilot trials cannot be published if the standard is lower than a full clinical trial requirement.

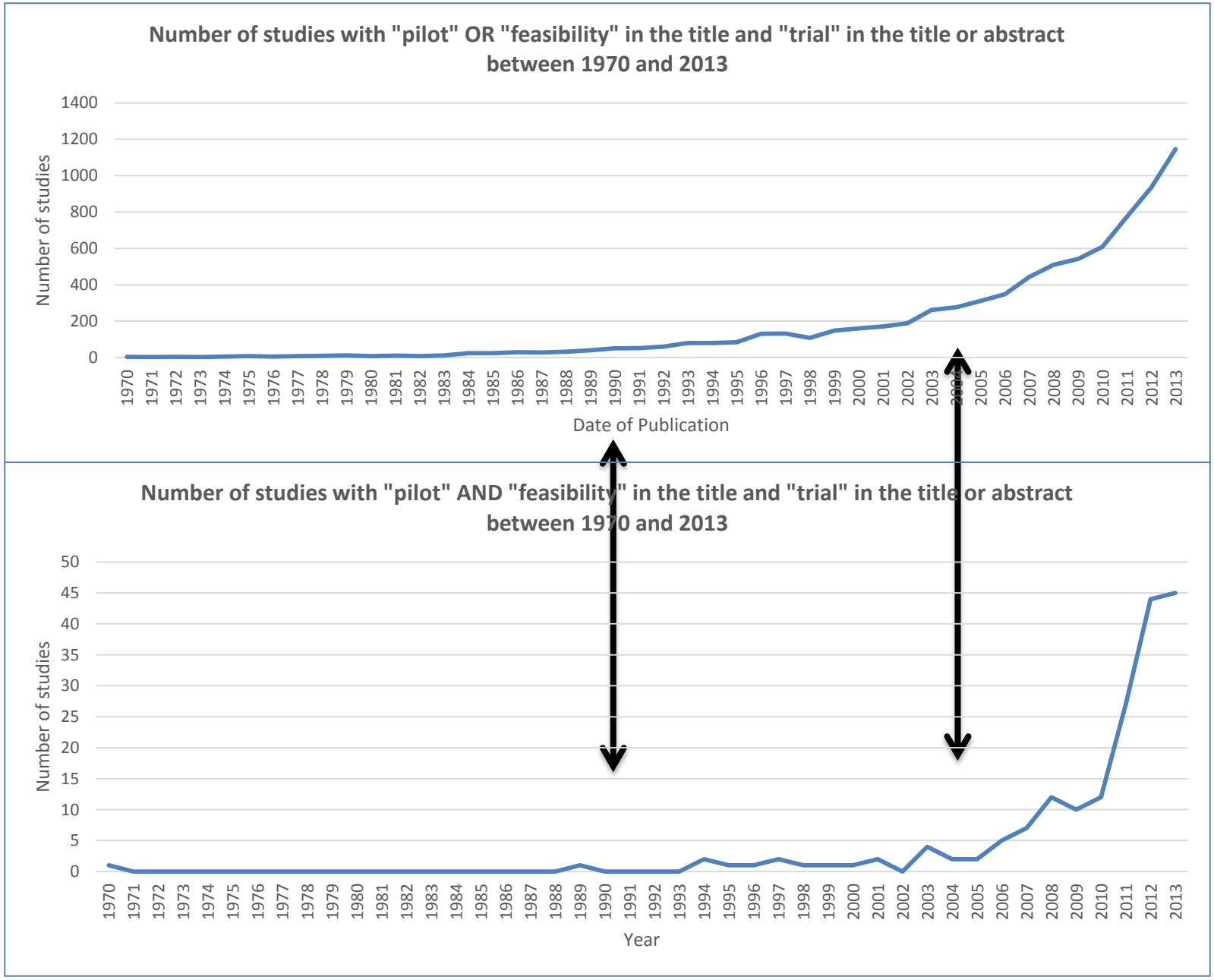
Most of the other journals do not encourage the publication of pilot studies because they consider them less rigorous than main studies.

Some editors accepted pilot studies for publication by compromising only on the requirement for a pre-study sample size calculation.

All other methodological issued were considered as important as for the full trials, with reporting according to the CONSORT guidelines.

Cautionary tale

- Editor 1: "…..it might be **more convincing if reported in more conventional style with p values**, appropriate attention to the calculation of sample size and both intention to treat and per protocol analyses"
- Editor 1: "….the fact remains that studies with **results that are definitive and clinically directive** will always have a better chance"
- Editor 2: "…..the **lack of objective outcomes** and the incomplete matching between groups"
- Editor 3: "We do appreciate the effort behind the study, and **its value to the scientific community**, but it can unfortunately not achieve sufficient priority to be considered"


Number of studies with "pilot" OR "feasibility" in the title and "trial" in the title or abstract between 1970 and 2013
Number of studies
1400
1200
1000
800
600
400
200
0
1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013
Date of Publication
Number of studies with "pilot" AND "feasibility" in the title and "trial" in the title or abstract between 1970 and 2013
Number of studies
50
45
40
35
30
25
20
15
10
5
0
1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013
Year

More work since 2011

Charlesworth *et al. BMC Medical Research Methodology* 2013, **13**:78
http://www.biomedcentral.com/1471-2288/13/78

BMC Medical Research Methodology

CORRESPONDENCE **Open Access**

Acceptance checklist for clinical effectiveness pilot trials: a systematic approach

Georgina Charlesworth[1,2*], Karen Burnell[3], Juanita Hoe[4], Martin Orrell[2,4] and Ian Russell[5]

Bugge *et al. Trials* 2013, **14**:353
http://www.trialsjournal.com/content/14/1/353

TRIALS

METHODOLOGY **Open Access**

A process for Decision-making after Pilot and feasibility Trials (ADePT): development following a feasibility study of a complex intervention for pelvic organ prolapse

Carol Bugge[1*], Brian Williams[2], Suzanne Hagen[3], Janet Logan[3], Cathryn Glazener[4], Stewart Pringle[5] and Lesley Sinclair[6]

Shanyinde *et al. BMC Medical Research Methodology* 2011, **11**:117
http://www.biomedcentral.com/1471-2288/11/117

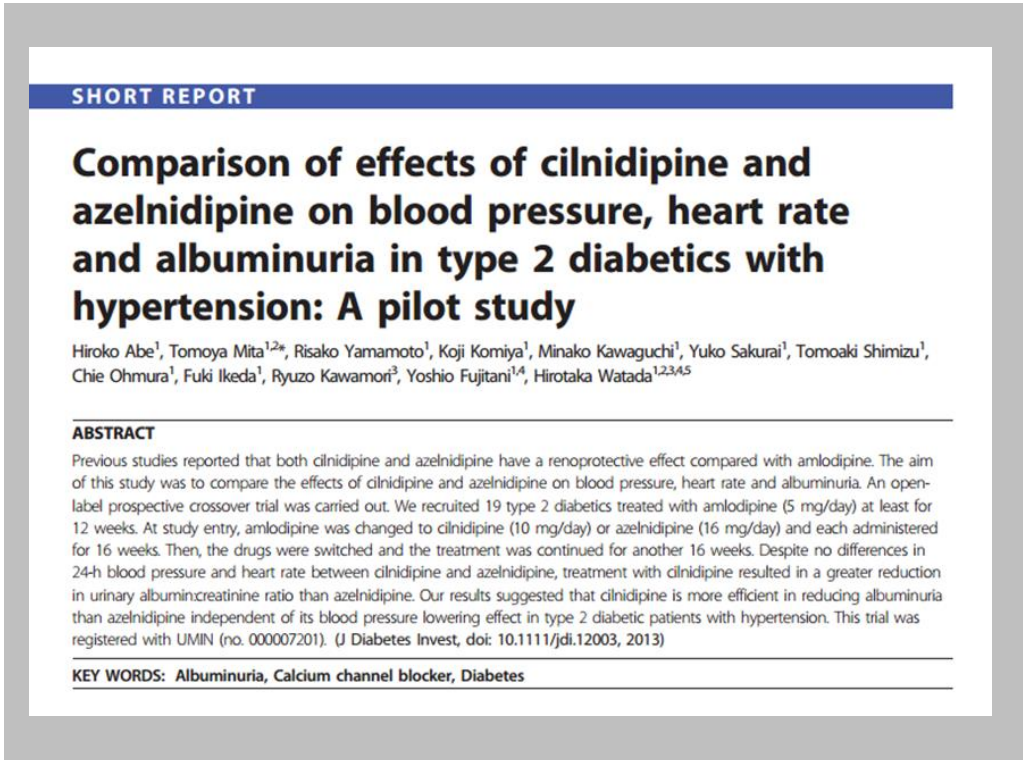
BMC Medical Research Methodology

RESEARCH ARTICLE **Open Access**

Questions asked and answered in pilot and feasibility randomized controlled trials

Milensu Shanyinde[1], Ruth M Pickering[1*] and Mark Weatherall[2]

Example: Small trial called a pilot



SHORT REPORT

Comparison of effects of cilnidipine and azelnidipine on blood pressure, heart rate and albuminuria in type 2 diabetics with hypertension: A pilot study

Hiroko Abe[1], Tomoya Mita[1,2]*, Risako Yamamoto[1], Koji Komiya[1], Minako Kawaguchi[1], Yuko Sakurai[1], Tomoaki Shimizu[1], Chie Ohmura[1], Fuki Ikeda[1], Ryuzo Kawamori[3], Yoshio Fujitani[1,4], Hirotaka Watada[1,2,3,4,5]

ABSTRACT

Previous studies reported that both cilnidipine and azelnidipine have a renoprotective effect compared with amlodipine. The aim of this study was to compare the effects of cilnidipine and azelnidipine on blood pressure, heart rate and albuminuria. An open-label prospective crossover trial was carried out. We recruited 19 type 2 diabetics treated with amlodipine (5 mg/day) at least for 12 weeks. At study entry, amlodipine was changed to cilnidipine (10 mg/day) or azelnidipine (16 mg/day) and each administered for 16 weeks. Then, the drugs were switched and the treatment was continued for another 16 weeks. Despite no differences in 24-h blood pressure and heart rate between cilnidipine and azelnidipine, treatment with cilnidipine resulted in a greater reduction in urinary albumin:creatinine ratio than azelnidipine. Our results suggested that cilnidipine is more efficient in reducing albuminuria than azelnidipine independent of its blood pressure lowering effect in type 2 diabetic patients with hypertension. This trial was registered with UMIN (no. 000007201). (J Diabetes Invest, doi: 10.1111/jdi.12003, 2013)

KEY WORDS: Albuminuria, Calcium channel blocker, Diabetes

Problems with Abe et al

- No sample size calculation
- Small trial (n=19)
- Short follow-up (32 weeks)
- Doesn’t lead to a bigger study

Summary

- Existing recommendations
- Increasing interest
- Increasing empirical trials
- Lack of understanding from journals
- Poor practice

Definitions and Objectives of pilot and feasibility studies

Gill Lancaster

Definitions

- Large and growing number of studies in the literature called feasibility or pilot studies
- Terms ‘pilot’ and ‘feasibility’ are not used consistently
- Makes providing guidance on robust reporting of these studies more challenging

2004

Journal of
Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

Design and analysis of pilot studies: recommendations for good practice

Gillian A. Lancaster MSc PhD[1,*], Susanna Dodd MSc[2], Paula R. Williamson PhD[3]

Article first published online: 4 JUN 2004
DOI: 10.1111/j..2002.384.doc.x

Issue

Journal of Evaluation in Clinical Practice
Volume 10, Issue 2, pages 307–312, May 2004

2010

BMC Medical Research Methodology

CORRESPONDENCE Open Access

What is a pilot or feasibility study? A review of current practice and editorial policy

Mubashir Arain[1], Michael J Campbell*[1], Cindy L Cooper[1] and Gillian A Lancaster[2]

Thabane *et al. BMC Medical Research Methodology* 2010, **10**:1
http://www.biomedcentral.com/1471-2288/10/1

BMC Medical Research Methodology

COMMENTARY Open Access

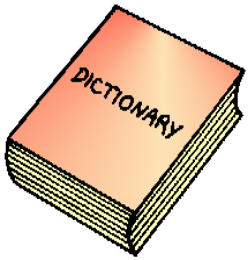
A tutorial on pilot studies: the what, why and how

Lehana Thabane[1,2*], Jinhui Ma[1,2], Rong Chu[1,2], Ji Cheng[1,2], Afisi Ismaila[1,3], Lorena P Rios[1,2], Reid Robson[3], Marroon Thabane[1,4], Lora Giangregorio[5], Charles H Goldsmith[1,2]

Guidance from the literature

- Lancaster et al. (2004) could find no formal guidance as to what constitutes a pilot study;
- Arain et al. (2010) - studies described as ‘feasibility’ studies had slightly different characteristics from those described as ‘pilot’
- Thabane et al. (2010) - number of definitions of pilot studies taken from health related websites
  - common idea of conducting a preliminary study
  - “*a pilot study is synonymous with a feasibility study intended to guide the planning of a large scale investigation*”

Guidance from a dictionary



Pilot

- Done as an experiment or test before being introduced more widely (Oxford dictionary)
- A small-scale experiment or set of observations undertaken to decide ***how and whether to launch a full-scale project***

Feasibility study

- Looks at the viability of an idea with an emphasis on identifying potential problems and attempts to answer one main question: ***will the idea work and should we proceed with it***
- An evaluation and analysis of the potential of the proposed project which is based on extensive investigation and research ***to support the process of decision making***

Guidance on complex interventions

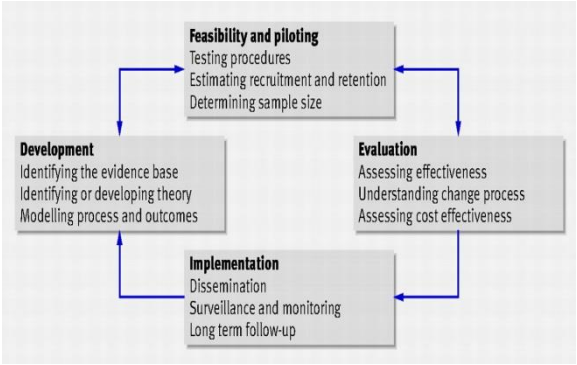
- MRC document
  'Developing and Evaluating Complex Interventions'

  www.mrc.ac.uk/complexinterventionsguidance

  Craig P. et al. BMJ 2008, 337:a1655
  - BMJ paper (Campbell NC et al. 2007, 334: 455-9)
    'Designing and Evaluating Complex Interventions to improve health care'
  - Case studies
- NIHR framework - glossary

  http://www.netscc.ac.uk/glossary/#glos6

MRC framework for complex interventions - feasibility and piloting



Puts pilot studies and all other types of feasibility studies together under one umbrella

Uses feasibility as an overarching term

"A pilot study need not be a 'scale model' of the planned main stage evaluation, but ***should address the main uncertainties*** *that have been identified in the development work."*

NIHR definition of feasibility study

Feasibility Studies are pieces of research done before a main study in order to answer the question "Can this study be done?". They are used to estimate important parameters that are needed to design the main study. For instance:

- *standard deviation of the outcome measure, which is needed in some cases to estimate sample size;*
- *willingness of participants to be randomised;*
- *willingness of clinicians to recruit participants;*
- *number of eligible patients, carers or other appropriate participants*
- *characteristics of the proposed outcome measure and in some cases feasibility studies might involve designing a suitable outcome measure;*
- *follow-up rates, response rates to questionnaires, adherence/compliance rates, ICCs in cluster trials, etc.*
- *availability of data needed or the usefulness and limitations of a particular database*
- *time needed to collect and analyse data*

NIHR definition of pilot study

Pilot studies are ***a smaller version of the main study*** *used to test* ***whether the components of the main study can all work together.*** *It is focused on the processes of the main study, for example to ensure that recruitment, randomisation, treatment, and follow-up assessments all run smoothly. It resembles the main study in many respects,* ***including an assessment of the primary outcome.*** *In some cases, this will be the first phase of the substantive study and data from the pilot phase may contribute to the final analysis; this can be referred to as an internal pilot. Or, at the end of the pilot study, the data may be analysed and set aside, a so-called external pilot.*

Hypotheses for a conceptual framework

- The words pilot and feasibility are both used within the literature to describe studies undertaken in preparation for a RCT of effectiveness
- It is not possible to apply unique definitions of pilot and feasibility studies in preparation for a RCT, consistent with the way authors describe their studies
- It is possible to identify some studies that are not pilot studies as defined within our conceptual framework, but that test the feasibility or acquire related information about applying an intervention in a future study.

Examples

To **assess feasibility** of RCT of management of reduced fetal movement (Heazell et al. BMC Preg Childbirth 2013)

- Recruitment , retention, acceptability , adherence to protocol, prevalence of poor perinatal outcomes

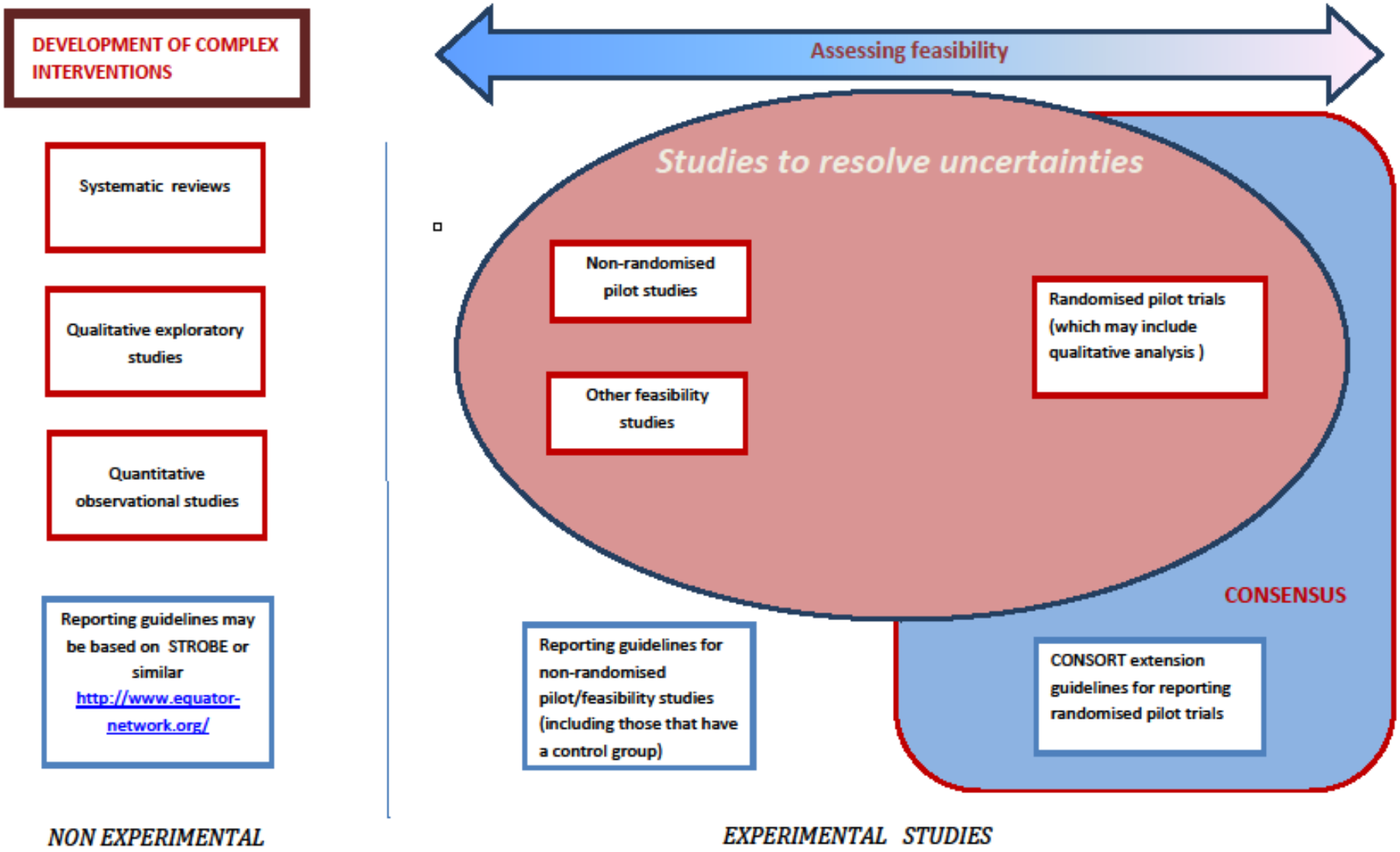
To **pilot an intervention** to avoid the use of syringes and contamination of materials amongst injecting drug users (Colon et al. AIDS Behav. 2009)

- Adoption of each of four components
- Whether pre-post changes in blood residues indicated that intervention merited further testing

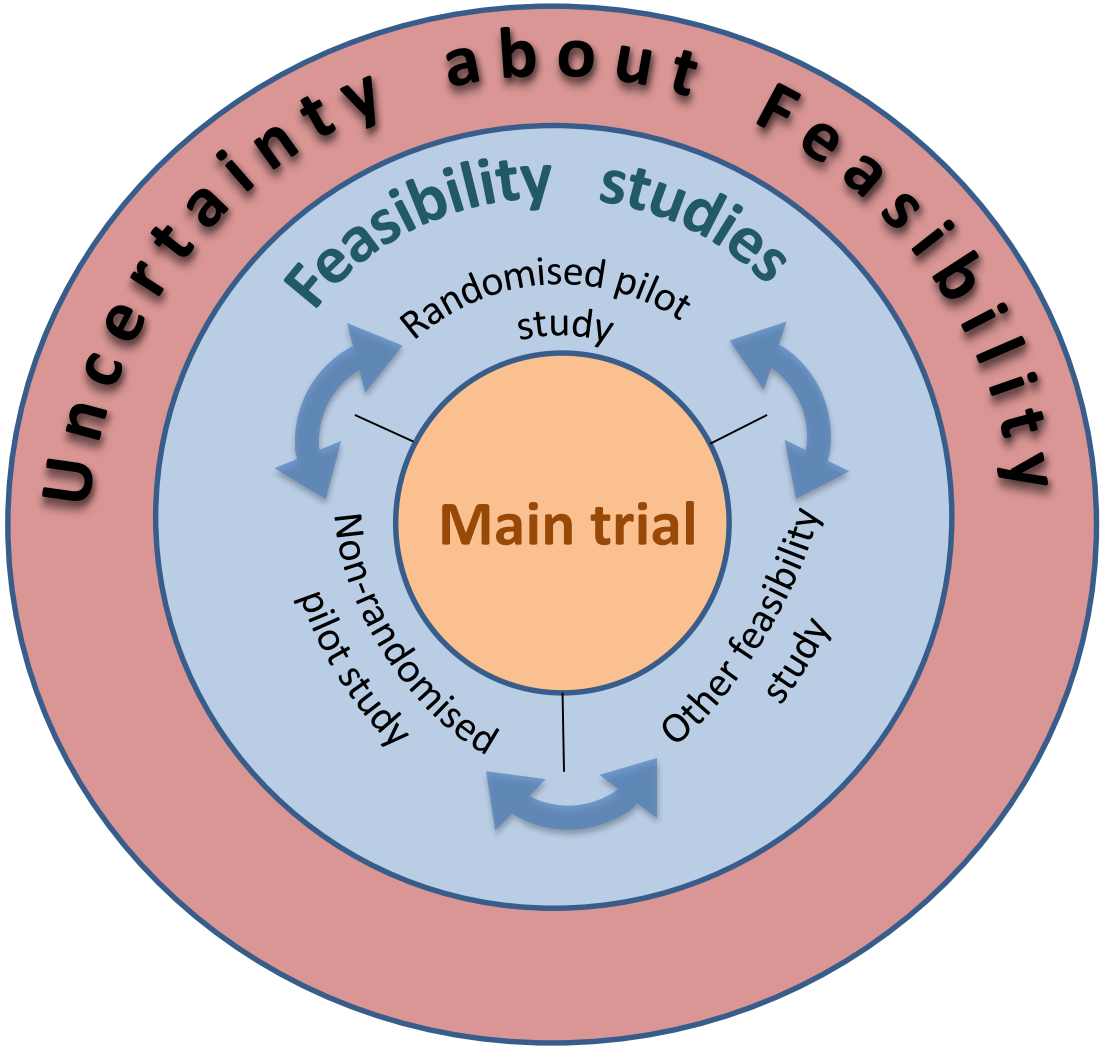
To **determine feasibility** of RCT comparing operative with non-operative treatment for femoroacetabular impingement surgery (Palmer et al. Bone Joint Res. 2013)

- Surgeon and patient opinion via a questionnaire

CONCEPTUALFRAMEWORK FOR TRIAL DEVELOPMENT WORK AND ASSOCIATED REPORTING GUIDELINES



Studies to resolve uncertainties



Workable definitions?

- Feasibility is a concept encapsulating ideas about whether something will work. A feasibility study asks whether this can be done and should we proceed with it (and if so how)
- A pilot study is a study in which a part or a whole of a future study is conducted on a smaller scale to see whether it will work
- Corollary: all pilot studies are feasibility studies but not all feasibility studies are pilot studies

Objectives of pilot and feasibility studies

Journal of
Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

Design and analysis of pilot studies: recommendations for good practice

Gillian A. Lancaster MSc PhD[1,*], Susanna Dodd MSc[2], Paula R. Williamson PhD[3]

Article first published online: 4 JUN 2004

DOI: 10.1111/j..2002.384.doc.x

Issue

Journal of Evaluation in Clinical Practice

Volume 10, Issue 2, pages 307–312, May 2004

Trials in primary care: statistical issues in the design, conduct and evaluation of complex interventions

GA Lancaster, MJ Campbell, S Eldridge, A Farrin, M Marchant, S Muller, R Perera, TJ Peters, AT Prevost, G Rait

Stat Methods Med Res 2010; 19, 4: 349-377

BMC Medical Research Methodology

CORRESPONDENCE Open Access

What is a pilot or feasibility study? A review of current practice and editorial policy

Mubashir Arain[1], Michael J Campbell*[1], Cindy L Cooper[1] and Gillian A Lancaster[2]

Thabane *et al. BMC Medical Research Methodology* 2010, **10**:1
http://www.biomedcentral.com/1471-2288/10/1

BMC Medical Research Methodology

COMMENTARY Open Access

A tutorial on pilot studies: the what, why and how

Lehana Thabane[1,2*], Jinhui Ma[1,2], Rong Chu[1,2], Ji Cheng[1,2], Afisi Ismaila[1,3], Lorena P Rios[1,2], Reid Robson[3], Marroon Thabane[1,4], Lora Giangregorio[5], Charles H Goldsmith[1,2]

Key messages

- Lancaster et al. (2004) – "*Pilot studies should have a well-defined set of aims and objectives to ensure methodological rigour and scientific validity*".
- Arain et al. (2010) – "*reporting of pilot studies was still poor … pilot studies have different objectives to RCTs and these should be clearly described*".
- Thabane and colleagues (2010) - based on reviewing submissions to ethics committees, "*no clear feasibility objectives; no clear analytic plans; and certainly no clear criteria for the success of feasibility*".

Pilot and feasibility studies

- Important if planning multi-centre study, complex intervention, cluster RCT
- May be pre-requisite for funding
- Subject to publication bias
- Test the integrity of the main study protocol
- Focus on ensuring processes of main study are understood and well-organised
- Internal or external pilot – needs to be specified beforehand
- Internal pilots are part of the main trial and should be planned as such

JECP 2004 – Systematic Review

Pilot studies published in 2000-2001 in selected journals*

Pilot study	BMJ	Lancet	JAMA	NEJM	BJC	BJOG	BJS	Total
Pilot in preparation for a trial	0	0	0	0	0	3 (3)	1 (1)	4 (4)
Piloting new treatment, technique, combination of treatments, Phase I/II trials	5 (3)	11 (8)	4 (1)	3	28 (25)	5 (1)	7 (1)	63 (39)
Piloting screening programme	1	3 (2)	0	0	1	0	0	5 (2)
Piloting guidelines, educational package, patient care strategy	5 (1)	1	2	0	0	2	1	11 (1)
Laboratory testing of activity of compounds eg. in vivo or in vitro assays	0	2 (1)	1	0	4	0	0	7 (1)
Total pilot studies	11 (4)	17 (11)	7 (1)	3	33 (25)	10 (4)	9 (2)	90 (47)
Total number of research papers**	372	1115	619	434	1132	381	396	4449

*Numbers in parentheses refer to the number of studies that mentioned the need for further study as a result of the findings of the pilot study.

**This is an approximate total, referring to a search of the total number of journal articles containing an abstract, excluding reviews, using PubMed (National Center for Biotechnology Information 2002).

Evidence-based key objectives

i. Test integrity of study protocol
ii. Sample size calculation
iii. Pilot data collection forms/questionnaires
   - Prepare and plan data collection and monitoring
iv. Acceptability of the intervention
   - Develop and test implementation and delivery of the intervention
   - Train staff in delivery and assessment
v. Selection of most appropriate outcome measures (endpoints)
vi. Recruitment and consent rates
vii. Randomisation procedure

(i) Integrity of study protocol

- Eg. In preparation for large multi-centre trial
- Randomised pilot study
- Enables all procedures to be put in place and tested
  - inclusion/exclusion criteria
  - drug preparation (if applicable)
  - storage and testing of equipment and materials
  - training of staff in administration
  - assessment of the intervention enrolment procedure
  - determine the number of research assistants necessary to provide 24 hour on-call cover

(ii) Sample size calculation

- Common reason for pilot study
- Need estimates for control group:
  - location (mean) and variability (sd)
  - proportion

  of primary outcome/endpoint
- Rule of thumb: need at least 30 patients (Browne 1995)
- Will be covered in later session

(iii) Testing data collection forms and questionnaires

- Particularly important when the patient has to self-complete a form or when several different assessors
- Ensures form is comprehensible and questions are well-defined, clearly understood and presented in a consistent manner
- Other forms such as patient information documents and consent forms can also be tested

  NB. Testing administration of a questionnaire is not the same as validating the instrument (see point v)

(iv) Acceptability of intervention

- When intervention may not appeal to all patients, it is wise to determine its acceptability
  eg. known side effects, difficult to administer, complementary therapy
- Of particular benefit in a paediatric population when drugs may be licensed and tested in adults but not necessarily in children, or when children need to stick to a dietary regime

(v) Selection of appropriate outcome measure(s)

- Distinguish between primary and secondary outcome measures
- Valid and reliable (repeatable & reproducible)
- Directly measured vs patient-reported
  - Include additional objective measures when self-reporting may be unreliable eg. self-assessed smoking cessation and biochemical measure
  - HRQL – use generic and disease-specific measure
- Individual level vs group (cluster) level
- Select most appropriate outcome for evaluating the effectiveness of the intervention

  eg. level of knee pain, knee function, ability to work, satisfaction with treatment

(vi) Recruitment

- Successful recruitment requires a co-ordinated approach and good pilot work
- Important to engage centres or practices early-on
  - Is research question important for Prim/Sec Care?
  - What is its priority compared to other issues?
  - How does it impact on patient-doctor relationship?
  - Is doctor confident to raise research issue within a sensitive consultation?
- Time constraints are a major issue
- Need to find efficient ways to identify the sample and gain consent
- Complex interventions can have different levels of recruitment (eg. practices & patients)

Principles of good recruitment

- Engage with all stakeholders (Clinicians, GPs, practice staff and participants)
  - Brand for trial (eg. BEAM, PANDA, SCAMPS)
  - Well-developed marketing strategy, good PR
    eg. Bell's Palsy trial used local celebrity in media
  - Well-written patient information documents
- Invitation to take part coming from own doctor
- Use trained staff other than doctor/GP to identify and consent participants eg. practice nurses
- Provide staff training in disease topic and research
- Get support from local research network eg. PCRN
  - 'Research Ready' accreditation scheme for GP practices
  - ePCRN (www.ePCRN.net now the TRANSFoRm EU project)
- NB. Participants are allowed to opt-out

(vii) Method of randomisation

- Test out randomisation procedure
  - By individual or by cluster eg. GP practices, households, nursing homes
  - relative costs and justification
- If CRT usually have relatively fewer clusters than individuals ➔ higher prob. of imbalance
  - in the size of each treatment arm
  - in baseline covariate distributions at individual level
- Complex interventions may have multiple components
  eg. simple parallel design vs factorial design

Randomisation procedure

- Test how the randomisation procedure is to work
- Preparation and storage of sealed envelopes
- Administration eg. through a hospital pharmacy where each envelope could be signed for at the pharmacy window to maintain objectivity
- Use of a specialist clinical trials unit to provide 24-hour randomisation service, or to provide phone coverage from 9am to 5pm
- Test acceptability of the concept of randomisation to the patient and best way of providing a suitable explanation and eliciting informed consent

Example 1 – UK BEAM trial

- UK Back Pain, Exercise, Active management and Manipulation trial (Farrin et al. 2005)
- To test the integrity of the study protocol using a series of sub-studies
- Planned as cluster randomised trial
- 3 treatments – active management (practice level); spinal manipulation and exercise (patient level) – 3 x 2 x 2 factorial design
- Qualitative and quantitative pilot work
  - Views, acceptability and needs of support staff
  - Sample size, staff training, data collection processes, treatment delivery

Example 1 cont.

Findings:

- Majority of methods were successful but highlighted where changes were needed
- Problem with differential recruitment between practices
- Twice as many recruited to intervention arm (active management) than control
- Less severe back pain, less depression, higher education, more in full-time work in intervention group than control at baseline

➔ changed to non-clustered design

Example 2 – Antibiotics use

- Optimising antibiotic use in nursing homes (Loeb, 2002)
- To develop diagnostic and treatment algorithms for use in delivering the intervention in nursing homes
- Multifaceted intervention to reduce prescriptions for antimicrobials for suspected urinary tract infections
- Randomised matched-pairs design
- Systematic review of literature, qualitative study to assess attitudes and perceptions

Example 2 cont.

- Findings:
  - Poor adherence to the algorithms in the nursing homes
  - Changed 'training the trainer' approach – used standardised training by research team rather than infection control practitioners to train nursing staff
  - Introduced regular on-site visits by research team to aid adherence to treatment algorithms

  ➔Developed the study protocol following the MRC complex intervention guidelines

  ➔ Protocol was published in BMC Health Services Research

Conclusion

- Specific aims and objectives of feasibility/pilot studies should be clearly presented
- Place definitions within a wider conceptual framework
- Methodologically rigorous framework safeguards against pilot studies being conducted simply because of small numbers
- Need guidelines for reporting feasibility/pilot studies

Discussion of the objectives in participants' examples

Pilot and feasibility studies
Sample size

Mike Campbell

Basic premise

- Since hypothesis testing of whether the intervention differs from control is not appropriate in a pilot study, power based sample size calculations are *not* appropriate

- However, we still need a sample size *justification*

NIHR Guidance re sample size

"Instead the sample size should be adequate to estimate the critical parameters (e.g. recruitment rate) to the necessary degree of precision.“

http://www.nihr.ac.uk/CCF/RfPB/FAQs/Feasibility_and_pilot_studies.pdf

External/Internal?

- External – when questions still exist over items such as the exact form of the intervention or what outcome measures are to be chosen, so that patients in the pilot will not be comparable to patients in the main study.
- Internal- when the intervention and the outcome measures are fixed, but questions remain as to , for example, the variance of the outcome measure , the recruitment rate or the drop out rate.

Is sample size a problem? Vickers et al (2003)

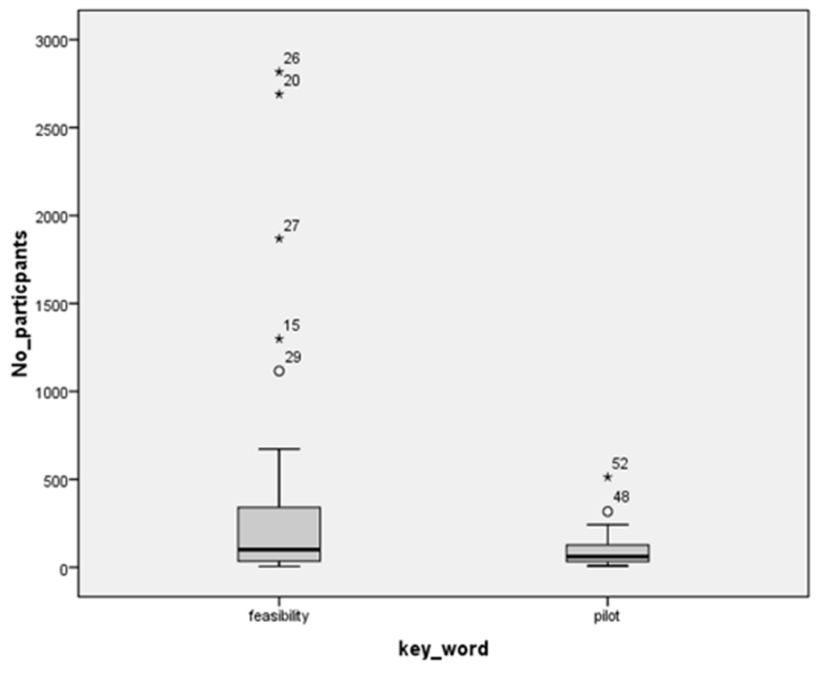
- A systematic review of published RCTs with continuous outcomes found evidence that the population variation was underestimated in 80% of reported endpoints in the sample size calculations compared to the variation observed when the trial was completed.
- They also found that 25% of studies were vastly underpowered and would have needed five times the sample size if the variation observed in the trial had been used in the sample size calculation.

How many p/f studies have used sample size calculations ?

Review by Arain et al (2010)
Literature review in 2007/8 using key words Pilot or Feasibility
(Select journals, 54 papers)

- 7/20 (35%) 'Pilot' studies used a sample size calculation
- 3/34 (8%) 'Feasibility' studies used a sample size calculation
- 9/26 (35%) papers described as 'pilot or feasibility studies for RCTs' including Phase II studies used a sample size calculation

Sample sizes of ‘pilot’ or ‘feasibility’ studies Arain et al



- **Feasibility studies bigger than pilot**
- **Upper limit 3000 for feasibility study**
- **Upper limit 500 for a pilot**

Hertzog (2008)

Research in Nursing & Health, 2008, 31, 180–191

Considerations in Determining Sample Size for Pilot Studies

Melody A. Hertzog*

College of Nursing, University of Nebraska Medical Center, Lincoln Division, 1230 "O" Street, Suite 131, P.O. Box 880220, Lincoln, NE 68588-0220

Accepted 12 August 2007

Abstract: There is little published guidance concerning how large a pilot study should be. General guidelines, for example using 10% of the sample required for a full study, may be inadequate for aims such as assessment of the adequacy of instrumentation or providing statistical estimates for a larger study. This article illustrates how confidence intervals constructed around a desired or anticipated value can help determine the sample size needed. Samples ranging in size from 10 to 40 per group are evaluated for their adequacy in providing estimates precise enough to meet a variety of possible aims. General sample size guidelines by type of aim are offered. Res Nurs Health 31:180–191, 2008

Review by Hertzog (2008)

Medline search in 2004

96 studies met criteria of pilot.

- Total sample sizes ranged from 3 to 419,
  median size 34.5.
- Those involving single groups, 13 were purely psychometric studies (median size 84),
- 35 were correlational/descriptive (median size 40)
- 21 were feasibility or efficacy studies (median size 18).
- 24 were two group comparisons -median size 20.5

Billingham(2013) – UKCRN audit

Billingham *et al. BMC Medical Research Methodology* 2013, **13**:104
http://www.biomedcentral.com/1471-2288/13/104

BMC
Medical Research Methodology

RESEARCH ARTICLE Open Access

An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database

Sophie AM Billingham[1], Amy L Whitehead[2] and Steven A Julious[2*]

Sample sizes in Billingham et al n= 79 trials

Table 2 Median sample size per arm according to type of study, funder and endpoint

		Sample size per arm		
		n	Median	(IQR) [Range]
Trial description	Pilot	50	30	(20, 45) [8, 114]
	Feasibility	25	36	(25, 50) [10, 300]
	Both	4	49	(36, 61) [23, 72]
Type of endpoint	Dichotomous	31	36	(25, 50) [10, 300]
	Continuous	45	30	(20, 50) [8, 114]
Funder	Industry	13	30	(16, 31) [8, 100]
	Public	47	36	(25, 60) [10, 300]
	Charity	19	30	(20, 45) [15, 52]

Advice from the literature

- Browne (1995) gave as a general rule to take a minimum of 30 patients to estimate a parameter
- Julious (2005) recommends a minimum sample size of 12 per group as a rule of thumb and justifies this based on rationale about feasibility and precision about the mean and variance;
- Hertzog (2008) gave a general discussion of different aspects of sample size. Suggested total pilot sample sizes of 20–40
- Stallard (2012) proposed that the sample size should be approximately 0.03 times that the sample size planned for the definitive study
- Sim and Lewis (2012) suggest a sample size of at least 50 per group based on upper CI of variance estimate
- Cocks and Torgerson(2013) suggest 9% of the sample size of the main planned study
- Teare et al (2014) suggest 35 per group to estimate SD or 60-100 per group for event rate

  "It is very much more efficient to use a larger pilot study, than to guard against the lack of precision by using inflated estimates

Teare *et al. Trials* 2014, **15**:264
http://www.trialsjournal.com/content/15/1/264

TRIALS

RESEARCH Open Access

Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study

M Dawn Teare*, Munyaradzi Dimairo, Neil Shephard, Alex Hayman, Amy Whitehead and Stephen J Walters

Abstract

Background: External pilot or feasibility studies can be used to estimate key unknown parameters to inform the design of the definitive randomised controlled trial (RCT). However, there is little consensus on how large pilot studies need to be, and some suggest inflating estimates to adjust for the lack of precision when planning the definitive RCT.

Methods: We use a simulation approach to illustrate the sampling distribution of the standard deviation for continuous outcomes and the event rate for binary outcomes. We present the impact of increasing the pilot sample size on the precision and bias of these estimates, and predicted power under three realistic scenarios. We also illustrate the consequences of using a confidence interval argument to inflate estimates so the required power is achieved with a pre-specified level of confidence. We limit our attention to external pilot and feasibility studies prior to a two-parallel-balanced-group superiority RCT.

Results: For normally distributed outcomes, the relative gain in precision of the pooled standard deviation (SD_p) is less than 10% (for each five subjects added per group) once the total sample size is 70. For true proportions between 0.1 and 0.5, we find the gain in precision for each five subjects added to the pilot sample is less than 5%

Need a compromise

- A *small* pilot will have large uncertainty, which has to be allowed for in main trial which may result in a larger trial than necessary
- A *large* pilot will have less uncertainty and so the main trial will be better planned, *but* may be less efficient overall.

Whitehead et al (in press) Rules of thumb

The fixed rules of thumb will have times when they will perform well in terms of total sample size of the the pilot and the main trial together and times when they result in a larger total sample size. It depends partly on the Effect size

Estimated stepped rules of thumb for required pilot trial sample size per treatment arm using the Non Central T approach to calculate the main trial sample size

Standardised Effect size	80% Powered Main Trial	90% Powered Main Trial
Extra Small <0.1	50	75
Small 0.2	20	25
Medium 0.5	10	15
Large 0.8	10	10

Can we use a pilot study to estimate an effect size?

- Effect sizes are not “what we expect” but rather what is clinically important
- Problem is that on occasion clinicians don’t know what is “clinically important”
- Usually a pilot is not our only source of information – should combine information from pilot with prior data
- Evidence from Kraemer et al (next slide)

A warning
Don’t use pilot studies to determine effect sizes

PERSPECTIVES

Caution Regarding the Use of Pilot Studies to Guide Power Calculations for Study Proposals

Helena Chmura Kraemer, PhD; Jim Mintz, PhD; Art Noda, MS; Jared Tinklenberg, MD; Jerome A. Yesavage, MD

Clinical researchers often propose (or review committees demand) pilot studies to determine whether a study is worth performing and to guide power calculations. The most likely outcomes are that (1) studies worth performing are aborted and (2) studies that are not aborted are underpowered. There are many excellent reasons for performing pilot studies. The argument herein is not meant to discourage clinical researchers from performing pilot studies (or review committees from requiring them) but simply to caution against their use for the objective of guiding power calculations. *Arch Gen Psychiatry. 2006;63:484-489*

Kraemer et al results

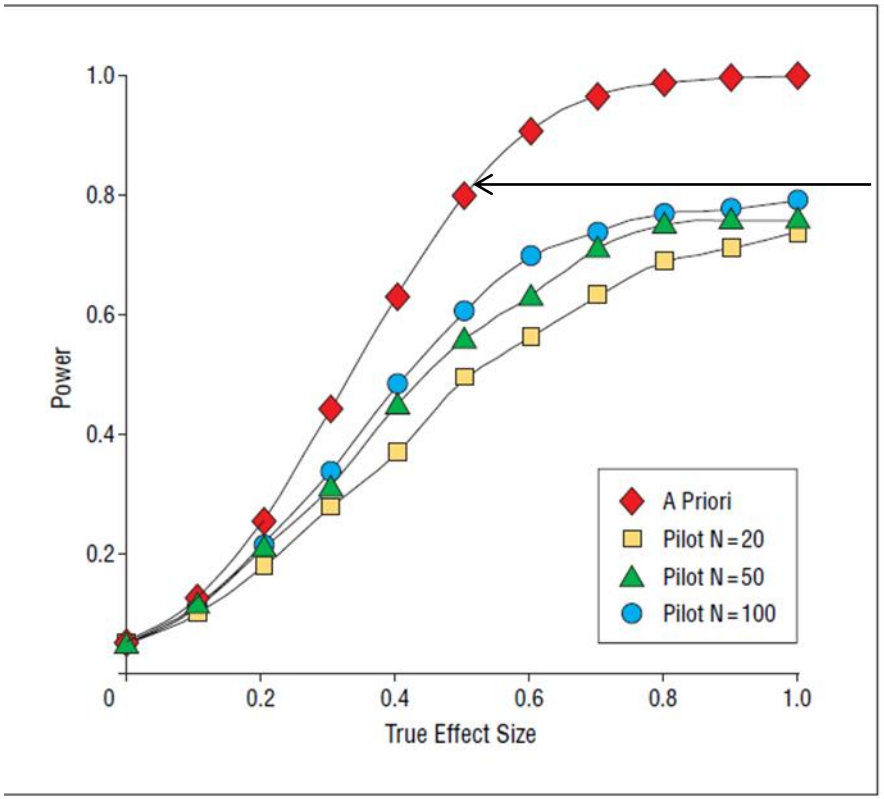


Figure 1. The power (probability of finding a statistically significant result) using a 1-tailed 5% test when the desired power to detect the effect size defining the threshold of clinical significance (δ*) is 0.5. The upper curve shows results using a priori power calculations. The 3 lower curves show results when the power calculation is based on effect sizes from pilot studies with sample sizes of 20, 50, and 100.

Sample size determined from power=80% , one sided significance 0.05 and standardised effect size of 0.5 (N=100)

For the non-red lines the effect size is determined from a pilot n=20,50,100, from a Population with standardised Effect size 0.5

Questions for discussion

- *What about studies called 'pilot' because the outcome is a surrogate or interim outcome? Do they need sample size calculations?*
- *Should we ever estimate the effect size in pilot studies?*
- *Is it sensible to use a larger external pilot, as suggested by Teare et al(2014) rather than allowing for uncertainty of estimates in a smaller pilot, considering that patients in the external pilot are lost for estimation of the true effect?*
- *If we are estimating a parameter (eg rate such as willingness to be randomised) what level of confidence should we use?*

Acknowledgement

- Thanks to Amy Whitehead for use of results in her thesis

References

- Arain M, Campbell MJ, Cooper CL, Lancaster GA: What is a pilot or feasibility study? a review of current practice and editorial policy. *BMC Med Res Method* 2010, 10:67
- Billingham S, An audit of sample sizes for pilot and feasibility trials being undertaken in the UK. *BMC Res Method*, 2013, 13:104
- Browne RH: On the use of a pilot sample for sample size determination. *Stat Med* 1995, 14:1933–1940.
- Cocks K amd Torgerson DJ: Sample size calculations for pilot randomized trials: a confidence interval approach. J. Clin Epi 66, 197-201.
- Hertzog MA: Considerations in determining sample size for pilot studies. *Res Nurs Health* 2008, 31:180–191
- Kraemer HC, et al Caution regarding the use of pilot studies to guide power calculations for study proposals. *Arch Gen Psychiatry,* 2006, 63, 484-489
- Julious SA, Patterson SD: Sample sizes for estimation in clinical research. *Pharm Stat* 2004, 3:213–215.
- Julious SA: Sample size of 12 per group rule of thumb for a pilot study. *Pharm Stat* 2005, 4:287–291.
- Sim J, Lewis M: The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. *J Clin Epidemiol* 2012, 65:301–308
- Teare MD et al (2014) Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study. *Trials*, 2014, 15: 264
- Vickers AJ: Underpowering in randomized trials reporting a sample size calculation. *J Clin Epidemiol* 2003, 56(8):717–720.

Analysis and progression criteria

Sandra Eldridge

EXISTING RECOMMENDATIONS

Recommendations from Lancaster 2004

The analysis of a pilot study should be mainly descriptive or should focus on confidence intervals.

Recommendations from Arain 2010

We conclude that pilot studies are still poorly reported, with inappropriate emphasis on hypothesis-testing. We believe authors should be aware of the different requirements of *pilot studies* and *feasibility studies* and report them appropriately. We found that in practice the definitions of feasibility and pilot studies are not distinct and vary between health research funding bodies and we suggest use of the NETSCC definition to clarify terminology.

Recommendations from Thabane 2010

Pilot studies should be well designed with clear feasibility objectives, clear analytic plans, and explicit criteria for determining success of feasibility. They should be used cautiously for determining treatment effects and variance estimates for power or sample size calculations. Finally, they should be scrutinized the same way as full scale studies, and every attempt should be taken to publish the results in peer-reviewed journals.

NIHR guidelines

"We expect that when pilot or feasibility studies are proposed by applicants, or specified in commissioning briefs, a clear route of progression criteria to the substantive study will be described. Listing ***clear progression criteria*** *will apply whether the brief or proposal describes just the preliminary study or both together. "*

No hypothesis tests of effect size

Normally no power calculation

Sample size too small to reliably detect important differences

A non-statistically significant result often simply reflects this

Even if randomisation has been used there may be baseline imbalances because sample size is small

But in Arain review 72% did perform hypothesis tests

Cautionary tale

- Editor 1: "…..it might be **more convincing if reported in more conventional style with p values**, appropriate attention to the calculation of sample size and both intention to treat and per protocol analyses"
- Editor 1: "….the fact remains that studies with **results that are definitive and clinically directive** will always have a better chance"
- Editor 2: "…..the **lack of objective outcomes** and the incomplete matching between groups"
- Editor 3: "We do appreciate the effort behind the study, and **its value to the scientific community**, but it can unfortunately not achieve sufficient priority to be considered"

What if a pilot/feasibility paper reports the results of a significance test?

Primary outcome	Surrogate or early outcome	Statistically significant result	No statistically significant result
Plausible sample size calculation for primary outcome	Plausible sample size calculation for surrogate or early outcome	No plausible sample size	No plausible sample size

Example (Boorgerd 2014)

Feasibility of an online treatment environment for adolescents with type 1 diabetes

62 adolescents aged 11–21 assigned to usual-care (n=31) or usual-care+intervention (n=31)

(i) Acceptability (do recipients use the intervention?)

(ii) Demand (do recipients continue to use the intervention?)

(iii) Practicability (can recipients access the intervention?)

(iv) Integration (does the intervention fit with guidelines for pediatric diabetes care?)

(v) Efficacy (what is the effect on adolescents' self efficacy?)

Objectives

(i) Acceptability (do recipients use the intervention?)

(ii) Demand (do recipients continue to use the intervention?)

(iii) Practicability (can recipients access the intervention?)

(iv) Integration (does the intervention fit with guidelines for pediatric diabetes care?)

(v) Efficacy (what is the effect on adolescents' self efficacy?)

Hypothesis testing

Assessment of efficacy revealed improvement in the intervention group in evaluation of care (Patients' Evaluation of Quality of Diabetes), $F(1,30)=5.35$, $p < 0.05$, and quality of life, communication (PedsQL), $F(1,30)=11.65$, $p < 0.05$.

No significant differences in change over time between the intervention and the control group concerning HbA1c ($F(1,61)=0.16$, $p=0.693$), confidence in diabetes self management ($F(1,61)=2.55$, $p=0.118$), diabetes knowledge ($F(1,61)=0.09$, $p=0.768$), health related quality of life ($F(1,61)=0.12$, $p=0.730$) and evaluation of diabetes care ($F(1,61)=0.08$, $p=0.781$)

Methods, data, analysis and progression criteria to match objectives

Analysis to meet other objectives

Feasibility was assessed in terms of **acceptability** and **demand**, practicability, integration, and efficacy. (Boogard 2014)

Investigator Judgement

hypothesis tests

Table 2. Total user data (in numbers) in intervention group in 9-month study phase

	Adolescents	Team
Page views (per month)	5795 (643.89)	3006 (334)
Chat visits (per month)	1050 (116.67)	253 (28.11)
Chat messages (per month)	3580 (397.78)	399 (44.33)
Initiated forum discussions (per month)	24 (2.67)	3 (0.33)
Forum messages (per month)	427 (47.44)	69 (7.67)
Initiated private discussions (per month)	24 (2.67)	47 (5.22)
Private messages (per month)	40 (4.44)	88 (9.78)

Pre-specified criteria to aid decision making about next stage

Example: DECISION+ pilot trial (Leblanc et al 2011)

Aim of main study: Optimal use of antibiotics for treating acute respiratory infections in primary care

Intervention: Education in shared decision-making among family physicians and patients

Objective of pilot trial: To assess feasibility and acceptability of study design, procedures, and intervention

Pre-specified criteria for judging whether to proceed to main trial

Family medicine groups participating >=50%

Recruited family physicians participating in all three workshops >=70%

Mean level of satisfaction from family physicians regarding the workshops >=65%

Missing data in each completed questionnaire <10%

Example result : Only 24% of family medicine groups agreed to participate

"Not reaching the pre-established criteria does not necessarily indicate unfeasibility of the trial but rather underlines changes to be made to the protocol"

Questions

- What sort of analyses are being proposed in your examples?
- Are these appropriate for addressing the stated objectives?
- How should criteria to make decisions about the next stage be chosen?
- How many criteria should there be?
- How should they be used to make the decision?

Reporting your study

Christine Bond

Aim of this session

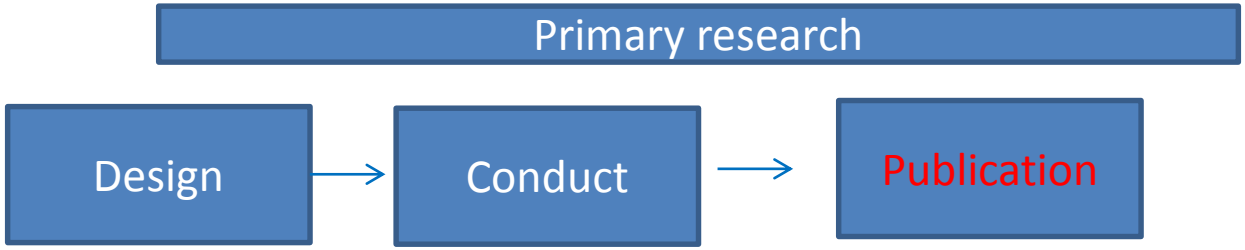
- Importance of good reporting
  - transparency in study design
- Problems of poor reporting of clinical research
  - why does it matter
- What can we do about it
  - development of reporting guidelines
  - CONSORT statement for randomized trials
  - implication for pilot randomized trials

Research article

- A published research article is a permanent record that will be used by users for many different purposes
- Some readers might be satisfied with scanning an article, or a brief summary
- Others will study it in detail for possible inclusion in a systematic review or to influence a clinical practice guideline
  - Only an adequately reported research study can be fully appraised and used appropriately
- Published research articles should be fit for multiple purposes
  - New ways of publishing (e.g., with online supplements with methodological information) can help to meet these varying needs

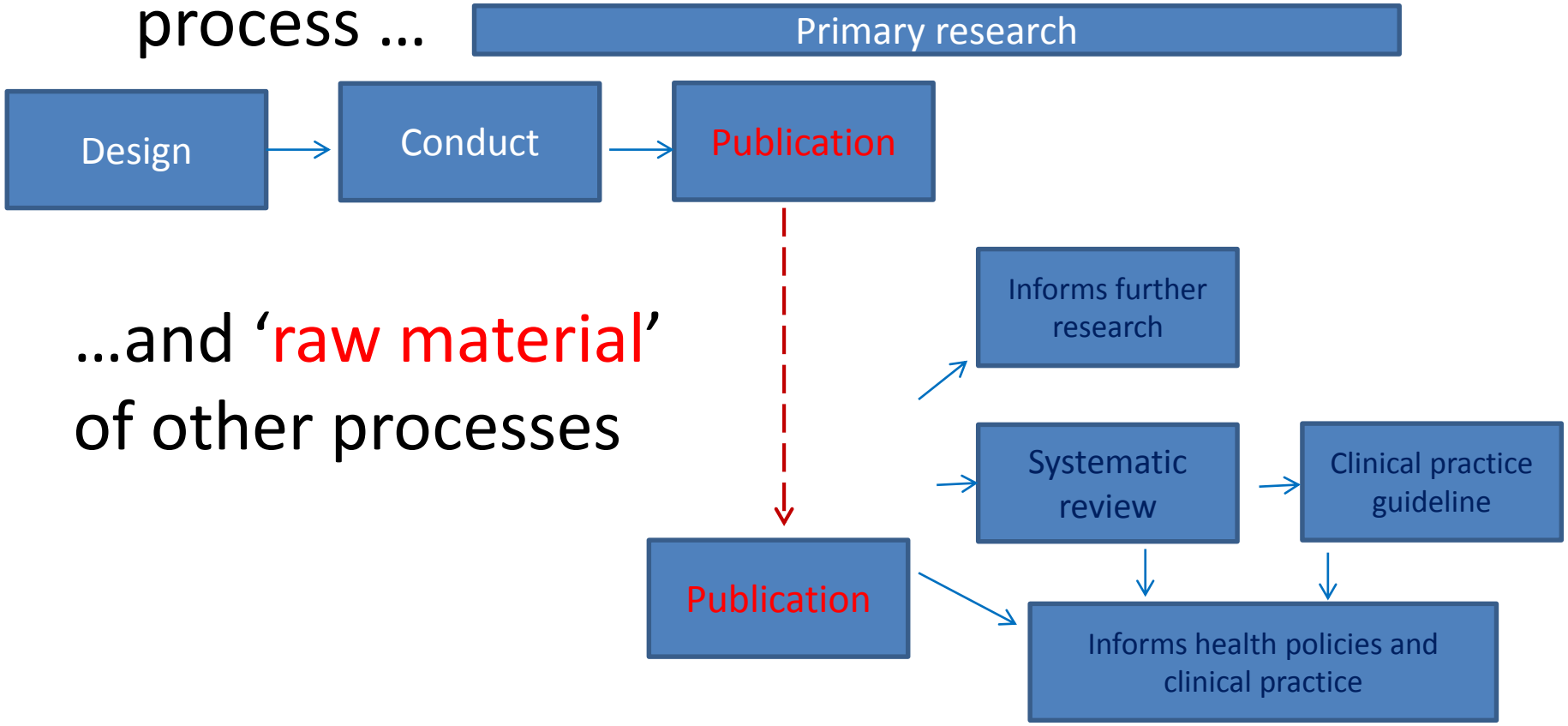
Research article

- Research article is ‘end product’ of one process …



Research article

- Research article is ‘end product’ of one process …



…and ‘raw material’ of other processes

Research article

- Scientific manuscripts should present sufficient data so that the reader can fully evaluate the information and reach his or her own conclusions about results
  - to assess reliability and relevance
- Readers need a clear understanding of exactly what was done
  - clinicians, Researchers, Systematic reviewers, Policy makers, …
- The goal should be transparency
  - should not mislead
  - should allow replication (in principle)
  - can be included in systematic review and meta-analysis

Taxonomy of poor reporting

- Non-reporting (or delayed reporting) of whole studies
  (even when some results have been presented in public)
- Misrepresentation of study design
  - e.g. study claiming is an RCT when is not
- Selective reporting
  - patient outcomes
  - analyses, e.g. subgroups, alternative analyses
- Incomplete publication
  - Omission of crucial aspects of research methods, e.g. interventions
  - Incomplete results: data cannot be included in meta-analysis
- Misleading interpretation (spin)
  - e.g. post hoc change of focus,
- Inconsistencies between sources
  - e.g. publication conflicts with protocol

In simple terms…

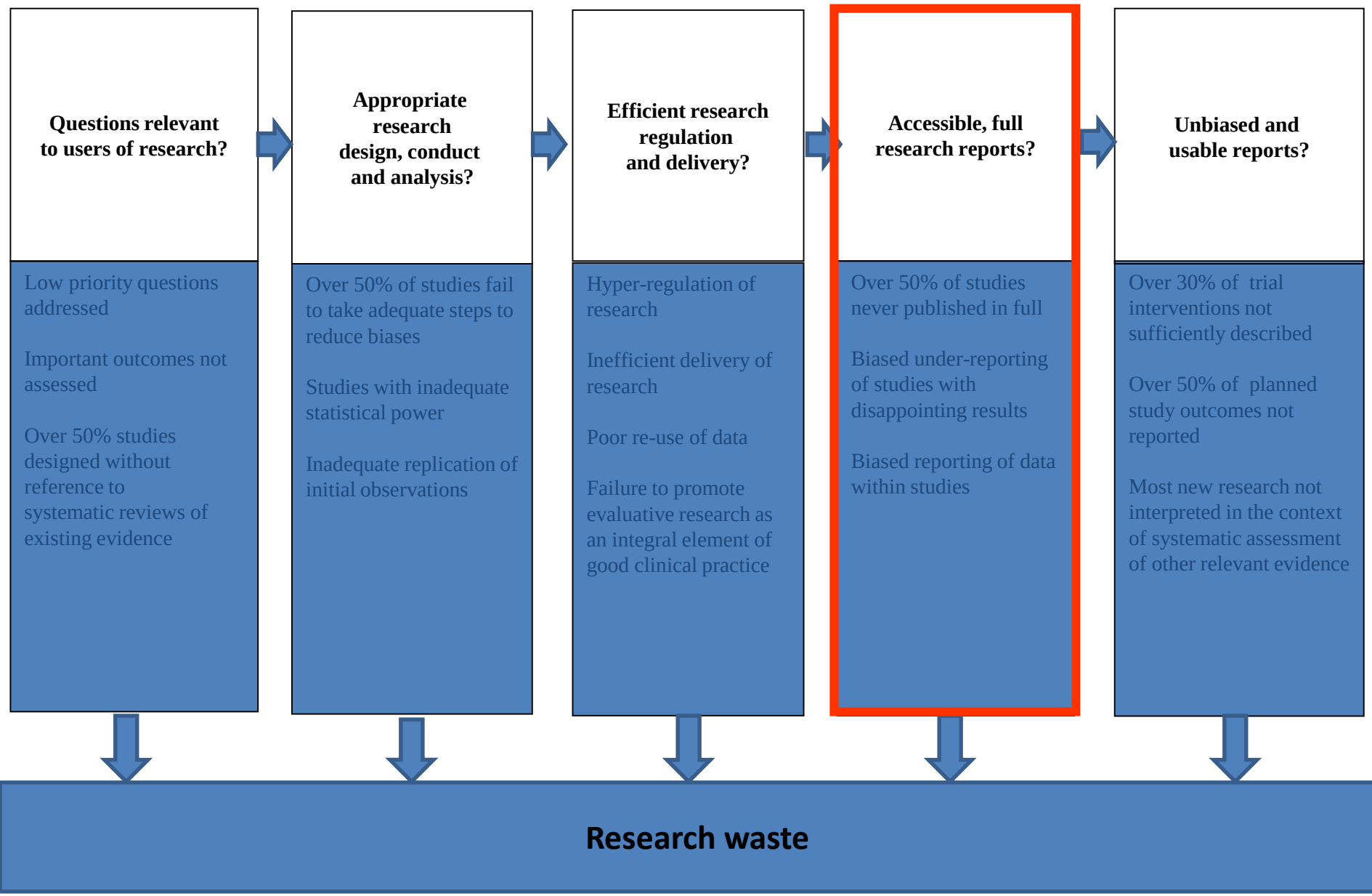
- Non-reporting
- Selective reporting
- Poor reporting

All are very common

Evidence of poor reporting

- There is considerable evidence that many published articles omit vital information
  - Hundreds of reviews of published research articles
- We often cannot tell exactly how the research was done
- These problems are generic
  - not specific to randomised trials
  - not specific to studies of medicines
  - not specific to research by pharmaceutical companies
  - … may be a particular problem for pilot studies

Avoidable waste in deciding what research to do, Lancet series, 2014



Whose fault is poor reporting?

- Poor reporting indicates a collective failure of authors, peer reviewers, and editors
  - … on a massive scale
- Researchers (authors) may not know what information to include in a report of research
- Editors may not know what information should be included

What help can be given to authors?

What help can be given to editors?

Welcome to the CONSORT Website

CONSORT stands for Consolidated Standards of Reporting Trials and encompasses various initiatives developed by the CONSORT Group to alleviate the problems arising from inadequate reporting of randomized controlled trials.

The CONSORT Statement

The main product of CONSORT is the CONSORT Statement, which is an evidence-based, minimum set of recommendations for reporting randomized trials. It offers a standard way for authors to prepare reports of trial findings, facilitating their complete and transparent reporting, and aiding their critical appraisal and

CONSORT 2010 Key Documents

- CONSORT 2010 Checklist
- CONSORT 2010 Flow Diagram
- CONSORT 2010 Statement
- CONSORT 2010 Explanation and Elaboration Document

Recent Tweets

Follow Us

EQUATOR Network @EQUATORNetwork 3 Mar
RT @trished: Declaration of transparency Now adopted by several journals as well as @bmj_latest & @BMJ_Open
equator-network.org/2014/08/12/dec...
Retweeted by CONSORT Statement

Special Communication

Improving the Quality of Reporting of Randomized Controlled Trials

The CONSORT Statement

Colin Begg, PhD; Mildred Cho, PhD; Susan Eastwood, ELS(D); Richard Horton, MB; David Moher, MSc; Ingram Olkin, PhD; Roy Pitkin, MD; Drummond Rennie, MD; Kenneth F. Schulz, PhD; David Simel, MD; Donna F. Stroup, PhD

JAMA, August 28, 1996

CONSORT statement

The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials

*David Moher, Kenneth F Schulz, Douglas G Altman, for the CONSORT Group**

Lancet 2001: **357:** 1191–94

RESEARCH METHODS & REPORTING

CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials

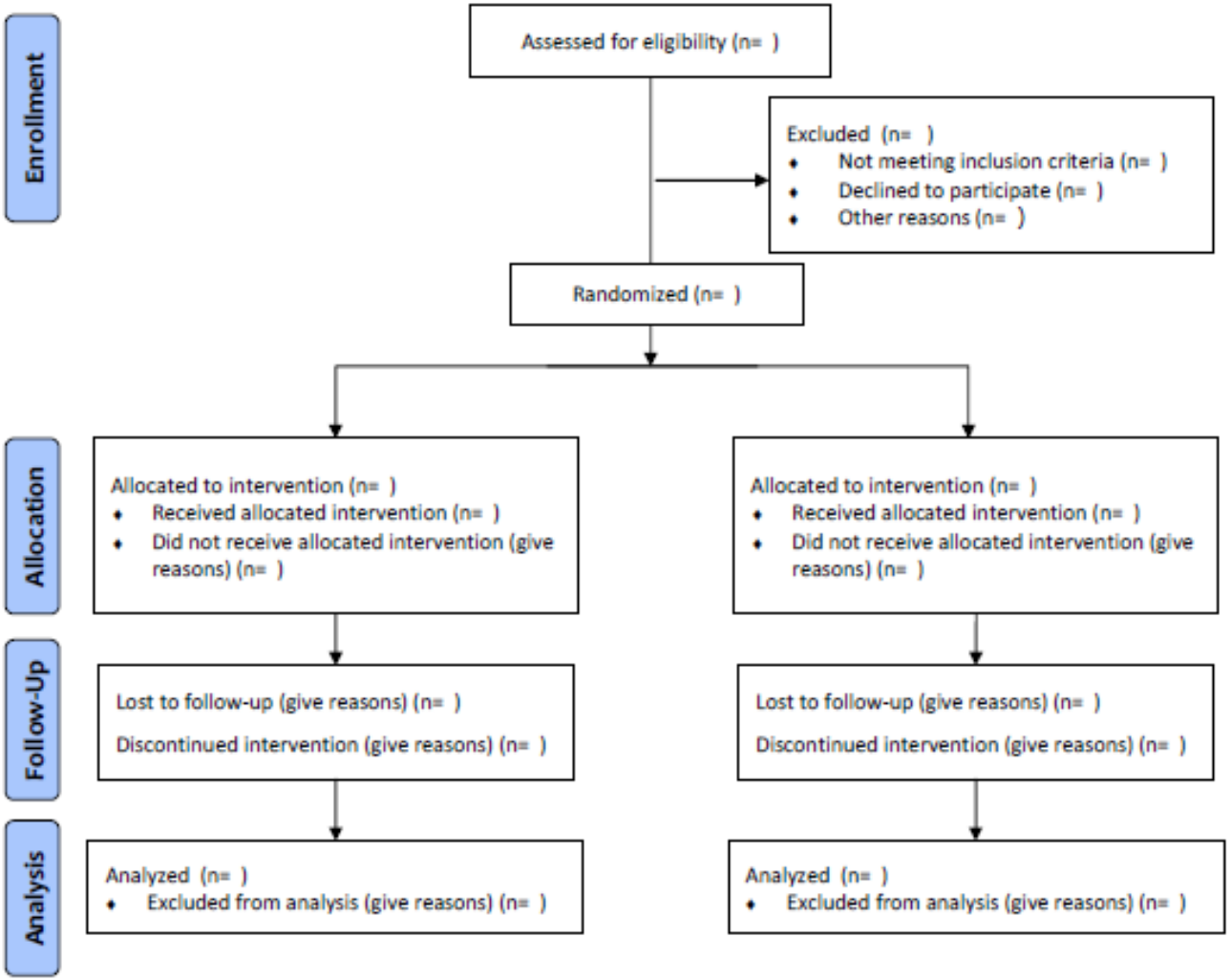
David Moher,[1] Sally Hopewell,[2] Kenneth F Schulz,[3] Victor Montori,[4] Peter C Gøtzsche,[5] P J Devereaux,[6] Diana Elbourne,[7] Matthias Egger,[8] Douglas G Altman[2]

CONSORT: checklist (25 items)

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	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)	
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	
	2b	Specific objectives or hypotheses	
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	
Participants	4a	Eligibility criteria for participants	
	4b	Settings and locations where the data were collected	
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	
Outcomes	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	
Sample size	7a	How sample size was determined	
	7b	When applicable, explanation of any interim analyses and stopping guidelines	
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	
Allocation concealment mechanism	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	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	

CONSORT 2010 flow diagram



CONSORT checklist 2010 (25 items)

TITLE & ABSTRACT

INTRODUCTION

- Background
- Objectives

METHODS

- Trial design
- Participants
- Interventions
- Outcomes
- Sample size
- Randomization
  - **Sequence generation**
  - **Allocation concealment**
  - **Implementation**
- Blinding (Masking)
- Statistical methods

RESULTS

- Participant flow
- Recruitment
- Baseline data
- Numbers analyzed
- Outcomes and Estimation
- Ancillary analyses
- Harms

DISCUSSION

- Limitations
- Generalisability
- Interpretation

OTHER INFORMATION

- Registration
- Protocol
- Funding

CONSORT items and examples

Methods

Trial design

3a Description of trial design (such as parallel, factorial) including allocation ratio

3b Important changes to methods after trial commencement (such as eligibility criteria), with reasons

Participants

4a Eligibility criteria for participants

4b Settings and locations where the data were collected

CONSORT items and examples

Methods

Item 3a. Description of trial design (such as parallel, factorial) including allocation ratio

Example—"This was a multicenter, stratified (6 to 11 years and 12 to 17 years of age, with imbalanced randomisation [2:1]), double-blind, placebo-controlled, parallel-group study conducted in the United States (41 sites)."[85]

Item 3b. Important changes to methods after trial commencement (such as eligibility criteria), with reasons

Example—"Patients were randomly assigned to one of six parallel groups, initially in 1:1:1:1:1:1 ratio, to receive either one of five otamixaban ... regimens ... or an active control of unfractionated heparin ... an independent Data Monitoring Committee reviewed unblinded data for patient safety; no interim analyses for efficacy or futility were done. During the trial, this committee recommended that the group receiving the lowest dose of otamixaban (0·035 mg/kg/h) be discontinued because of clinical evidence of inadequate anticoagulation. The protocol was immediately amended in accordance with that recommendation, and participants were subsequently randomly assigned in 2:2:2:2:1 ratio to the remaining otamixaban and control groups, respectively."[86]

CONSORT items and examples

Methods

Interventions

5 The interventions for each group with sufficient details to allow replication, including how and when they were actually administered

Outcomes

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 items and examples

Methods

Item 6a. Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed

Example—"The primary endpoint with respect to efficacy in psoriasis was the proportion of patients achieving a 75% improvement in psoriasis activity from baseline to 12 weeks as measured by the PASI [psoriasis area and severity index] Additional analyses were done on the percentage change in PASI scores and improvement in target psoriasis lesions."[103]

Item 6b. Any changes to trial outcomes after the trial commenced, with reasons

Example—"The original primary endpoint was all-cause mortality, but, during a masked analysis, the data and safety monitoring board noted that overall mortality was lower than had been predicted and that the study could not be completed with the sample size and power originally planned. The steering committee therefore decided to adopt co-primary endpoints of all-cause mortality (the original primary endpoint), together with all-cause mortality or cardiovascular hospital admissions (the first pre-specified secondary endpoint)."[112]

CONSORT items and examples

Methods

Randomisation

- Sequence generation

8a Method used to generate the random allocation sequence

8b Type of randomisation; details of any restriction (such as blocking and block size)

- Allocation concealment mechanism

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

CONSORT items and examples
Methods

Item 8a. Method used to generate the random allocation sequence

Examples—"Independent pharmacists dispensed either active or placebo inhalers according to a computer generated randomisation list."[63]

Item 8b. Type of randomisation; details of any restriction (such as blocking and block size)

Examples—"Randomization sequence was created using Stata 9.0 (StataCorp, College Station, TX) statistical software and was stratified by center with a 1:1 allocation using random block sizes of 2, 4, and 6."[137]

CONSORT items and examples

Methods

Item 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

Examples—"The doxycycline and placebo were in capsule form and identical in appearance. They were prepacked in bottles and consecutively numbered for each woman according to the randomisation schedule. Each woman was assigned an order number and received the capsules in the corresponding prepacked bottle."[146]

CONSORT extensions

ACADEMIA AND CLINIC | Annals of Internal Medicine

CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomized Trials

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

The CONSORT (Consolidated Standards of Reporting Trials) statement is used worldwide to improve the reporting of randomized, controlled trials. Schulz and colleagues describe the latest version, CONSORT 2010, which ...

Ann Intern Med. 2010;152:726-732. www.annals.org
For author affiliations, see end of text.

* For the CONSORT Group contributors to CONSORT 2010, see the Appendix (available at www.annals.org)

BMJ

BMJ 2012;345:e5661 doi: 10.1136/bmj.e5661 (Published 4 September 2012) Page 1 of 21

RESEARCH METHODS & REPORTING

Consort 2010 statement: extension to cluster randomised trials

The Consolidated Standards of Reporting Trials (CONSORT) statement was developed to improve the reporting of ...

Reporting of Patient-Reported Outcomes in Randomized Trials

The CONSORT PRO Extension

Melanie Calvert, PhD
Jane Blazeby, MD
Douglas G. Altman, DSc
Dennis A. Revicki, MD
David Moher, PhD
Michael D. Brundage, MD
for the CONSORT PRO Group

The CONSORT (Consolidated Standards of Reporting Trials) Statement aims to improve the reporting of randomized controlled trials (RCTs); however, it lacks guidance on the reporting of patient-reported outcomes (PROs), which are often inadequately reported in trials, thus limiting the value of these data. In this article, we describe the development of the CONSORT PRO extension based on the methodological framework for guideline development pro-

RESEARCH METHODS & REPORTING

Improving the reporting of pragmatic trials: an extension of the CONSORT statement

Merrick Zwarenstein,[1,2,3] Shaun Treweek,[4,5] Joel J Gagnier,[5,6] Douglas G Altman,[7] Sean Tunis,[8,9,10] Brian Haynes,[11] Andrew D Oxman,[5] David Moher,[12,13] for the CONSORT and Pragmatic Trials in Healthcare (Practihc) groups

Pragmatic trials are designed to inform decisions about practice, but poor reporting can reduce their usefulness. The **CONSORT and Practihc groups** describe modifications to the CONSORT guidelines to help readers assess the applicability of the results

Abstract

Background
The CONSORT statement is intended to improve reporting of random ised controlled trials and focuses on minimising the risk of bias (internal validity). The applicability of a trial's results (generalisability or external validity) is also important, particularly for pragmatic trials. A pragmatic trial (a term first used in 1967 by Schwartz and Lellouch) can be broadly defined as a randomised controlled trial whose purpose is to inform decisions about practice. This extension of the CONSORT statement is intended to improve the reporting of such trials and focuses on applicability.

Methods
At two, two-day meetings held in Toronto in 2005 and 2008, we reviewed the CONSORT statement and its extensions, the literature on pragmatic trials and applicability, and our experiences in conducting pragmatic trials.

Recommendations
We recommend extending eight CONSORT checklist items for reporting of pragmatic trials: the background, participants, interventions, outcomes, sample size, blinding, participant flow, and generalisability of the findings. These extensions are presented, along with illustrative examples of reporting, and an explanation of each extension. Adherence to these reporting criteria will make it easier for decision makers to judge how applicable the results of random ised controlled trials are to their own conditions. Empirical studies are needed to ascertain the usefulness and comprehensiveness of these CONSORT checklist item extensions. In the meantime we recommend that those who support, conduct, and report pragmatic trials should use this extension of the CONSORT statement to facilitate the use of trial results in decisions about health care.

Randomised controlled trials are used to assess the benefits and harms of interventions in health care. If conducted properly, they minimise the risk of bias (threats to internal validity), particularly selection bias.[1,2] There is, however, considerable evidence that trials are not always well reported,[3,4] and this can be associated with bias, such as selective reporting of outcomes.[5]

The usefulness of a trial report also depends on the clarity with which it details the relevance of its interventions, participants, outcomes, and design to the clinical, health service, or policy question it examines. Furthermore, a trial may be valid and useful in the healthcare setting in which it was conducted but have limited applicability (also known as generalisability or external validity) beyond this because of differences between the trial setting and other settings to which its results are to be extrapolated.

Schwartz and Lellouch[6] coined the terms "pragmatic" to describe trials designed to help choose between options for care, and "explanatory" to describe trials designed to test causal research hypotheses—for example, that an intervention causes a particular biological

Table 1 | Key differences between trials with explanatory and pragmatic attitudes, adapted from a table presented at the 2008 Society for Clinical Trials meeting by Marion Campbell, University of Aberdeen

Question	Efficacy—can the intervention work?	Effectiveness—does the intervention work when used in normal practice?
Setting	Well resourced, "ideal" setting	Normal practice
Participants	Highly selected. Poorly adherent participants and those with conditions which might dilute the effect are often excluded	Little or no selection beyond the clinical indication of interest
Intervention	Strictly enforced and adherence is monitored closely	Applied flexibly as it would be in normal practice
Outcomes	Often short term surrogate or process measures	Directly relevant to participants, funders, communities, and healthcare practitioners
Relevance to practice	Indirect—little effort made to match design of trial to decision making needs of those in usual setting in which intervention will be implemented	Direct—trial is designed to meet needs of those making decisions about treatment options in setting in which intervention will be implemented

Health Services Sciences, Sunnybrook Hospital, Toronto, Ontario, Canada
Institute for Clinical Evaluative Sciences, Toronto
Department of Health Policy, Management and Evaluation, University of Toronto, Toronto
Division of International Health (IeHOH), Karolinska Institutet, Stockholm, Sweden
Clinical and Population Sciences Department, University of Dundee, Dundee
Norwegian Knowledge Centre for the Health Services, Oslo, Norway
Faculty of Medicine, University of Toronto
Centre for Statistics in Medicine, University of Oxford, Oxford
Centre for Medical Technology Policy, Baltimore, MD, USA
Division of General Internal Medicine, Johns Hopkins School of Medicine, Baltimore, MD
Center for Healthcare Policy, Stanford University School of Medicine, Palo Alto, CA, USA
Department of Clinical Epidemiology and Biostatistics and Department of Medicine, McMaster University Faculty of Health Sciences, Hamilton, ON, Canada
Clinical Epidemiology Program, Ottawa Health Research Institute, Ottawa, Canada
Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa
Correspondence to: M Zwarenstein merrick.zwarenstein@ices.on.ca

Cite this as: BMJ 2008;337:a2390 doi: 10.1136/bmj.a2390

Annals of Internal Medicine | ACADEMIA AND CLINIC

Reporting Randomized, Controlled Trials of Herbal Interventions: An Elaborated CONSORT Statement

Joel J. Gagnier, ND, MSc; Heather Boon, PhD; Paula Rochon, MD, MPH; David Moher, PhD; Joanne Barnes, PhD, MRPharmS FLS; and Claire Bombardier, MD, for the CONSORT Group*

Herbal medicinal products are widely used, vary greatly in content and quality, and are actively tested in randomized, controlled trials (RCTs). The authors' objective was to develop recommendations for reporting RCTs of herbal medicine interventions, based on the ... RCTs of herbal medicines. Item 4, concerning the herbal medicine intervention, required the most extensive elaboration. These recommendations have been developed to improve the reporting of RCTs using herbal medicine interventions.

Reporting of Noninferiority and Equivalence Randomized Trials

Extension of the CONSORT 2010 Statement

Gilda Piaggio, PhD
Diana R. Elbourne, PhD
Stuart J. Pocock, PhD
Stephen J. W. Evans, MSc
Douglas G. Altman, DSc

The CONSORT (Consolidated Standards of Reporting Trials) Statement, which includes a checklist and a flow diagram, is a guideline developed to help authors improve the reporting of the findings from randomized controlled trials. It was updated most recently in 2010. Its primary focus is on individually randomized trials with 2 parallel groups that assess the possible

OPEN ACCESS Freely available online | PLOS MEDICINE

CONSORT for Reporting Randomized Controlled Trials in Journal and Conference Abstracts: Explanation and Elaboration

Sally Hopewell[1,2*], Mike Clarke[1,3], David Moher[4,5], Elizabeth Wager[6], Philippa Middleton[7], Douglas G. Altman[2], Kenneth F. Schulz[8], and the CONSORT Group

1 UK Cochrane Centre, Oxford, United Kingdom, 2 Centre for Statistics in Medicine, Wolfson College, Oxford University, Oxford, United Kingdom, 3 School of Nursing and ... ute, Ottawa, Canada 5 Department of ... ugh, United Kingdom, 7 Discipline of ... ath Carolina, United States of America

ACADEMIA AND CLINIC | Annals of Internal Medicine

Methods and Processes of the CONSORT Group: Example of an Extension for Trials Assessing Nonpharmacologic Treatments

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*

Background: The conduct of randomized, controlled trials of nonpharmacologic treatments presents specific challenges that are not adequately addressed in trial reports.

Objective: To develop an extension of the CONSORT (Consolidated Standards of Reporting Trials) Statement for trials of nonpharmacologic treatments.

Design: A consensus meeting was organized to develop an extension of the CONSORT Statement that addresses randomized trials of nonpharmacologic treatments. To prepare for the meeting, a survey was conducted to identify the specific issues for discussion.

Setting: Consensus meeting in Paris, France.

Participants: A total of 33 experts attended the meeting. The experts were methodologists ($n = 17$); surgeons ($n = 6$); editors ($n = 5$); and clinicians involved in rehabilitation ($n = 1$), psychotherapy ($n = 2$), education ($n = 1$), and implantable devices ($n = 1$).

Measurements: Experts indicated which of the 22 items on the CONSORT checklist should be modified or which additional items should be added specifically for nonpharmacologic treatments. During a 3-day consensus meeting, all items were discussed and additional methodological issues related to nonpharmacologic research were identified.

Results: The consensus was that 11 items on the CONSORT checklist needed some modifications for nonpharmacologic trials: item 1 (title and abstract), item 3 (participants), item 4 (interventions), item 7 (sample size), item 8 (randomization), item 11 (blinding), item 12 (statistical methods), item 13 (participant flow), item 15 (baseline data), item 20 (discussion: interpretation), and item 21 (generalizability). In addition, the meeting participants added 1 item related to implementation of the intervention.

Limitation: Evidence was not always available to support the inclusion of each checklist item.

Conclusion: The methods and processes used to develop this extension could be used for other reporting guidelines. The use of this extension to the CONSORT Statement should improve the quality of reporting randomized, controlled trials assessing nonpharmacologic treatments.

Ann Intern Med. 2008;148:W-60-W-66. www.annals.org
For author affiliations, see end of text.
*For contributors to the CONSORT Extension for Nonpharmacologic Treatment Interventions, see the Appendix.

Randomized, controlled trials (RCTs) are considered the gold standard for evaluation of drugs, devices, and procedures. To help improve the quality of reporting of these trials, the CONSORT (Consolidated Standards of Reporting Trials) Statement, a 22-item checklist and flow diagram, was developed. Use of this evidence-based guideline is associated with improved quality of reporting of RCTs (1, 2). The original CONSORT Statement proposed guidelines for reporting 2-group parallel RCTs (3, 4). The CONSORT Statement has subsequently been extended to cover specific variants of this design, such as cluster randomized trials (5) and noninferiority and equivalence trials (6); certain interventions, such as herbal therapies (7); and data, such as reporting of harms (8).

Nonpharmacologic treatments cover a wide range of interventions, including surgery, technical procedures (for example, angioplasty), implanted devices (for example, pacemakers), nonimplantable devices, rehabilitation, physiotherapy, behavioral therapy, psychotherapy, and complementary and alternative medicine. Although the CONSORT Statement can be applied to reports of these trials, certain issues, such as the complexity of the intervention, expertise of the care provider, and difficulties with blinding (9), present specific challenges that the revised CONSORT Statement and the accompanying explanation and elaboration document do not address in depth (3, 4, 9–13).

Because these important study aspects are often inadequately reported (3), we developed an extension of the CONSORT Statement for trials of nonpharmacologic interventions (14–17). This article describes the methods and processes used by the CONSORT Group to develop this extension.

Methods

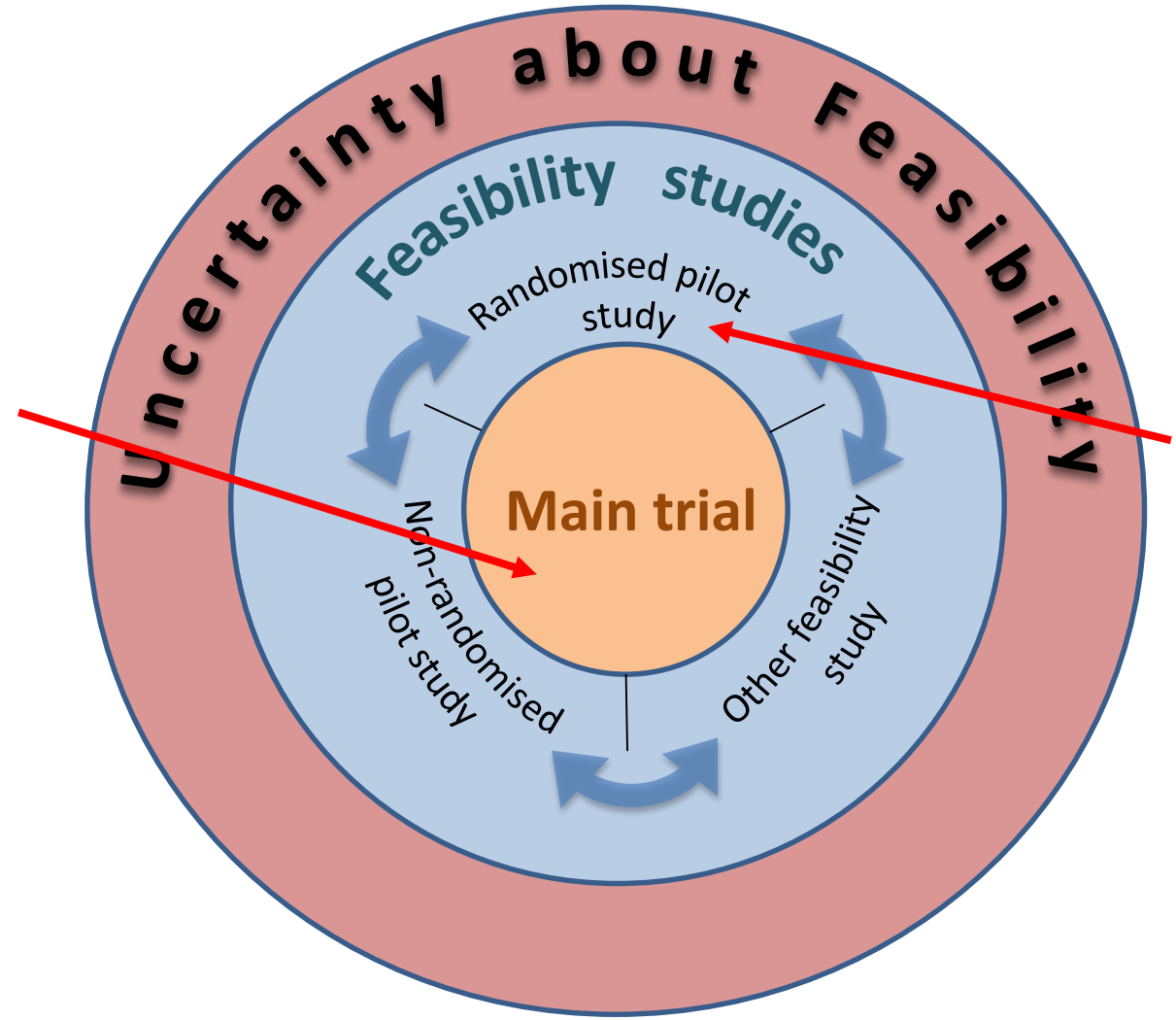
To develop the CONSORT extension for nonpharmacologic treatments, we used general guideline development principles (18) and drew on the experience gained from developing previous CONSORT extensions (19).

Steering Committee

A steering committee was ultimately responsible for the development of this reporting guide. They secured

See also:
Print
Related article 295
Web-Only
Appendix
Conversion of graphics into slides

Implications for reporting randomized pilot trials



CONSORT statement for randomized trials

CONSORT extension for pilot randomized trials

Group exercise

- Using the CONSORT checklist
  - which items would you change (modify) for reporting a randomized pilot trial?

- conducted in preparation for a future definitive trial
- primary aim is to test the feasibility of the future definitive trial

CONSORT extension for randomized pilot trials

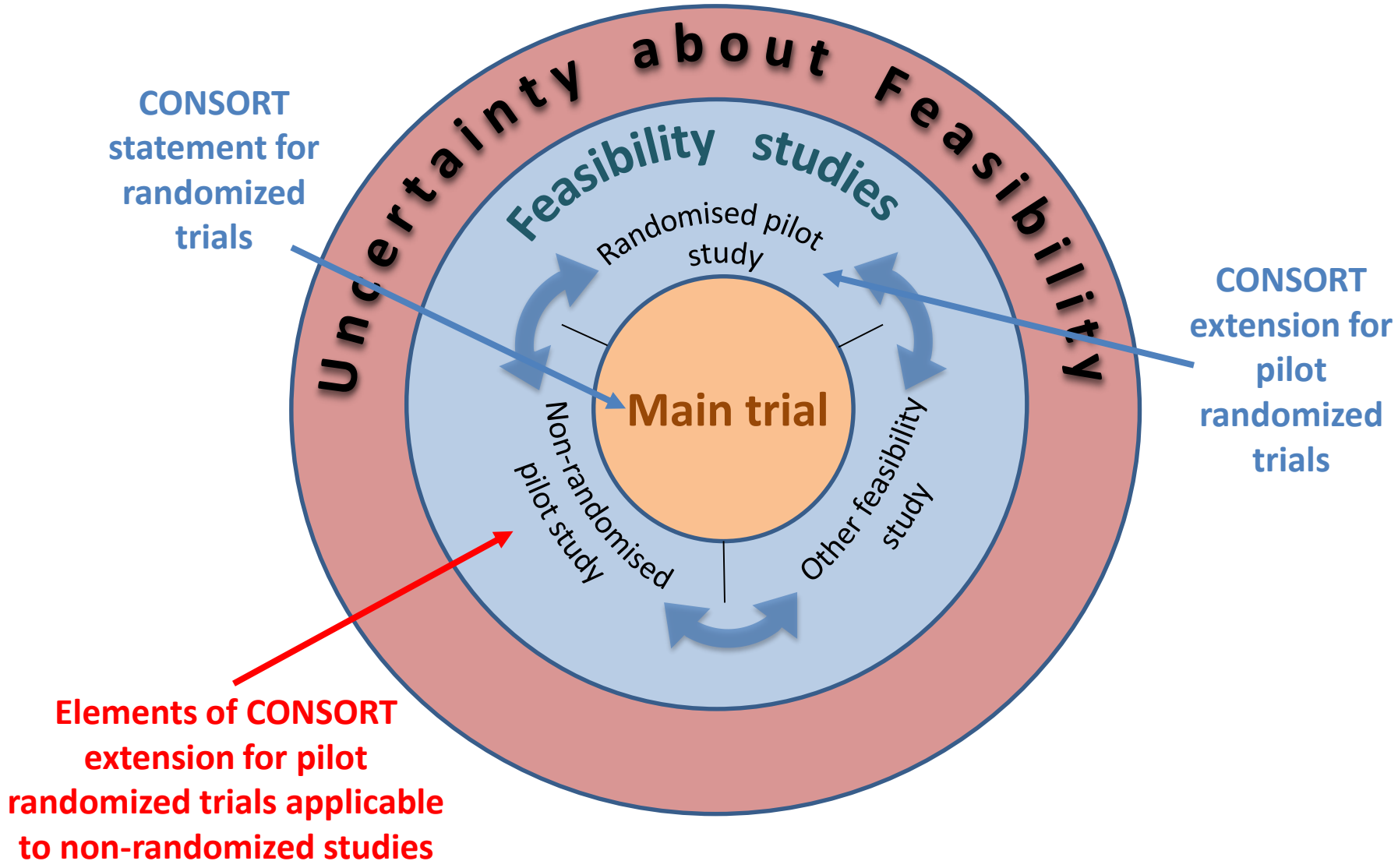
Checklist applies to:

- Randomized trials
- Conducted in preparation for a future definitive trial of effectiveness or efficacy
- Primary aim: feasibility of the future definitive trial
- No restrictions on terminology used to describe the preparatory trial
- No restrictions on the design of either trial

It does not apply to internal pilot studies.

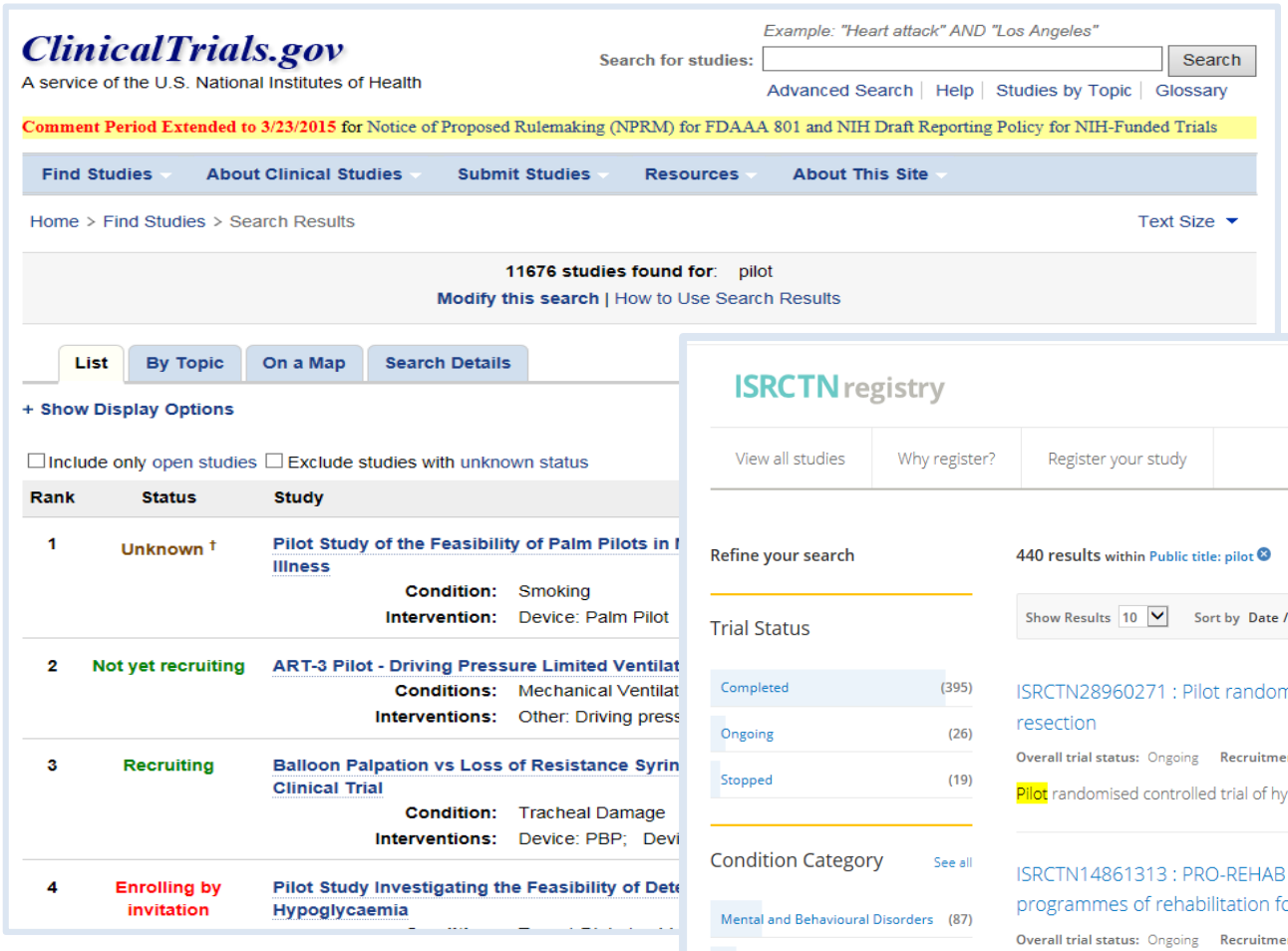
CONSORT extension for randomized pilot trials – The checklist

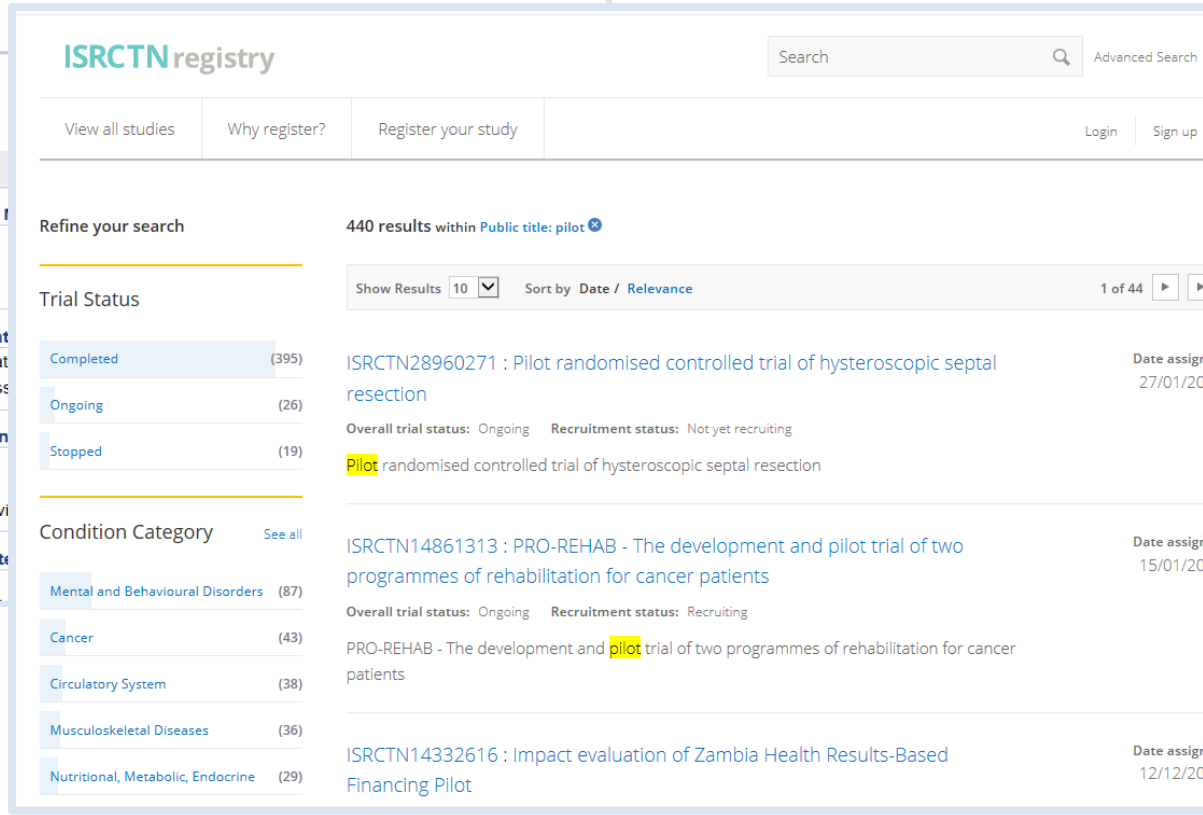
- The next few slides are not in the pack because we have not published the checklist yet


CONSORT statement for randomized trials
Uncertainty about Feasibility
Feasibility studies
Randomised pilot study
Main trial
Non-randomised pilot study
Other feasibility study
CONSORT extension for pilot randomized trials
Elements of CONSORT extension for pilot randomized trials applicable to non-randomized studies

Importance of publishing results of pilot studies

.. and prospective registration





Any questions?

Closing remarks

Gill Lancaster

Challenges with Pilot Studies

Challenges with Pilot Studies

- ✓ **Most are not well designed**
  - ❑ No clear feasibility objectives
  - ❑ No clear rationale for piloting
  - ❑ No clear analytic plans
  - ❑ No clear criteria for success of feasibility
- ✓ **Most are not reported/published**
- ✓ **It can be dangerous to use pilot studies to estimate treatment effects**
  - ❑ Estimates may be unrealistic/biased
- ✓ **If not used cautiously, results of pilot studies can potentially mislead sample size/power calculations**
  - ❑ Arch Gen Psychiatry 2006;63:484-489.

Common misconceptions

Common misconceptions

- ✓ **A small study that can be completed quickly**
  - ❑ A small study done by a student/intern
- ✓ **A small study that does not require any funding**
  - ❑ I don’t have any funding to do a big study!
  - ❑ My boss told me to do it!
- ✓ **A small study that has limited funding**
  - ❑ I have funding for only 10 patients!
  - ❑ I have limited SEED funding!

Common misconceptions

- ✓ **A small single centre study**
  - ❑ I don't have the resources for a large multi-centre study!
- ✓ **A small study that is similar in size as someone else's published study**
  - ❑ So-and-so did a similar study with 6 patients and got statistical significance – ours uses 12 patients (double the size)!
- ✓ **We did a similar pilot before (got it published!)**
  - ❑ Pilot studies should always be viewed in the context of the main study

Frequently Asked Questions

Publishing Results of Pilot Studies

- ✓ **Can I publish the results of a pilot study?**
  - ❑ Yes, every attempt should be made to publish them

- ✓ **Why is it important to publish the results of pilot studies?**
  - ❑ To provide information about feasibility to the research community
  - ❑ To save resources (avoid duplication of efforts)
  - ❑ We have ethical and scientific obligation to do so

Most pilot studies do not show statistically significant results

"no evidence of effect" is not "evidence of no effect"

BMJ 1995; 311: 485.

The focus in reporting the results of a pilot should be on **feasibility, NOT statistical significance**

New CONSORT checklist for reporting will emphasize this point

Other Important Issues

- ✓ **Can I combine data from a pilot with data from the main study?**
  - ❑ Yes, provided the sampling frame is the same and so is the methodology
- ✓ **Can I combine the results of a pilot with the results of another study or in a meta-analysis?**
  - ❑ Yes, same conditions as above
  - ❑ Also depends on whether the main study is reported
- ✓ **Can a pilot ever exist on its own?**
  - ❑ Yes, if the results show that it is not feasible to go to the main study or there is no funding for the main study
- ✓ **Can I apply for funding for a pilot study?**
  - ❑ Yes, like any grant it is important to justify the need for piloting
  - ❑ The pilot has to be placed in the context of the main study

Other Important Issues

- ✓ **Can I randomize patients in a pilot study?**
  - ❑ Yes; to assess how a randomization might work in main study or whether it might be acceptable to patients
  - ❑ In general, it is always best for a pilot to maintain the same design as the main study

- ✓ **Can I use the pilot to estimate the sample size for the main trial?**
  - ❑ Yes, but be cautious
  - ❑ Consider supplementing with qualitative discussions
  - ❑ Use SS table to capture prevailing uncertainty

- ✓ **Can I use the results of pilot study to treat my patients?**
  - ❑ Not a good idea!
  - ❑ Pilot studies are primarily for assessing feasibility

- ✓ **What can I do with a failed or bad pilot study?**
  - ❑ No study is a complete failure, it can always be used as bad example!

Our focus is on external pilots, not internal

STATISTICS IN MEDICINE, VOL. 9, 65–72 (1990)

THE ROLE OF INTERNAL PILOT STUDIES IN INCREASING THE EFFICIENCY OF CLINICAL TRIALS

JANET WITTES AND ERICA BRITTAIN

Biostatistics Research Branch, Division of Epidemiology and Clinical Applications, National Heart, Lung, and Blood Institute, Federal Building, Room 2A11, Bethesda, MD 20817, U.S.A.

Pilot vs Proof-of-concept study

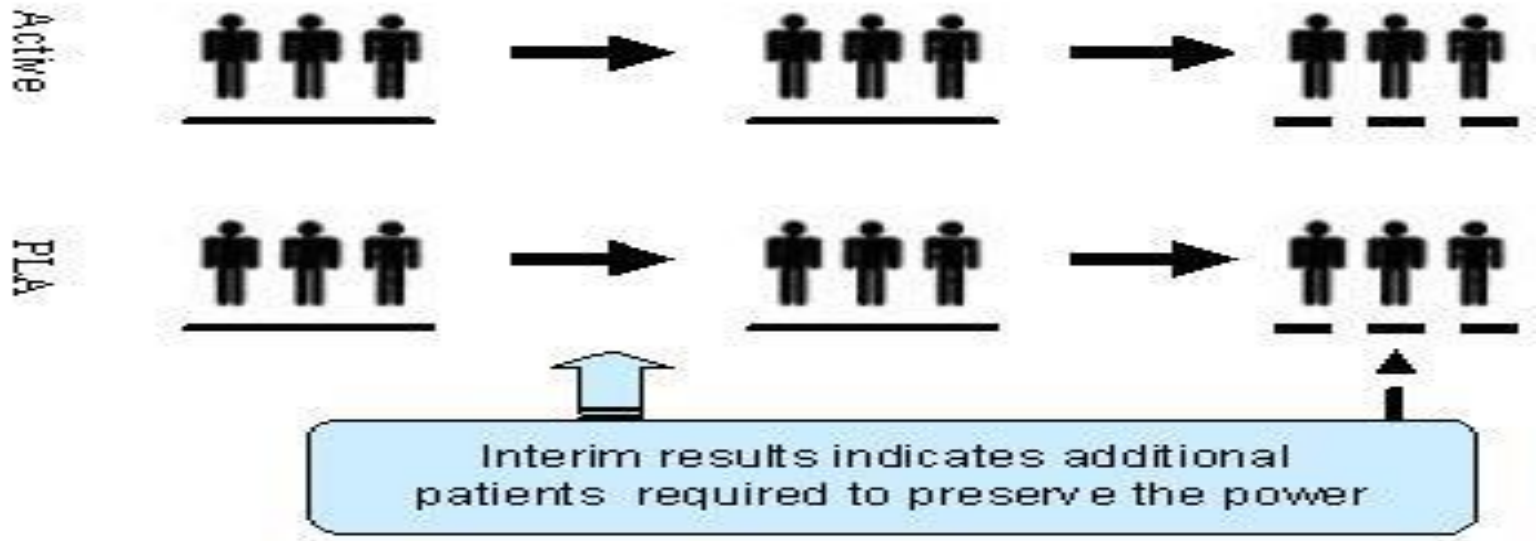
- **Proof-of-concept (POC) study:**
  - to determine if a treatment (drug) is biologically active or inactive
  - Stat Med 2005;24:1815-35
- **Usually based on surrogate makers as endpoints**
- **Usually Phase I/II studies**
  - assessing safety, dose levels and response to new drugs

Proof of concept is not necessarily proof of feasibility

Adaptive Trial Designs and Piloting

(Chow C-S, Chang M. Adaptive design methods in clinical trials – a review. *Orphanet J Rare Dis.* 2008; 3: 11)

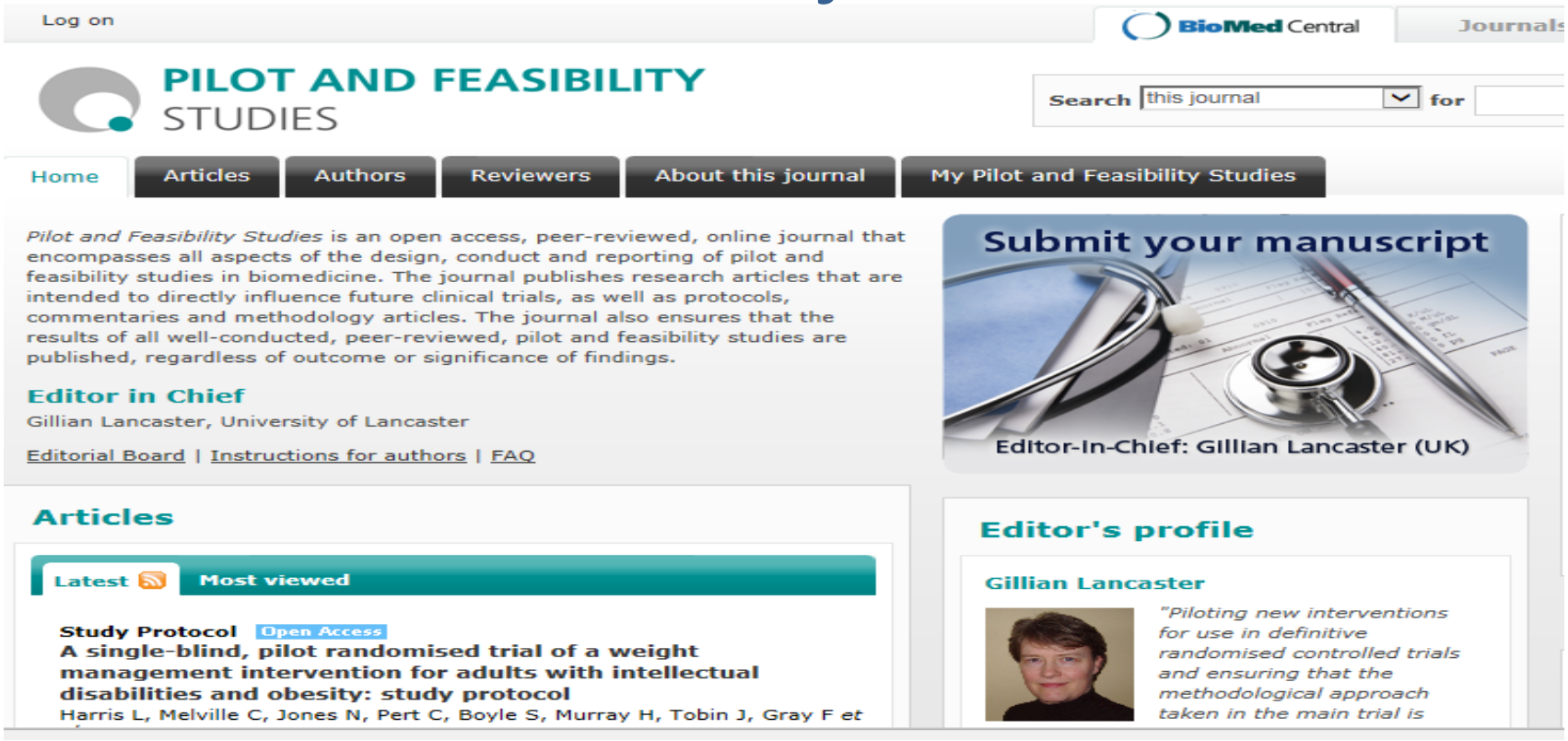
- **Adaptive trial design:** Modification or change made to
  - trial design or statistical procedures during the conduct of a clinical trial



Usually used in internal pilot studies

Designed to inform sample size calculation for the main study

A new journal



Lancaster *Pilot and Feasibility Studies* 2015, **1**:1
http://www.pilotfeasibilitystudies.com/content/1/1/1

PILOT AND FEASIBILITY STUDIES

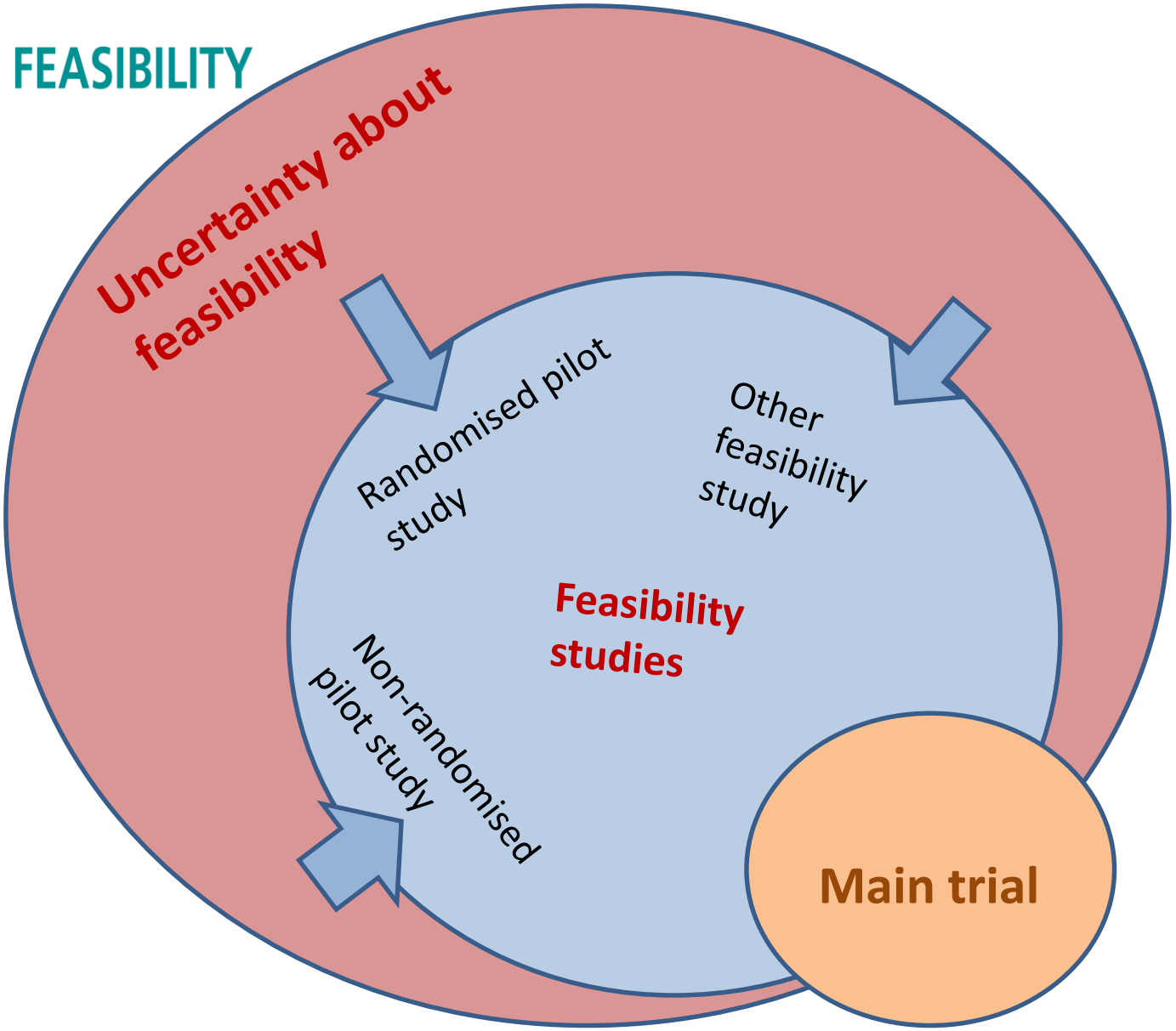
EDITORIAL Open Access

Pilot and feasibility studies come of age!

Gillian A Lancaster

Our current thinking…

PILOT AND FEASIBILITY STUDIES



African Proverb (Ashanti, Ghana)

You never test the depth of a river with both feet