# Understanding the potential for pilot and feasibility work to optimise surgical trials

**Katherine Fairhurst**

A dissertation submitted to the University of Bristol in accordance with the requirements for award for the degree Doctor of Philosophy in the Faculty of Health Sciences

**Department of Population Health Sciences**

**Bristol Medical School**

**October 2019**

Word count: 62 213

# *Abstract*

Improved surgical trial design and conduct is sought both to enhance the likelihood of main trial success and to reduce research waste. Surgical trials face unique and complex challenges that often hinder and obstruct their successful completion. Pilot and feasibility studies (PFS) are increasingly acknowledged as a methodological solution by presenting an opportunity to explore and address uncertainties around conducting a future main trial. Little research, however, has explored the optimal design, conduct and necessity of PFS for surgical trials. This work sought to provide a detailed understanding of the potential for PFS to optimise future surgical trials and provide clear recommendations for surgeons and study teams.

The thesis comprises three phases of work: a targeted review and systematic analysis of NIHR funded surgical PFS protocols; a qualitative interview study exploring the perceptions, views and experiences of key stakeholders involved in designing and conducting surgical PFS and; a synthesis and interpretation of findings from the first two phases, to produce recommendations for the optimisation of PFS in surgery.

The findings indicate a cyclical model of sub-optimisation, in which misunderstanding leads to sub-optimal design, conduct, reporting and consequent devaluation of PFS amongst the surgical research community. Confounding factors further preventing optimisation include cultural challenges historically inherent to the surgical community, and a lack of targeted and accessible guidance.

A set of recommendations have been produced for all key stakeholders for optimising the design and conduct of PFS to inform main trials in surgery. To operationalise and apply these recommendations, a brief set of practical top tips have also been developed targeted specifically at surgeons involved in designing and conducting surgical PFS. A key recommendation includes a proposal to develop and seek endorsement for formal, consensus-based accessible guidance for surgeons on the design and conduct of surgical PFS. Additional recommendations focus on improving cross-disciplinary collaboration within the surgical trials community, the reporting and dissemination of PFS, and the research funding infrastructure. Future implementation of these recommendations will enhance the quality of definitive surgical trials, strengthen the surgical evidence base and, ultimately, improve surgical practice and patient outcomes.

# *Author's declaration*

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... .......... DATE: 03.10.19

The following paper was published from the research completed in this thesis.

**Fairhurst K, Blazeby JM, Potter S, Gamble C, Rowlands C, Avery KNL. Value of surgical pilot and feasibility study protocols. Br J Surg. 2019;106(8):968-78.**

A copy of this paper is included in Appendix VIII.

K Fairhurst is the first author of this paper and primarily performed the data collection, analysis and interpretation as well as drafting the finished article. The other authors contributed to this paper in terms of duplicate screening the databases (Rowlands), supervising the work (Avery/Blazeby/Potter/Gamble) and contributing to revisions of the paper (Potter/Avery/Blazeby). K Avery was the senior author of this paper.

SIGNED: .......... DATE: 03.10.19

SIGNED: .......... DATE: 03.10.19

# *Acknowledgments*

I would like to acknowledge the support and assistance of the following people who helped make this work possible.

**The Chief Investigators** of the RfPB studies who kindly supplied their study protocols in order that this work could be completed. Listed in alphabetical order:

Pradeep Bhandari; Jane Blazeby; Martin Birchall; John de Caestecker; Angharad Care; Malcolm Crundwell; Tim Davis; Anjan Dhar; Sam Eldabe; Kevin Franks; Anju Goyal; Damian Griffin; Emma Hall; Peter Hull; Khaled Ismail; Charles Knowles; Adrian Marchbank; Hugh Mostafid; Steven Pereira; Thomas Pinkney; Shelley Potter; Rajendra Prasad; Paul Roderick; Alastair Sutcliffe (NB protocol accessible online); Martin Tickle; Sumita Verma.

**Martin Ashton-Key and Amanda Blatch-Jones**, at the NIHR Evaluation Trials and Study Co-ordination Centre (NETSCC) for their assistance with the Microsoft Excel database used for screening for the HTA study cohort, and for commenting on the protocol for the systematic analysis of NIHR funded pilot and feasibility work.

**Ceri Rowlands,** fellow PhD student**,** for his help with screening the NIHR HTA and RfPB databases of funded studies.

**Jennie Hejdenberg and Ben Morgan,** at the NIHR Central Commissioning Facility (CCF), for their assistance in confirming the publication and outcomes of the RfPB study cohort, where the data was publicly available.

**The PEPSTAR participants,** though they remain anonymous, for without them this work would not have happened. Exceptionally busy people gave up their valuable time to talk to me, and I am extremely grateful for their generosity and engagement with this work.

**Claudette Blake & Steve Beech**, Research Secretaries, University of Bristol, for their expert and rapid transcription of the PEPSTAR interviews.

**Chris Metcalfe**, Professor of Medical Statistics and **Leila Rooshenas**, Lecturer in Qualitative Health Science, University of Bristol, for their time and expertise in performing my annual reviews, and their guidance on the structure of this work and the thesis chapters.

**My supervisors,** for their collective expertise, patience and support throughout my PhD. Specifically:

**Carrol Gamble,** Professor of Medical Statistics, University of Liverpool for her oversight of the project and contribution to supervision meetings.

**Jane Blazeby**, Professor of Surgery, University of Bristol, provided supervisory and oversight support for the project, particularly in directing the structure of the first phases and facilitating recruitment of the PEPSTAR participants. Thank you for the opportunity to work within the Centre of Surgical Research at the University of Bristol and explore a potential academic career.

**Shelley Potter,** NIHR Clinician Scientist & Consultant Senior Lecturer in Oncoplastic Breast Surgery, University of Bristol, for her enthusiasm, guidance, and resolute belief in my potential. Thank you for the no-nonsense approach and inspiring me to persevere.

**Kerry Avery,** Senior Lecturer in Health Services Research, University of Bristol for her attention to detail, guidance and open-door policy as my primary supervisor. Kerry has been unwavering in her support, availability and encouragement throughout my PhD journey. Thank you for weathering the highs and lows, teaching and believing in me.

**My family,** who make it all worthwhile.

# *Table of Contents*

# *List of Figures*

# *List of Tables*

# *List of Appendices*

# CHAPTER 1

## *Why it is necessary to optimise surgical trials*

## 1.1 Introduction

This chapter will describe the historical development of surgical trials and why they need to be optimised. The multiple challenges specifically facing surgical trials will be discussed and explained including those in relation to recruitment, surgical interventions, outcomes and surgical culture. Some identified solutions to these challenges will also be discussed, and pilot and feasibility studies presented as one such potentially important solution for optimising future surgical trials.

## 1.2 The importance of evidence to inform health policy and change clinical practice

Clinical decision-making, health policy and best clinical practice guidelines are informed by evidence from research. Evidence based medicine was first introduced as a concept in the early 1990's when clinicians began to evaluate the evidence from clinical research to inform the practice and treatment of patients, rather than relying on experience and theory taught from basic science principles alone [1]. To assist clinicians with this new way of practising, hierarchical pyramids of evidence were described based on study design (Figure 1a). These range from expert opinion and case series as the weakest forms of evidence at the bottom of the pyramid, to randomised controlled trials (RCTs), systematic reviews and meta-analyses, as the strongest forms of evidence at the top. Whilst such systems are helpful, they can also be misleading.

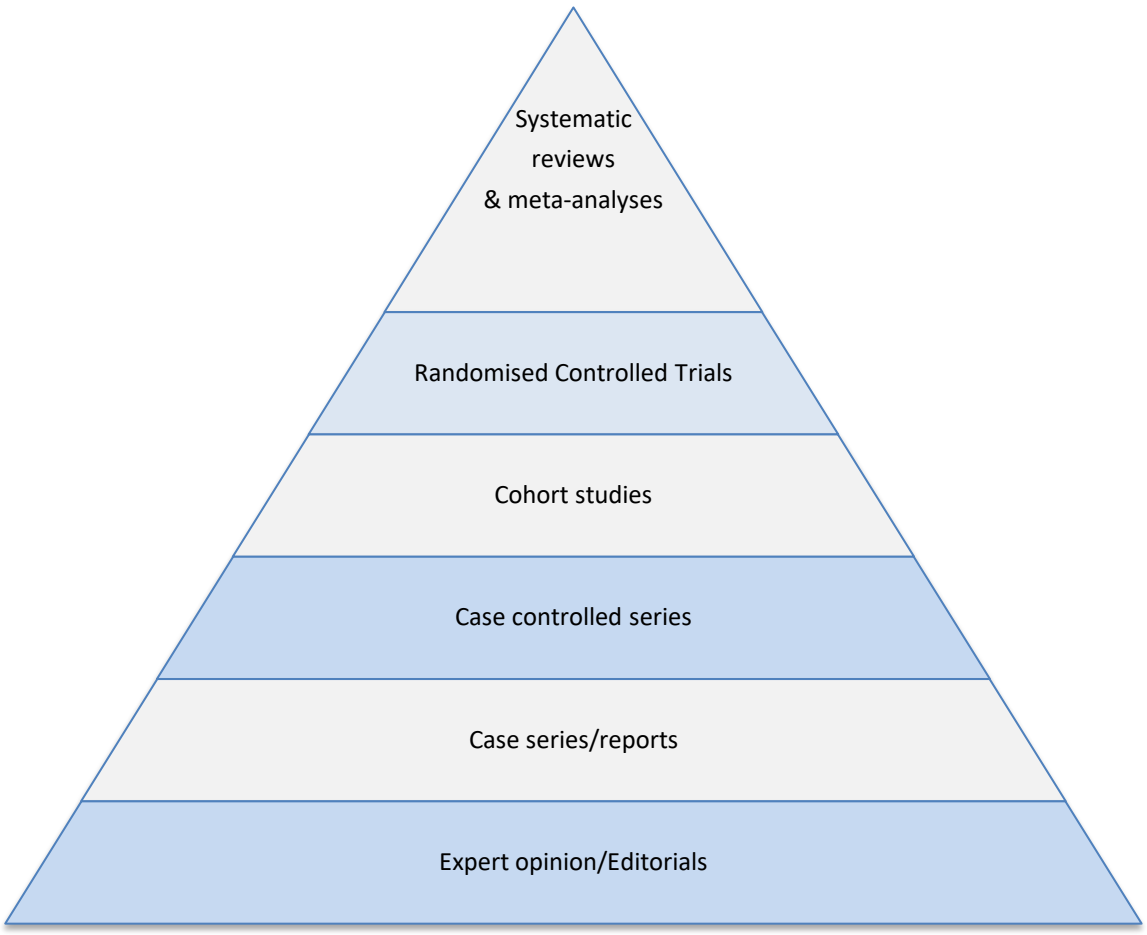



***Figure 1a*** *The evidence pyramid.*

Systematic reviews and meta-analyses are only as good as the studies they include: if you put poor quality evidence in, you get poor quality evidence out. Therefore, the evidence from a large pragmatic RCT may be less biased (internally valid), and more generalisable (externally valid), than a meta-analysis of multiple small, underpowered poorly designed RCTs. This has led others to challenge the traditional pyramidal hierarchy of evidence, most recently by suggesting that systematic reviews and meta analyses are a lens through which other forms of evidence should be viewed [2]. Others have suggested that evidence based medicine is 'in crisis' [3], and has perhaps been misinterpreted to result in a plethora of guidelines and protocols which are too restrictive on real world practice. Consequently, it is advocated that a return to patient centred shared decision making is needed. To acquire more valuable and relevant evidence to inform this decision-making process, improved RCTs asking relevant questions

that are important to patients and clinicians are vital. However, RCTs can be difficult, time consuming and expensive to perform.

Improved trial design, conduct and evaluation is sought to enhance the likelihood of main trial success, therefore avoiding trial failure and reducing research waste. When trials fail, it is a significant waste of time, effort and resources. Trial failure can also result in loss of morale and momentum within research teams and even disappointment from patients that their involvement has not been worthwhile. Improved trials will generate the best evidence to inform practice and improve outcomes for patients in the future. Trials in surgery, however, face unique challenges not encountered in trials in other clinical areas, and therefore need special consideration.

## 1.3 Challenges and some solutions in designing and conducting surgical trials

*"…I should like to shame [surgeons] out of the comic opera performances which they suppose are statistics of operations."*

*Major Greenwood, 1923* [4].

*"…moving from slapstick to symphonies, and it's symphonies that we are after today."*

*Jane Blazeby, 2016* [5].

It is more than 20 years since Richard Horton published his commentary in the Lancet on poor quality of surgical research in general, the lack of RCTs and a preponderance to case series as evidence for efficacy [4]. In this commentary, some challenges to conducting RCTs in surgery were identified. Those challenges identified include the potential difficulties of multicentre collaborative research, complexities in the standardisation of surgical techniques and the difficulties of selecting appropriate study outcomes and acceptable study designs. The surgical

community replied vociferously in their own defence [6], and in doing so, perhaps overlooked these valuable insights as to why surgical research was indeed challenging.

A decade on, a group of methodologists and surgeons met and formed the Balliol Collaboration, through taking part in three conferences concerning surgical innovation and evaluation at Balliol College, Oxford University. In 2009, the Balliol Collaboration published three seminal papers in The Lancet [7-9] describing the process of surgical innovation (termed the IDEAL paradigm or framework) [7], the challenges facing the evaluation of surgical innovation [8] and recommendations for the design and conduct of future surgical research [9]. The IDEAL recommendations propose different study designs and methods of reporting for each stage of surgical innovation. These stages span from the initial '**I**dea' (Stage 1), to '**D**evelopment' (Stage 2a), '**E**xploration' (Stage 2b), '**A**ssessment' (Stage 3), and '**L**ong-term monitoring' (Stage 4), with different study designs recommended at each stage. The purpose of setting out the IDEAL recommendations was to improve the quality of research in surgery and ensure that there was 'no surgical innovation without evaluation'[9].

The challenges of designing, conducting and evaluating surgical research, beyond those typical of any RCT, have been well documented in the literature by several authors [8, 10-12]. Surgical research is challenging fundamentally, because surgery is a complex intervention, meaning simply that it consists of multiple interacting components acting both independently and interdependently [13]. Complex interventions are widely used throughout the health service and include interventions from primary care and public health to surgery. What makes an intervention complex rather than simple is not well defined and it may, therefore, be more helpful to consider all interventions as having different levels of complexity [13]. The evaluation of complex interventions presents numerous methodological and practical challenges. Many of these are pertinent to the assessment of all complex, non-pharmacological interventions, but unique to surgery is the fact that multiple challenges often coincide [8].

A surgical intervention has been defined by others in the literature, but at the start of this work, there was no universally accepted definition of exactly what a surgical intervention constituted. Others had defined surgical interventions in the literature [12, 14, 15], but these definitions are limited with regard to the scope of modern surgical practice. For the purposes of this thesis it was therefore necessary to define a surgical intervention. A definition was created with more relevance to modern day surgical practice inclusive of a range of techniques, procedures and delivering personnel. Explanation of development detail and the parameters of this definition are discussed in Chapter three, section 3.3.2. In brief, a surgical intervention is defined for this thesis as: *A diagnostic, therapeutic or adjunctive invasive intervention performed by a trained clinician, using hands, instruments and/or devices.*

Trials of surgical interventions may, therefore, be considered as trials of complex interventions, but with distinct challenges. Broadly these challenges may be considered as being related to methodological and cultural issues, though these may overlap and interact. Methodological difficulties may include for example, those with recruitment, intervention stability and standardisation, and outcome selection and/or measurement. The cultural challenges of surgical research and practice may be separate or intertwined with these methodological challenges, for example a lack of equipoise amongst surgeons, or inexperience of recruiting to RCTs and working collaboratively. These challenges are discussed in detail below.

### 1.3.1 Challenges relating to recruitment

Clinicians must be able to portray equipoise (even if they are not in equipoise) in order to explain to patients that there will be no advantage or disadvantage to a patient taking part in a trial and that the lack of evidence means they are unable to make a decision about which treatment to offer [16]. Whether surgeons are in equipoise will depend on what the intervention is and how different this is to the comparators. Surgeons have traditionally not worked in a highly evidence-based field and are more used to making treatment decisions based mostly on their

own preferences. They may not even be aware of their own equipoise, or lack of, and how this relates to community equipoise, which may contribute to difficulties recruiting and randomising patients in the context of a trial [17-19]. Surgeons may also have limited appreciation of the methodological flaws of data derived from non-randomised studies and the possibilities of pragmatic trial designs [20], which will also limit their appreciation of the role of RCTs in surgery. Surgical procedures are very emotive for patients and patients too may exercise preference for a certain surgical treatment, which can make randomisation even more challenging [21].

For example, the QUEST (quality of life after mastectomy and breast reconstruction) feasibility trials considered the impact of the type and timing of latissimus dorsi breast reconstruction on health-related quality of life when post mastectomy radiotherapy was unlikely (QUEST A) or highly probable (QUEST B). The primary outcome was to demonstrate acceptable randomisation rates of at least 25% of patients screened. Both feasibility trials closed early because of poor recruitment with the authors citing patient preferences and non-acceptance of randomisation as the major reasons for failure. Issues with clinical equipoise of the recruiting surgeons were also reported as well as clinician biases with the concept of randomisation itself [22]. The ESTEeM Trial (Endocrine +/- Surgical Therapy for Elderly women with Mammary cancer) was also stopped as it failed to recruit. Again, the authors cited patient preferences and non-acceptance of randomisation as the major reason for poor recruitment, but also stated that clinician preference and the influence of family and friends may have played a part [23]. Both these trials demonstrate that patient and surgeon preferences may be co-influential as a challenge to trial recruitment.

Many of the issues described above around recruitment are specific to RCTs. Work over the last two decades has led to the development of qualitative methods to improve recruitment to RCTs [24, 25]. These methods have been highly effective and successful in improving patient recruitment to surgical trials by surgeons, through identifying opportunities for training and improving

recruiting practices [26]. It is advocated that such methodology is integrated in the pilot/feasibility stage of the trial when work can be done to consider both patient and recruiter views and/or beliefs about the trial. In addition, because RCTs are not always required, possible or practical in surgery, there is a growing body of work and evidence regarding incorporating a wider range of study types for the evaluation of surgery [7-9]. These include, for example, using surgical registries for real world evaluation of surgical devices and implants [27, 28]

### 1.3.2 Challenges relating to interventions

Buxton's law stated in 1987 that: 'It's always too early (for rigorous evaluation), until unfortunately it's suddenly too late' [29]. The phenomenon explained by this law is a process where evaluating interventions whilst they are developing and changing is difficult, meaning interventions progress seamlessly to becoming stable. At this point, the delivering clinicians become so convinced of the benefit of the intervention, that they lack equipoise and would not accept randomised evaluation in the context of an RCT. The IDEAL recommendations describe the study designs which may be appropriate at different stages in the development of an intervention so that there is 'no innovation without evaluation' [9]. The appropriate evaluation of surgical devices and techniques before widespread adoption has become even more topical following evidence of emerging harm with some surgical implants. These implants include Poly Implant Prothesis (PIP) breast implants [30], metal on metal hip replacements [31], and transvaginal mesh [32]. Knowing what study design is appropriate and possible at different stages in the development of surgical interventions is key, so that the chance for randomised evaluation is not missed.

Because surgical interventions are invasive, they also present unique questions on ethical issues such as when to introduce a new intervention to clinical practice and how to obtain informed consent from patients who are undergoing novel surgical procedures [33]. These issues in turn may present distinct difficulties in terms of gaining ethical approval for proposed research on surgical interventions.

Trials in surgery may be hindered by a lack of standardisation of the surgical procedure and difficulties in recognising when a surgeon has significantly deviated from the intended procedure. Surgeons often adopt procedural variations to best suit their personal skills, the skills and experience of the team, the facilities of the local clinical environment and the needs of individual patients. Surgical interventions are therefore dependent on the expertise and actions of many operators, including the surgeon, anaesthetist, operating department practitioners and theatre nurses. Whilst all health care professionals are increasingly aware of the need for quality assurance in the NHS in general [34], there are many challenges as to how this should be achieved. In surgery, there is also recognition that quality assurance should not be interpreted as simply ranking the 'best' surgeons [35]. Instead, quality assurance should constitute clear protocols for surgical interventions so that outcomes can be measured, compared and monitored. The complexity in achieving this lies in deciding which elements of surgical procedures should be standardised, and protocolised in the context of trials, and which outcomes are key in the assessment of benefit. Extensive methodological work has been done to develop a framework to standardise surgical interventions in the context of RCTs [36]. Work is ongoing to extend this framework to include co-interventions, and the identification of factors which may influence the extent of standardisation required.

The intervention may be altered by a surgeon's capability of learning new procedures, the so-called 'surgical learning curve', and by individual surgical aptitude [12, 37]. A learning curve can simply be defined as improvement of a surgeon's performance over time [37]. In the conduct of trials of novel interventions, a learning curve can impact significantly on the results, as the intervention may be prone to more complications whilst evolving [38, 39].

Two approaches to address the issue of the surgical learning curve in the context of RCTs have been described [12, 37]. The design approach involves stipulating entry level criteria for operators in terms of the number of procedures performed and a minimum professional level. This 'expertise' approach may mean that the

results of the RCT are less externally valid. In contrast, the analysis approach uses statistical methods to quantify the impact of a learning curve on outcomes. The latter approach has a high data requirement and may therefore only be practical in very large trials [12]. The use of a questionnaire to elicit information from surgeons participating in a trial about their current practice, experience and beliefs about learning curves, has been advocated to aid the interpretation and generalisability of trial results [40]. Limitations to this approach include that personal beliefs and attitudes will influence the responses to the questionnaire.

One way of dealing with the challenges and complexities of standardising surgical practice, is to perform large pragmatic surgical trials. Unlike pharmaceutical trials where patients can be guaranteed to receive the same drug, no two surgical treatments are the same leading to complexity with intervention implementation, standardisation and fidelity [8, 14]. Surgeons will never all practice in exactly the same way and care surrounding the surgical procedure will not necessarily be provided in the same way. The methodological arrangement of pragmatic trials allows for this, meaning the results may be more generalisable to practice in the 'real world'. Challenges with this approach, include the extent to which a trial is pragmatic; for example, how rigid should the trial protocol be, which elements should be standardised and how should co- comitant interventions be considered.

Consideration of co-interventions are therefore of particular relevance to the design of trials of surgical interventions. There are many which have the potential to impact on outcomes such as pre-optimisation programmes, the type and duration of anaesthetic given, what perioperative care is received (for example, high dependency, intensive care or ward care), and enhanced recovery programmes. Whilst the potential impact of such co-interventions is recognised by those with expertise in surgical trials methodology [14], consideration has not yet appeared in guidelines seeking to standardise the reporting of interventions such as the TIDieR (Template for Intervention Description and Replication) checklist [41]. The purpose of the TIDieR checklist is to encourage authors to

describe interventions in sufficient detail to allow for their replication. The checklist is an extension of both the CONSORT (CONsolidated Standards Of Reporting Trials) 2010 statement [42] and the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 statement [43].

Randomised controlled trials are designed specifically to avoid biased results. Despite randomisation, different forms of bias can still occur, such as through missing data, an imbalance in baseline prognostic variables, or in not performing an intention to treat analysis [44]. Blinding of both providers and recipients of healthcare interventions and those assessing the primary outcomes is desirable to avoid biasing the results. However, it is often difficult to blind any members of a surgical team to the intervention received by a patient as they are typically all intricately involved in providing that intervention [45].

The different methods of blinding available for non-pharmacological interventions have been reviewed and classified [46], and include the use of sham and/or placebo procedures and blinding participants or assessors to the study hypothesis. It is unknown whether complex blinding procedures translate into improvements in the quality of evidence, but work is underway to consider this in general and abdominal surgery [47]. It is important to consider the effects of blinding or not blinding members of the trial team, and how challenges may be overcome, and effects modified. Pilot and feasibility work may be well placed to trial novel or complex blinding procedures [48].

### 1.3.3 Challenges relating to outcomes

The selection and measurement of outcomes is less clear in surgical trials, and if not well thought out can lead to less meaningful results. Traditionally outcomes from surgery have focussed on short term clinical measures, for example mortality and morbidity. The Dutch TIME trial (traditional invasive versus minimally invasive esophagectomy), for example, compared two or three phase oesophagectomy with totally minimally invasive oesophagectomy. Short term outcomes were reported of a lower incidence of in-hospital pulmonary infection with minimally invasive oesophagectomy (12% versus 34%) and a shorter

hospital stay (median 11 days versus 14 days) when compared to open surgery [49]. The publication of this paper concluded that minimally invasive surgery was advantageous, but was met with criticism from the surgical community who argued that had various components of the surgery been different, the results would have been altered [50-53].

An additional challenge is that the definitions of clinical outcomes are both numerous and not standardised. The sheer volume of outcome definitions and their heterogeneity, therefore, makes comparison between studies and institutions difficult. For example, in 2001, a National Institute for Health Research (NIHR) funded review was published on the measurement and monitoring of surgical adverse events [54]. A total of 41 different definitions and 13 grading scales of surgical wound infection were identified from 82 studies. Similarly, 56 definitions of anastomotic leak were found from 107 studies of upper gastrointestinal, hepatopancreaticobiliary and lower gastrointestinal surgery. In terms of mortality, although the review found the definition to be relatively consistent, most systems reported in-hospital mortality only and few had the potential to link deaths to national registers [54]. The recording of outcomes in national registries has been shown to be important for transparent public reporting, research, and improving the quality of care for patients [31, 55].

The absence of standardised definitions for surgical outcomes has led to the development of scores such as the Clavien-Dindo classification of surgical complications, which grades adverse events by clinical severity and the treatment required [56]. Such scores are widely used though there is no universal consensus, and outcomes continue to be reported with huge variation.

In addition, whilst the clinical measures of morbidity and mortality are undoubtedly important, there is growing recognition of the need to also consider the perception of patients when choosing trial outcomes. The NIHR promotes and requires Patient and Public Involvement (PPI) in all stages of trials research in the UK, and funds the INVOLVE programme to actively support public involvement in NHS, public health and social care research [57] . Specific to

selecting research outcomes, PPI involvement in the early stages of study design is vital to ensure that important clinical questions are answered and that outcomes are patient-focussed. With relevance to surgery, what patients consider to be important in terms of outcome, might be different to outcomes of interest to surgeons. For example, long-term function following hand surgery may be of more interest to patients, than whether or not they develop a short-term wound infection or how straight their fingers are.

In terms of trial outcomes, the real issue of importance is to assess the longer-term clinical effectiveness of surgery, and thus consider the true value of these procedures. Clinical, patient reported and economic outcomes that are relevant to patients, health care professionals and other key stakeholders such as funders and health policy makers all need to be reported to achieve this. The selection of appropriate outcomes is therefore pivotal to the process of improving the relevance of surgical trials so that they can influence health care decision making.

The COMET (Core Outcome Measures in Effectiveness Trials) Initiative aims to develop agreed, standardised sets of outcomes, known as core outcome sets (COS), which represent the minimum number of outcomes that should be measured and reported in any trial of a specific condition [58]. These sets are not restrictive but allow the results of trials and studies to be compared and combined in terms of the core set of outcomes relevant to the clinical area.

There are now 52 (of a total of 1249) references to planned, ongoing and completed work to develop COS with surgery in title within the COMET database [58]. Whilst COS standardise outcome reporting, thereby improving the ability for comparison between studies and institutions, they are often specific to a disease or condition. This means that strategies must be employed to ensure outcome selection for each individual study is appropriate. An example of such a strategy, is the development of a new standardised measure to evaluate surgical site infection, which has built on the COS methodology and been completed within the context of pilot work [59].

### 1.3.4 Challenges & solutions relating to surgical culture

Historically, there have also been cultural challenges surrounding surgical research, such as a lack of collaborative working and a lack of equipoise amongst surgeons [19]. Similar to the examples of methodological work to improve surgical trials highlighted in the above sections, there has also been considerable movement to more broadly address these cultural issues and promote research around and participation in surgical trials.

One method initiated by the Medical Research Council (MRC), was the investment and development of a Network of Regional Hubs for Trials Methodology Research [60]. One of the five hubs, called the ConDuCT-II Hub [61] (Collaboration and innovation in Difficult and Complex randomised controlled Trials In Invasive procedures), was specifically tasked with researching the methodology of trials of complex interventions, and in particular, trials in surgery. This initiative brought together the expertise of surgeons and methodologists to work collaboratively on optimising trials in surgery, thereby improving both the methodology and the culture surrounding the design and conduct of surgical trials.

Such schemes are a stride forward, but perhaps remain the domain of the few rather than the many. A shift in understanding is also needed amongst the surgical community more generally, so that participation in trials becomes embedded within routine surgical practice. The Royal College of Surgeons of England (RCSEng) has lead a Surgical Trials Initiative [62] including the investment in Surgical Trials Centres throughout the UK and the appointment of Surgical Specialty Leads; individual surgeons to champion surgical research in their clinical specialty. The 2017-2018 Surgical Research Report published by RCSEng [63] lists the initiative's achievements as: 85 current or completed clinical trials (42 open and recruiting, 34 in follow-up and 9 completed) with a further 35 trials in the setting up phase, and over 25,000 patients recruited to this joint portfolio of trials in the ten different specialties.

This development has coincided with the now widespread surgical trainee collaboratives, which have an unprecedented record in recruiting patients to RCTs [64]. For example, the West Midlands Research Collaborative completed the ROSSINI trial, recruiting ahead of schedule, which showed there was no difference in rates of surgical site infection following laparotomy when using a wound edge protection device compared with standard care [65]. Challenges exist as to how the trainee collaborative model will develop in terms of trial delivery in the future, including the establishment of new collaboratives and the development of infrastructure such as access to online databases and statistical advice [64]. However, trainee collaboratives have and continue to assist with the changing culture surrounding surgical trials. Through successive generations of trainees with understanding, interest and experience in being part of surgical trials and working collaboratively, taking up consultant posts, the network of experienced surgeon trialists will continue to grow thus building greater capacity for collaborative surgical research.

Complementary to the growing collaborative research structure in the UK [66], is a demonstration of the importance of teamwork in the performance of RCTs. One study [67] showed that increased length of time working together as a research team combined with the experience of the team members, and a healthy competitive instinct towards recruitment, contributed to a successfully functioning trial research team.

A significant challenge facing surgical trials, is the difficulty of integrating trials and research into busy clinical practice. Many surgeons, as well as not having formalised academic sessions as part of their NHS contract, find it difficult to engage in the complexities and nuances of trial processes alongside their clinical work. Treating as many patients as efficiently and effectively as possible is the main pressure faced by surgeons, and time to engage with clinical trials, for example to recruit patients, is very limited [68]. More work is needed, to improve research infrastructure and supporting busy surgeons in engaging with clinical trials.

Whilst the culture around surgical trials is evolving and growing into a more established network of surgeon trialists, the challenge now is to build on momentum and established networks, so that participation in surgical research becomes routine for all surgical patients and clinicians.

## 1.4 Summary

The unique challenges of surgical trials have been identified and discussed in this chapter. Addressing the challenges described requires development of: (i). Methods to optimise recruitment; (ii). Methods to define and standardise interventions, monitor them, and select appropriate comparator interventions and; (iii). Methods to develop, select, measure and report relevant outcomes. Optimising surgical trials, therefore, requires both changes to methodological processes, and a shift in the culture surrounding the performance of surgical trials. Chapter two will review the role of pilot and feasibility work, as one solution to optimise trials more generally, and provide discussion of the importance of understanding why and how to optimally design and conduct pilot and feasibility work for surgical trials.

# CHAPTER 2

## *Review of the role of pilot and feasibility work in optimising trials*

## 2.1 Introduction

Chapter one described the many unique methodological and cultural challenges of surgical trials. This chapter will begin by discussing the applicability of pilot and feasibility work to optimising trials more generally, in the context of the methodological and theoretical literature surrounding pilot and feasibility work. In view of this knowledge, the rationale for understanding why and how pilot and feasibility work may be a potentially valuable method for improving surgical trial design and conduct will be presented at the end of this chapter, followed by a list of aims and objectives for this work.

But what is pilot work? The Oxford English Dictionary defines the word 'pilot' as "*…an experimental undertaking designed to assess the viability of a full-scale project or activity…That serves as a prototype or trial prior to a full-scale operation or activity*" [69]. Pilot work could, therefore, be simply described as work undertaken *a priori*, to inform future work.

However, defining what pilot and feasibility work is in the context of trials generally, has been inconsistent and fraught with resulting difficulties. This chapter will explore the published literature including the development in understanding of what pilot and feasibility work is; how it might be used and optimised to improve trial design and conduct and; why it might have been undervalued in the past. Throughout, the relevance of these questions to a surgical trials context will be considered.

This is a narrative literature review, exploring the most important aspects of current knowledge and research practice around pilot and feasibility work. Theoretical and methodological work is examined, discussed and considered in terms of surgery, as well as looking at the historical development of research practice around designing and conducting pilot and feasibility work. The aim was not to answer a clearly formulated question, by systematically examining all of the published literature, but to provide context for this thesis and consider inconsistencies or gaps in the body of current knowledge. This process allowed a rationale for this thesis to be described in terms of how exploring the issues surrounding pilot and feasibility work, may be integral to improving the design and conduct of surgical trials specifically. The chapter will conclude with the overall aims and objectives of this thesis.

## 2.2 Definitions of pilot & feasibility work

There are several definitions of pilot work in the literature and, currently, terms are often used interchangeably and without universally accepted definitions [70-72]. Pilot work could simply be defined as any preliminary work undertaken in preparation for conducting a main study [73-76]. The National Institute of Health Research (NIHR) defines pilot studies as "*…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, randomization, treatment, and follow-up assessments all run smoothly. It resembles the main study in many respects, including an assessment of the primary outcome*" [77]. In addition, the NIHR distinguishes feasibility studies from pilot studies, describing them as: "*…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*" [78].

In contrast, the Medical Research Council (MRC) guidance on developing and evaluating complex interventions [13] makes no distinction between pilot and feasibility studies, describing how "*the pilot and feasibility stage* (of evaluating a complex intervention) *includes testing procedures for their acceptability, estimating*

*the likely rates of recruitment and retention of subjects, and the calculation of appropriate sample sizes…A pilot study need not be a 'scale model' of the planned mainstage evaluation, but should address the main uncertainties that have been identified in the development work.*" A key message in this MRC definition is that pilot and feasibility work should focus on identifying and addressing areas of uncertainty or ambiguity in the study protocol, to resolve any issues before the main trial begins.

More recently, Eldridge *et al*. have developed a conceptual framework to define pilot and feasibility studies in preparation for RCTs [71] (see Figure 2a). This work was integral to the primary research project of developing guidelines for reporting pilot and feasibility studies [79]. The core components of the primary research were a large Delphi study and an international expert consensus meeting [80]. During this research, the authors felt the lack of clarity, differences of opinion and interest in defining pilot and feasibility studies, warranted re-evaluation and prompted the development of the definitions framework.

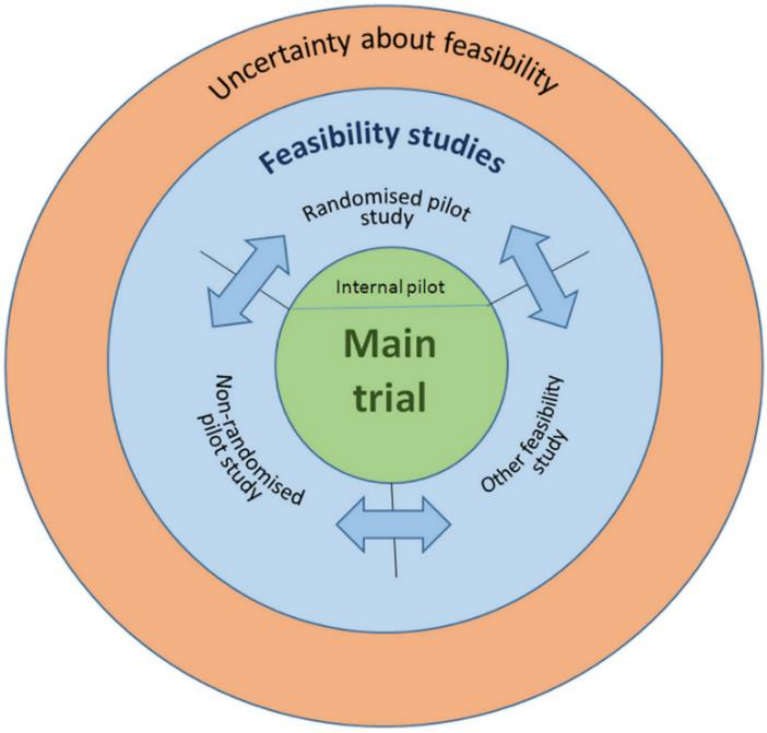



***Figure 2a*** *Conceptual framework to describe the classification of pilot and feasibility studies (Figure reproduced from [71] – see acknowledgements)*

Within the framework, feasibility studies are defined as an umbrella term to

include all types of studies that ask: can this trial be done, should it be done and how? [71, 81]. Feasibility studies are then further sub classified into three groups (see Table 2.1). A pilot study, whether randomised or non-randomised, may therefore be categorised as a subset of feasibility work that is intended to ask such questions about the viability and conduct of a proposed main trial but that has specific design features [81].

***Table 2.1*** *Sub classification of feasibility studies by Eldridge et al. [71]*

| Study type | Definition |
|---|---|
| Randomised pilot studies | *"… studies in which the future RCT, or parts of it, including the randomization of participants, is conducted on a smaller scale (piloted) to see if it can be done. Thus randomized pilot studies can include studies that for the most part reflect the design of a future definitive trial but, if necessary due to remaining uncertainty, may involve trying out alternative strategies, for example, collecting an outcome variable via telephone for some participants and online for others. Within the framework randomized pilot studies could also legitimately be called randomized feasibility studies."* |
| Non-randomised pilot studies | *"… studies in which all or part of the intervention to be evaluated and other processes to be undertaken in a future trial is/are carried out (piloted) but without randomization of participants. These could also legitimately be called by the umbrella term, feasibility study. These studies cover a wide range from those that are very similar to randomized pilot studies except that the intervention and control groups have not been randomized, to those in which only the intervention, and no other trial processes, are piloted."* |
| Feasibility studies that are not pilot studies | "…(studies) *in which investigators attempt to answer a question about whether some element of the future trial can be done but do not implement the intervention to be evaluated or other processes to be undertaken in a future trial, though they may be addressing intervention development in some way."* |

These design features include determining the sample size of the main trial; addressing recruitment and randomisation, testing the implementation of interventions; assessing data collection and outcomes and considering harms, benefits and potential effectiveness of the intervention [82]. These definitions by Eldridge *et al*, illustrate that pilot and feasibility work may not be seen as mutually exclusive but rather as part of the same paradigm that includes any work done in preparation for a main trial.

Importantly, the definitions defined by Eldridge *et al*, refer only to so-called external pilot studies; standalone studies that are completed before the start of an RCT and in which, according to the NIHR definition '*….the data may be analysed and set aside'*[77]. The definitions do not address so called internal pilot studies, which are described by the NIHR as '*…the first phase of the substantive study and data from the pilot phase may contribute to the final analysis'* [77]. The design of and data from internal pilot studies are, therefore, integral to the main study. Internal pilot studies may consequently be considered more appropriate when there are fewer uncertainties about the trial and study protocol. Internal pilot studies provide an opportunity for improved efficiency in terms of time, participants and funding, as the study may continue to the main phase if it meets the pre-defined progression criteria [83, 84].

These distinctions are important as, whilst different types of pilot work do exist and examples of such studies are diverse, there is no established guidance to aid selection of the most appropriate type of pilot or feasibility study to perform prior to undertaking a definitive main trial.

Most recently, Cook *et al.* argue that the terms internal and external pilot are unhelpful and outdated [85]. The authors emphasise the importance of researchers reporting and publishing in detail, the refinement of trial processes so that others can learn from the findings. An alternative taxonomy is proposed by Cook *et al.* to illustrate when each study type should be used (see Table 2.2). Here, all pilot and feasibility studies are described as distinct studies in their own right. The authors suggest that as all trials have some form of internal assessment process allowing refinement to the study design, it is particularly important to improve clarity and transparency around what is actually done, including formal reporting. The overlapping terms of 'stop go assessment' and 'routine monitoring' are defined as terms to illustrate these internal processes, rather than the term 'internal pilot'.

***Table 2.2*** *Approaches to developing and modifying the trial design with particular reference to the recruitment strategy (Table reproduced from [85] – see acknowledgements)*

| Approach | Conception | When to use | Key purpose | Main output |
|---|---|---|---|---|
| Feasibility study | Separate study | Substantial uncertainty about whether it is possible or how to implement aspects of the trial design and conduct (e.g. delivery of the interventions and recruitment process) | Exploration of the feasibility of a trial | Assessment of the feasibility of a trial |
| Pilot trial | Separate study | Limited refinement of the trial design, processes and recruitment strategy anticipated | Refinement of trial design and processes prior to commencement of the main trial | Minor modifications to trial design and confirmation of practicality |
| Stop-go assessment | Early in the recruitment period of the main trial | Substantial adjustment of recruitment strategy and study processes is considered likely | One-off within study assessment of recruitment and study processes to allow adjustment of the strategy (including the possibility of early cessation of recruitment) | Single reassessment of the recruitment and data collection strategy |
| Routine monitoring | Throughout the recruitment period of the main trial | Minor modifications to study processes and recruitment strategy anticipated | Regular monitoring to allow periodic within study modification | Periodic updating of the recruitment strategy and study processes |

The examples provided above represent some of the more widely cited definitions of pilot and feasibility work in the context of healthcare research. A shared characteristic of all these published definitions is that they describe pilot and feasibility studies as preliminary work, performed in preparation for a main trial. The debate over what pilot and feasibility studies are and how they should be defined continues. Defining the different types of pilot and feasibility studies is relevant when designing trials, but also when creating guidance for the conduct, reporting, evaluation and publication of research. This is because to avoid further confusion in the literature, which provides the resource for further work, it is important to be clear about to which type of study the guidance refers. In the absence of a consensus on terminology currently, it is perhaps most

important that researchers fully describe the study methodology used when describing the design and/or findings of pilot and feasibility work, so that it can be transparently understood, replicated or refined.

It is noted, however, that none of these definitions are defined within the context of surgical trials specifically. Some surgical examples are used in the conceptual framework describing the different types of pilot and feasibility studies [71]. Whilst the list of participants for the Delphi survey did include 'clinicians', it is unclear if surgeons were amongst those invited to take part. The later parts of the consensus process occurred at methodological conferences, so significant surgeon involvement is less likely.

In the absence of a universally accepted definition of pilot and feasibility work for clinical trials, for the purposes of this thesis, if was necessary to define a clear and inclusive working definition. This definition is explained in more detail in Chapter 3 (Methods, Section 3.3.3) but for clarity whilst directly discussing definitions here, the term pilot/feasibility study (PFS) was coined for the purposes of this thesis and defined as:

**Any research that is undertaken before a main study and is explicitly intended to inform the design and/or conduct of a future main study** where:

**Main study** is defined as a definitive study (e.g. RCT) of an intervention(s).

## 2.3 Broadening the potential scope of pilot and feasibility work

Whilst the definitions of pilot and feasibility work have emphasised the intention to address areas of uncertainty or ambiguity, historically the scope of pilot work has often been narrow, focusing typically on issues related to safety and efficacy [86, 87] or recruitment [88]. It has been acknowledged, for example, that pilot work has not comprehensively evaluated methodological issues and has rarely detailed the intention or process of proceeding to a main trial [74, 89]. For example, in 2011 Shanyinde *et al*. performed a literature review of 50 pilot RCTs and illustrated that only 56% of the studies reviewed addressed methodological issues in any

depth [89]. There is, however, increasing awareness of the greater potential for pilot work to inform the design and conduct of main trials more generally. In particular, there is awareness of the potential for pilot and feasibility studies to address the many methodological issues, uncertainties and ambiguities in the study process which have not been formally addressed previously [89]. The investigation of these uncertainties may, therefore, further enhance the likelihood of success of the main trial [73, 75, 90]. As such, perhaps the full potential of pilot work has arguably not yet been realised.

The knowledge that there are many other areas that could be addressed to increase the benefits of pilot work has lead several authors to identify and categorise specific reasons for undertaking pilot work, in an effort to guide researchers into systematically considering these areas for their own studies (see Table 2.3).

In 2001, van Teijlingen *et al.* [91] created a long list of 19 reasons for conducting pilot and feasibility work compiled from several sources (see Table 2.3). The main purpose of this article was to report learning from the conduct of a series of small pilot studies performed to determine the most effective way of surveying maternity care in Scotland and, therefore, did not consider surgical studies specifically. The need to assess the logistics of the study protocol is clear from this list of objectives. In particular, the authors emphasised at this early stage, the importance and value of pilot work both to the study team and to a much wider audience, therefore advocating mandatory reporting and publication of pilot studies.

In 2004, Lancaster *et al.* [72] performed a literature review of six pilot and feasibility studies published in 1998 to 2002, from which they described seven possible key objectives for performing external pilot studies (see Table 2.3). Of these six studies, none considered surgical interventions specifically, with four considering pharmacological interventions, and the other two considering blood pressure monitoring and skin-to-skin contact postpartum as the intervention. This list was extended to add a further three key objectives to the list in 2010, in a

review article published by Lancaster *et al.* regarding statistical issues in the design and evaluation of complex interventions in primary care [92]. These ten key objectives are perhaps more specific than those published by van Teijlingen [91], explicitly mentioning randomisation, recruitment, consent and the selection of a primary outcome measure.

Also in 2004, Thabane *et al.* [75] condensed the list produced by van Teijlingen *et al.* [91] and categorised the focus of pilot work as concentrating on the four main areas of processes, resources, management and scientific issues (see Table 2.3). The fourth of these categories, 'scientific', is described as an assessment of treatment safety, dose and effect. However, the authors themselves conclude in the discussion that pilot studies should be used with caution for estimating treatment effects. They make recommendations for reporting the results of pilot studies, by creating a checklist adapted from the CONSORT (2001) statement [42], which emphasises appropriate labelling of the study as 'pilot' and defining and reporting feasibility objectives and outcomes. Whilst the focus of this work is again complex interventions, surgical studies are not specifically considered or mentioned.

Shanyinde *et al.* [89] in 2011 expanded on this work by reviewing a random sample of 50 studies, which include 28 'drug' and 22 'non drug' pilot or feasibility studies, of which just two studies were of surgical interventions. From this, the authors produced a checklist of seven methodological issues that may/should be addressed in randomised pilot work specifically, prior to conducting an RCT. The authors contrasted this list with seven further issues that could be evaluated in other types of feasibility study (see Table 2.3). Specific additions in this list include an assessment of blinding procedures and considering the logistics of a multicentre trial. This is important as the work was published ten years after that of van Teijlingen *et al.* [91], and gives insight into the developing understanding of trials methodology.

In 2011 also, Leon *et al.* [93] discussed the different aspects of feasibility that can be examined in pilot work which are similar to those already listed (see Table 2.3).

The paper emphasises again that pilot studies are not powered to undertake hypothesis testing, and should not be used to evaluate, safety, efficacy, or effectiveness.

Bugge *et al*. in 2013 [94] used the list of 14 methodological issues published by Shanyinde *et al.* [89] (see Table 2.3) to develop the ADePT tool (A process for Decision-making after Pilot and feasibility Trials). The authors justify the list published by Shanyinde *et al.* as 'the best available', though acknowledge it is not clear exactly how it was developed. The work is centred on the conduct of a pilot study for an RCT of pelvic floor muscle training for pelvic organ prolapse. The 14 methodological issues were used as a checklist to assess how their own pilot study had addressed each area. Based on the areas of difficulty and uncertainty identified, they then developed the ADePT tool to consider: i) the type of problem (only for the trial, only for the real world, or both); ii) potential solutions to the problem; and iii) an assessment of which would be the best solution to the problem. The authors stated that they plan to evaluate the tool in terms of its usefulness for developing an intervention and designing a trial to test it. This work represents a growing understanding of how pilot work might be used to improve trial design, though not specifically considering the context of surgical trials.

Most recently, Kistin *et al*. [95] in 2015 discussed five main reasons for performing pilot work. The key addition here, is the third point of using pilot work to understand '*…barriers and facilitators to eventual dissemination and implementation of the results*', thus exemplifying the importance of trial results being implementable in clinical practice.

In summary, the literature has documented over the last 18 years, the multiple methodological issues that pilot and feasibility work may address prior to conducting a main trial. This work has been undertaken partly to address a growing understanding that pilot and feasibility studies are often narrow in scope. This work also reflects a gradual acknowledgement of the importance of having a clear rationale and objectives for pilot and feasibility studies to examine

specific issues prior to conducting a main trial. There is increasing understanding that because RCTs of complex interventions such as surgery face unique challenges (chapter one), pilot and feasibility studies may have greater benefit in such situations. However, whilst it is likely that many or all of the methodological issues described and discussed here are relevant to pilot and feasibility studies in surgery, none of the work reviews or explores surgery specifically. Next, the published guidance on optimal design and conduct of pilot and feasibility studies generally, with consideration of any guidance specific to surgery, will be discussed.

***Table 2.3*** *Reasons for conducting pilot work as reported in the literature.*

| Author / publication | Reasons for conducting pilot/feasibility work identified |
|---|---|
| Van Teijlingen *et al* 2001[91] | 1) Developing and testing adequacy of research instruments;<br>2) Assessing the feasibility of a (full-scale) study/survey;<br>3) Assessing people's willingness to participate/potential or likely response rates;<br>4) Designing a research protocol;<br>5) Assessing whether the research protocol is realistic and workable;<br>6) Assessing whether data collectors understand the protocol;<br>7) Establishing whether sampling frame and technique are adequate and effective;<br>8) Identifying logistical problems which might occur using proposed methods;<br>9) Estimating variability in outcomes to help determining sample size;<br>10) Collecting preliminary data;<br>11) Determining what resources (finance/staff) are needed for main study;<br>12) Assessing the proposed data analysis techniques to uncover potential problems;<br>13) Developing a research question and/or research plan;<br>14) Assessing whether or not each data collector obtains similarly valid information;<br>15) Training researchers in as many elements of the research process as possible;<br>16) Educating students about research methods and the research process;<br>17) Convincing funding bodies that research team is competent and knowledgeable;<br>18) Convincing funding bodies that the main study is feasible and worth funding;<br>19) Convincing other stakeholders (service providers, ethics committees, managers, politicians, etc.) that the proposed main study is worth supporting. |
| Lancaster *et al* 2004[72] | 1) To perform a sample size calculation;<br>2) To test the integrity of the study protocol;<br>3) To test data collection forms and/or questionnaires;<br>4) To test the randomisation procedure;<br>5) To assess rates of recruitment and consent;<br>6) To determine the acceptability of the intervention and;<br>7) To select the most appropriate primary outcome measure; |
| Lancaster *et al* 2010[92] | 8) To develop and test the implementation and delivery of the intervention;<br>9) To train staff in delivery and assessment procedures and;<br>10) To prepare and plan data collection and monitoring procedures. |

| | |
|---|---|
| Thabane *et al* 2010[75] | 1) Process: analysing the feasibility of different steps in the trial such as assessment of numbers of eligible patients and the acceptability of randomisation;<br>2) Resources: assessing the use of study resources such as, the usability of proposed software or the time it takes to fill in data forms;<br>3) Management: considering human and data management problems for example, data entry, missing values, study personnel difficulties and;<br>4) Scientific: an assessment of the treatment safety, dose and effect. |
| Shanyinde *et al* 2011[89] | 1) Recruitment;<br>2) Consent;<br>3) Randomization procedures;<br>4) Blinding procedures;<br>5) Retention;<br>6) Logistics of a multi-centre trial and;<br>7) Testing that all components of the protocol work together.<br>8) Sample size calculation;<br>9) Eligibility;<br>10) Compliance and/or adherence to the intervention;<br>11) Acceptability of the intervention;<br>12) Cost and duration of the intervention;<br>13) Outcome assessment and;<br>14) Selection of the most appropriate outcomes. |
| Leon *et al* 2011[93] | 1) Screening;<br>2) Recruitment;<br>3) Randomisation;<br>4) Retention;<br>5) Treatment adherence;<br>6) Treatment fidelity and;<br>7) The assessment process. |
| Kirsten *et al* 2015[95] | 1) To field test logistical aspects of the future study and to incorporate these aspects into the study design;<br>2) To optimize intervention delivery with specific attention to adherence and fidelity;<br>3) To increase understanding of the barriers and facilitators to eventual dissemination and implementation of the results;<br>4) To obtain empirical estimates of statistical parameters to inform power calculations and other design effects of a subsequent trial and;<br>5) To estimate other critical study parameters such as the proportion of study participants who experience a key outcome event spontaneously without intervention, or the level of correlation over time for a repeated outcome measure, or the extent of a clustering effect. |

## 2.4 Optimal design and conduct of pilot and feasibility work

There are some guidelines for the design of pilot work published in medical disciplines, including occupational therapy [96] and nursing [97]. The main characteristics of these guidelines are, however, based on the key works of others already discussed such as Thabane *et al.* [75] in 2010 and Arain *et al*. [74] in 2010, so are not consensus-based or surgery specific. Furthermore, these guidelines are not widely cited as useful for guiding the design of surgical pilot and feasibility studies.

Some published recommendations, however, have provided more practical guidance regarding the use of pilot and feasibility work in complex interventions that might have specific relevance to surgery. The Medical Research Council (MRC) guidance on developing and evaluating all complex interventions, offers a framework to guide researchers through the development-evaluation-implementation process [13]. Conducting pilot or feasibility work of complex interventions prior to full-scale evaluation in a definitive main trial is considered 'vital preparatory work' as part of this framework and is, therefore, recommended for the full evaluation of all complex interventions [13]. Current developments in this area include work also funded by the MRC, to develop guidance specifically for pilot and feasibility studies in complex public health interventions [98, 99].

Targeted specifically at surgery, the IDEAL framework [100] describes a pathway for new surgical interventions from first in man (Stage 1) to long-term study (Stage 4). Within this framework, pilot and feasibility studies are considered at Stage 2a (development) and 2b (exploration) focusing on addressing uncertainties prior to Stage 3 assessment in a definitive RCT. The initial publication however, was largely theoretical with little practical guidance as to how pilot and feasibility studies should be operationalised [101]**.** Recently published updated IDEAL recommendations [102] now provide some clarification regarding the role of pilot and feasibility studies in surgery, as a result of recognition that the original IDEAL guidance published in 2009 [101] had little impact on the design and conduct of surgical pilot and feasibility studies [103]. The authors of the updated IDEAL

recommendations [102] also state that the development of a reporting checklist for IDEAL studies is underway, to improve appropriate reporting of all studies using the IDEAL framework in the literature.

Whilst these strategic publications discuss pilot and feasibility work as part of a larger framework for developing, evaluating and implementing new complex interventions, such as surgery, it is still not clear exactly when and how pilot and feasibility work should be conducted. Specific to surgery, there is currently no formal guidance on how to design and conduct pilot or feasibility studies in surgery, in what situations this work should be done, and how or if pilot and feasibility work should be modified to differing circumstances, for example for different types of surgical intervention. It is identified, therefore, that further research is needed to understand the potential for pilot and feasibility studies to optimise future surgical trials.

## 2.5 Rationale for this thesis

Evidence of feasibility is increasingly required by funders, before agreement to fund a main trial is reached. This is based on the knowledge that poor research design, conduct and analysis are known to contribute to significant research waste [104-108]. This effect is further compounded by the poor reporting and dissemination of research findings [107], particularly those from clinical trials [109, 110]. Interest in the design and conduct of pilot and feasibility work is, therefore, intensifying with the increasing recognition that such preliminary work may contribute extensively to the success of subsequent definitive main trials. Through avoiding common problems such as the inability to recruit and a corresponding reduction in statistical power, or excessive attrition due to intolerable procedures, it has been suggested that pilot studies may also reduce the proportion of failed trials [93].

Practicality dictates that pilot studies should be a pre-requisite to trials of novel treatments or interventions or novel applications of treatments or interventions [93] or in areas where trials have been historically difficult to do, because of the inevitable uncertainty surrounding these types of trial. As described in Chapter 1, surgical trials face unique complexities, and often interacting uncertainties

surrounding the design, conduct and completion of trials, meaning there is an even greater need to consider if and how surgical trials can be improved.

Pilot and feasibility work may be well placed to explore the complex uncertainty surrounding surgical trials. Whilst there may be growing recognition that pilot and feasibility work is beneficial to main trial success and should precede all main trials of complex interventions, there has been little research to explore how to design and conduct pilot and feasibility studies in different situations.

Published guidance from the MRC and the IDEAL collaboration, describes the importance of pilot and feasibility work in overcoming the challenges of complex interventions, which includes surgical trials. However, the conduct, reporting and publication of pilot and feasibility work in surgery remains rare [111, 112]. Furthermore, guidance tailored specifically to optimally design and undertake pilot and feasibility studies for surgical trials is lacking. For example, there is no surgery-specific guidance endorsed by funding bodies or professional membership organisations such as the Royal College of Surgeons (RCS), aimed at surgeons participating in or contemplating designing pilot and feasibility studies. Research is needed, therefore, to understand the potential for pilot and feasibility studies to optimise main trials, and develop guidance for surgeons regarding best practice in this area.

## 2.6 Aim of thesis

The overall aim of this thesis is to develop a detailed understanding of the potential for pilot and feasibility work to optimise surgical trials and provide clear recommendations for surgeons and study teams for how to improve research practice.

## 2.7 Objectives of thesis

Specific objectives of this thesis are:

**Objective 1** To undertake a targeted review and systematic analysis of protocols of pilot and feasibility studies funded by the National Institute for

Health Research (NIHR) to understand current practice in using pilot and feasibility studies to inform main trials in surgery;

**Objective 2** To use in-depth qualitative interviews to explore the experiences, perceptions and views of key stakeholders towards the role, value and challenges of designing and performing pilot and feasibility studies in surgery;

**Objective 3** To synthesise and interpret the findings of the targeted review and systematic analysis of protocols of pilot and feasibility studies funded by the NIHR and the in-depth qualitative interview findings, in light of the methodological and theoretical literature around pilot and feasibility studies (as reviewed in chapter one and two) and;

**Objective 4** To develop recommendations from this work for all key stakeholders and, specifically, accessible and practical top tips for surgeons to optimise the design and conduct of pilot and feasibility studies to inform main trials in surgery.

Chapter three will describe in detail, the methods used for this work, and rationalise the choice of techniques in context of the breadth of methodological approaches available.

# CHAPTER 3

# *Methods*

## 3.1 Introduction

This chapter will describe the methods used in this thesis. The aims and objectives of this work are outlined on the previous page at the end of Chapter two. **Phase one** of this work was a targeted review and systematic analysis of NIHR funded pilot and feasibility studies to establish current practice in this area (objective 1). **Phase two** was a qualitative interview study with key stakeholders, to explore opinions, perceptions and views on the role, value and challenges of performing pilot and feasibility studies in surgery (objective 2). **Phase three** of this work synthesises and interprets the findings of the first two phases, in light of the evolving methodological and theoretical literature (objective 3), to produce recommendations for improving practice in this area (objective 4). This chapter will describe the methods and rationale for the use of these methods in each phase of the work.

# Phase I: A targeted review and systematic analysis of NIHR funded pilot and feasibility studies

## 3.2 Rationale for Phase I of this thesis

Chapter two described how major research funders are increasingly recognising the role of well-designed pilot and feasibility studies in informing the design and conduct of future surgical trials, but clear and accessible guidance on how and when to perform surgical pilot and feasibility studies is currently lacking. It was, therefore, hypothesised that reviewing and appraising completed examples of pilot and feasibility studies would be valuable to consider and evaluate exactly

what the issues regarding design and conduct are. It was also hypothesised that through defining and highlighting examples of well-designed pilot and feasibility studies, this would inform others and assist in designing future preparatory studies. It was considered, that examples of pilot and feasibility studies could have been searched for in the published literature. However, it is known that pilot and feasibility studies have been both inconsistently and under reported historically, meaning a review of the published literature would be unlikely to include examples of the full range of current research practice. The difficulties surrounding using the general literature as a resource to determine current practice in this area are explained below.

### 3.2.1 Issues with reporting of pilot and feasibility work

There are examples of good reporting in the scientific literature of how pilot work has explored specific uncertainties and methodological issues to usefully inform the viability and conduct of a main trial. The PROTECT thromboprophylaxis trial, for example, aimed to examine, the effect of low molecular weight heparin versus heparin on the primary outcome of proximal leg deep vein thrombosis (DVT) among medical and surgical critically ill patients. The published main trial protocol and analysis plan [113], clearly demonstrated how the main trial protocol had been improved as a result of the internal pilot study [114]. Analysis of data from the internal pilot, allowed, for example, the participant exclusion criteria to be broadened, making a main trial feasible and improving the ultimate applicability of the results.

In public health, the 'Walk to Work' RCT sought to examine the effectiveness and cost effectiveness of an employer-led scheme to increase walking during the commute [115], and was funded as a result of a successful feasibility study [116]. The first phase of the feasibility study was development of the intervention including a resource review and focus group work. The second phase was a pilot exploratory cluster RCT to test the intervention and evaluate the processes and costs involved. The results demonstrated the feasibility of both the proposed intervention and its evaluation and secured further NIHR funding for a full-scale cluster RCT.

Surgical examples also exist. The ROCSS external feasibility study randomising patients to standard vs biological mesh closure of stoma site reinforcement, aimed to evaluate recruitment, randomisation and deliverability of the new surgical technique [117]. This study successfully completed to inform a main trial funded by the NIHR [118].

There are, therefore, some detailed examples in the published literature demonstrating how pilot work has informed main trial design. More commonly, however, pilot and feasibility studies have been vastly both *inconsistently* and *under* reported in the scientific literature. The reasons for this appear to be multifactorial and are discussed below.

### *Inconsistent reporting of pilot and feasibility work*

It is possible that the lack of globally accepted definitions of pilot and feasibility studies (as discussed in chapter two) and the lack of understanding of their potential uses, may contribute to inconsistencies in the reporting literature. There are many examples of published surgical studies cited as 'pilot studies', that on closer inspection, are actually single centre, small RCTs [119-123] masquerading as pilot or feasibility studies. Interrogation of the work shows that the studies are not addressing areas of uncertainty preceding a main trial and are lacking investigation of possible uncertainties about the design or conduct of a main trial. Instead, the authors have undertaken formal hypothesis-testing and reported estimation of effect sizes [119-123]. Studies may interpret the data to support or refute a hypothesis, even though many are not powered to test a hypothesis or adequately refute one. These studies therefore do not report on feasibility issues in order to inform definitive trials, and are also insufficient to inform evidence. Such studies can therefore be misleading and prevent main trials being undertaking by reporting definitive results.

Certainly, in the past, the focus of much pilot work has been on hypothesis testing and estimating intervention effectiveness, but it is now generally accepted that pilot studies are not an appropriate setting in which to evaluate such issues [13, 72, 74, 75, 92, 93, 95, 96]. There has been acknowledgement amongst field experts that data on intervention safety or efficacy may be collected during a pilot or feasibility study

but should only be reported in descriptive terms. Such data should be clearly reported with the caveat of no firm conclusions being drawn because of a lack of statistical power [72, 75, 92, 124, 125], and changes in practice should therefore, not be based on the results of pilot work alone. Whilst misunderstanding may be the main reason why hypothesis testing has been performed within a pilot or feasibility study, another reason identified in the literature is judging such work may be more likely to be published, or more likely to be awarded funding or ethical approval if 'results' are given [74].

The terms pilot and feasibility have also been used very loosely in the literature to mean for example, a study conducted by a junior or student [75], or a study looking at the technical feasibility of performing a new surgical procedure [126, 127]. In addition, there is evidence in the literature of studies being redefined as pilot studies, *a posteriori* at the request of editorial boards. This practice has been undertaken seemingly to highlight the uncertainty of the results [128], which may include, for example, main trials that encountered difficulties with recruitment or intervention delivery. This is reiterated by Shanyinde *et al* [89] who suspected that many of the so called pilot RCTs in their review (n=50 of which 2 were surgical, also discussed in chapter 2), were named 'pilot' *a posteriori* when hypothesised effects were not seen or inadequate sample sizes were collated.

Other reasons for mislabelling studies include the perceived lack of resources to perform a large adequately-powered study and replicating other small published 'pilot' studies, thereby perpetuating the problems of misunderstanding and resource waste [75]. There has also been no requirement for formal registration of pilot and feasibility studies in the past, though improved regulation and guidance on research transparency and the requirement for formal registration of all clinical trials in order to gain ethical approval should improve this [129].

### *Under reporting of pilot and feasibility work*

Previously, pilot work may frequently have been published as only a fleeting mention in the main trial report. There is now growing agreement that pilot and feasibility work, if conducted well, should always be published regardless of whether the results are 'positive' or 'negative' for conducting a main trial [74], to

allow others to learn about the best research processes and avoid future repetitions of 'failed' methods. A pilot study that shows a trial is not possible, may be considered to have had a negative outcome, but should still be published if the methodology is thorough and insightful [74, 130]. Such studies may be important and useful to other researchers by, for example, avoiding repetition of the same mistakes and allowing development of methodological processes for future pilot and feasibility studies. In addition, some have suggested that authors have an ethical obligation to report all issues arising from a study, including those from the pilot phase [91]. All trials involve some risk to participants, which also increases the ethical obligation to publish the results [130].

Although guidelines exist for the reporting of definitive main clinical trials [42, 131], until recently there have been no equivalent guidelines for reporting pilot and feasibility studies. Extensive methodological work amongst experts has now led to the development of an extension to the 2010 CONSORT statement [131], to standardise the reporting of external randomised pilot and feasibility studies done to inform a future RCT [79, 132]. The statement does not, however, extend to internal pilot studies that are built into the design of a main trial or to non-randomised pilot and feasibility work, though the authors state that many of the principles will also apply to these types of study [79, 132]. The work involved establishing a working group and conducting a literature review, engaging stakeholders (including the CONSORT group, journal editors and publishers, the clinical trials community and funders), a Delphi process and a consensus meeting over the course of 5 years [80]. The work also resulted in a new open access journal 'Pilot and Feasibility Studies' coming into print in January 2015, specifically to encourage the publication of pilot and feasibility study reports and protocols in all areas of research [133, 134]. It is hoped that adhering to the CONSORT extension for the reporting of pilot and feasibility studies will significantly improve the dissemination of such work, and consequently improve the consistent and transparent sharing of ideas, methods, findings and lessons learned.

In summary, the combination of a lack of established and widely used definitions, the narrow scope of issues explored, and both inconsistent and under reporting, may have previously hindered the perceived value of pilot and feasibility work.

Whilst the importance of pilot and feasibility work is acknowledged amongst experts in the field, and extensive methodological work has been done to try and improve the reporting of pilot and feasibility work, this view is as yet, not understood or accepted in the wider trials community. Trials, therefore, may continue to be designed and conducted without the optimisation that learning from pilot and feasibility work would afford. With reference to surgery specifically, previous literature searches of the top 10 medical journals and the journal 'Trials' (01/01/11 to 31/08/13) identified 300 pilot studies of which, just five (6%) were in surgery. This work suggests the performance, accurate reporting and comprehensive publication of pilot and feasibility work in surgery is rare [111, 112].

In light of the knowledge accumulated from the published scientific literature both in general and specific to surgery, it was hypothesised that a traditional systematic review, sourcing published surgical pilot and feasibility studies from the scientific literature, would have added little to current knowledge and understanding of practice in this area.

### 3.2.2 Rationale for a targeted review and systematic analysis of protocols of funded pilot/feasibility studies

As established above, a narrative or systematic review of the literature regarding reported surgical pilot and feasibility studies would not have been helpful due to the ingrained challenges of inconsistent and under reporting of pilot and feasibility work in the literature. It was hypothesised, therefore, that a targeted review and systematic analysis of protocols of pilot and feasibility studies submitted for peer-review and successfully funded by established national bodies, might be more abundant and have more thorough methodological designs and describe those methods in more detail, than reports of pilot and feasibility studies published in the general scientific literature. The protocols of funded pilot and feasibility studies may, therefore, be more likely to contain relevant detail to aid understanding of how pre-surgical trial work may optimally inform future

definitive studies, and therefore, be of greater relevance to the overall aim of this thesis.

A narrative review of published reports of funded pilot and feasibility projects would also have had significant limitations. Narrative literature reviews do not systematically select articles for inclusion, and are therefore subject to bias [135], as all the available data sources are not necessarily identified, included and objectively analysed. Narrative reviews may also be criticised for lack of transparency as the unsystematic approach means that the methods are not easily reproducible. A narrative review of funded surgical pilot and feasibility studies would also not have yielded the detailed information desired, to explicitly identify the issues around the design and conduct of such studies.

Focusing on reviewing research funded by the National Institute for Health Research (NIHR) was hypothesised to be the most efficient use of time and resources within the limitations of this doctoral research. Other potential funders of surgical pilot and feasibility studies include, for example, the British Heart Foundation (BHF) and the Medical Research Council (MRC). However, whilst pilot and feasibility studies are funded by other groups, the NIHR is the major funder of studies of surgical interventions in the UK. The limitations of considering only NIHR funded pilot and feasibility work for this thesis, are further discussed in Chapter 7, section 7.4.1. The methods used for a targeted review and systematic analysis of protocols of surgical pilot and feasibility studies funded by the NIHR from are described in detail below.

## 3.3 Definitions used in this thesis

To begin this analysis of pilot and feasibility study protocols, it was necessary to first consider what a surgical trial is, and consequently therefore, what a surgical intervention is, to enable formation of inclusion and exclusion criteria for the review. It was also necessary to define pilot and feasibility work for the purposes of this thesis, in the absence of a universally accepted definition. The working definitions of a surgical trial, surgical intervention and a pilot/feasibility study are described in detail below. These definitions were developed by the author of this

thesis in light of consideration of definitions already available, the limitations of these definitions and the author's personal experience of surgical practice and what this entails. Iterative discussion with other researchers (Avery/Blazeby) involved in this work also took place, informed by early scoping searches of the NIHR databases, to ensure that the definitions of both a surgical intervention and a pilot or feasibility study, were inclusive and relevant to modern surgical practice.

### 3.3.1 What is a surgical trial?

There is currently no consensus on what a surgical trial constitutes, because there is no consensus on the definition of what a surgical intervention is. The noun 'surgery' is conventionally defined by the Oxford English Dictionary as "*The art or practice of treating injuries, deformities, and other disorders by manual operation or instrumental appliances*" [136]. Simplistically, a surgical trial could therefore be defined as one involving surgery as an intervention [137]. However, surgical trials could also include those where surgery is performed, for example, as a co-intervention to oncological treatment [138] or other medical therapies [139]. Such trials are traditionally considered as oncological or pharmaceutical trials of medical therapies. However, surgery is often performed as a co-intervention to chemo- or radiotherapy or other pharmacological treatments in a pragmatic way, with no assessment or control. Additionally, surgical trials could include those where the intervention is not surgery in the traditional sense, but the intervention is performed by a surgeon. This would include trials of therapeutic endoscopy procedures and the medical management of surgical conditions such as acute pancreatitis or diverticulitis [140]. Trials of the management of surgical conditions through the intervention of other healthcare professionals, may also be considered as surgical trials. This would include interventional radiological procedures, for example, to place a drain for an infected intra-abdominal collection, angioplasty for an ischaemic limb, or a stent for an obstructed colon. Patients with these conditions would be admitted under the care of surgeons and have their treatment decisions and management co-ordinated by surgeons, but the management plan

may include the intervention of other professionals, for example, physiotherapists or radiologists.

In summary, surgical trials could be considered to more broadly encompass: i) trials where surgery is the main intervention; ii) trials where surgery is a co-intervention; iii) trials where no surgery is performed but surgeons are administering or managing other treatments or; iv) trials where other health professionals are administering treatments to patients under the care and responsibility of surgeons. When we consider trials, it is the complexity of the intervention, which defines the challenges of trial design and the methodology used to design, implement and conduct them. It may be more helpful, therefore, to consider and define what a surgical intervention is.

### 3.3.2 What is a surgical intervention?

Surgical interventions have been described in many ways and there is currently no universally agreed definition. Cook described surgical interventions as: *"…those which involve physically changing body tissues and organs through manual operation such as cutting, abrading, suturing or the use of lasers"* [12]. Blencowe *et al.* expanded on this definition and described a surgical intervention as one: *"…that cuts or physically alters a patient's tissues (whether using a scalpel, stapler, laser or another instrument or device) and involves the use of a sterile environment, anaesthesia, antiseptic conditions and suturing or stapling"* [14]. In a subsequent systematic review published by some of the same authors, surgical interventions were defined as *"…procedures involving an incision with instruments usually performed in an operating theatre and normally involving anaesthesia and/or respiratory assistance"* [15].

The key points in these definitions include cutting or incising a patient's tissues in a sterile environment, involving anaesthesia. They seem to define a more traditional operation in an operating theatre. As surgical techniques continue to evolve, especially in the areas of endoscopy and interventional radiology, these definitions may no longer be broad reaching enough. They lack wider consideration of who can perform a surgical intervention, where it can be performed; using what if any form of anaesthesia, and for what reasons.

For the purposes of this thesis, a surgical intervention was defined as (see also Figure 3a):

**A diagnostic, therapeutic or adjunctive invasive intervention performed by a trained clinician, using hands, instruments and/or devices** where:

**Diagnostic** is defined as – Performed to aid in reaching a diagnosis and;

**Therapeutic** is defined as - Performed to attempt to treat and/or cure pathology and;

**Adjunctive** is defined as – Performed to allow another intervention to proceed and;

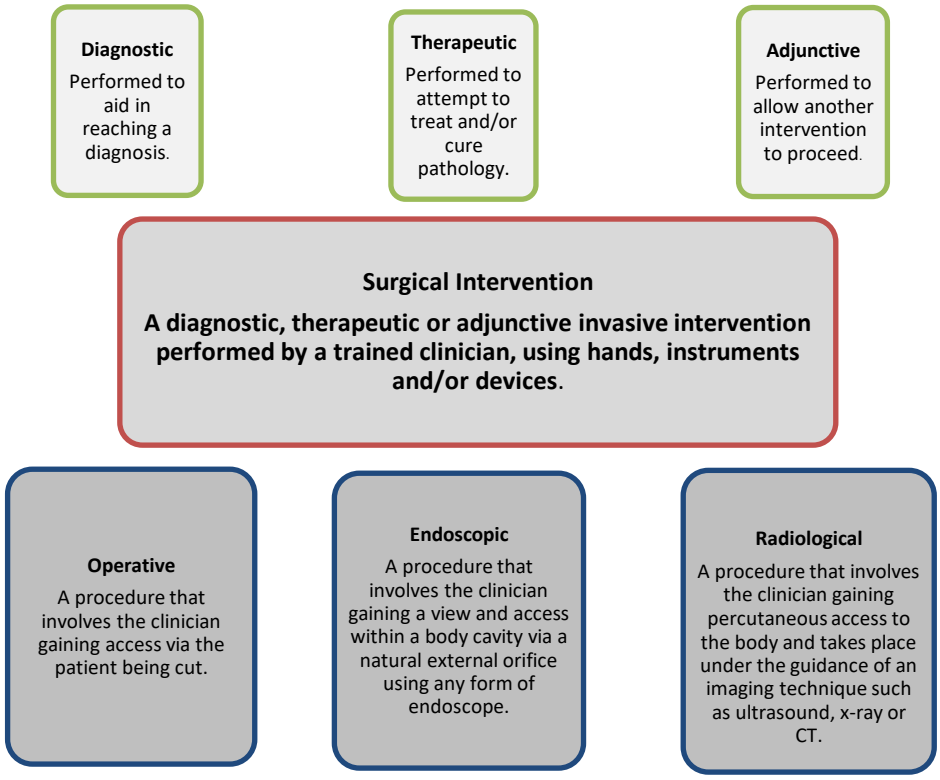
**Invasive** is defined as - Penetrating the body of a patient either via the patient being cut or via a natural external orifice and;

**Trained clinician** is defined as – A registered healthcare professional qualified and trained to perform surgical interventions. This will include surgeons, interventional radiologists, endoscopists, dentists and surgical care practitioners. Importantly, a surgical intervention is not necessarily always performed by a surgeon, but must be performed by a registered doctor, dentist or nurse and;

**Hands, instruments and/or devices** encompasses all surgical equipment and techniques.

A surgical intervention may be performed using anaesthesia (local or general), sedation or with the patient awake and fully co-operative. To clarify further, surgical interventions may also be considered in terms of their method of performance and be sub classified as operative and/or endoscopic and/or radiological (see Figure 3a and Table 3.1 for further clarification and examples).

*Figure 3a* ***Classification of surgical interventions***




**KEY**: CT = Computerised Tomography

This definition does not include diagnostic imaging performed with the administration of intravenous contrast through a simple single-use needle or cannula which is not inserted under image guidance, for example, computerised tomography (CT), angiography or fluoroscopy. Similarly, the administration of drugs via a simple needle or cannula being percutaneously inserted into a blood vessel without image guidance (for example, chemotherapy, intravenous fluids or haemodialysis) are not considered to be surgical interventions. Acupuncture is an alternative therapy administered though the placement of needles at certain points, but this is not done under image guidance, so would not be considered a form of surgical intervention.

Finally, this definition excludes manipulations done by hand, for example external cephalic version (ECV) of a breech or transverse lying foetus, or manipulation of a fractured bone. These procedures do not require penetration of the patient's body and are therefore non-invasive. ECV may be performed by any trained person

with appropriate ultrasound skills and in an environment with access to theatre facilities should complications arise. Many manipulations of fractures are simply performed in the accident and emergency department by trained physicians or nurses, with only children needing to undergo general anaesthesia routinely.

Some procedures cross multiple categories. The adaption of endoscopic or radiological techniques to gain access through the skin or via the gastrointestinal tract lumen to a normally closed body cavity, for example. Such procedures include percutaneous necrosectomy (the insertion of the drain is an adjunctive radiological surgical intervention and the necrosectomy is a therapeutic operative surgical intervention), transgastric necrosectomy (the insertion of the cystgastrostomy is an adjunctive endoscopic surgical intervention and the necrosectomy a therapeutic endoscopic and operative surgical intervention) and percutaneous nephrolithotomy (again the insertion of the stent is an adjunctive radiological surgical intervention and the nephrolithotomy is a therapeutic operative surgical intervention.

The definition described above and used for the work in phase one of this thesis, was used as the preliminary definition for the work published by Cousins *et al* in 2019 [141], to describe and define what an invasive procedure is. The final definition of an invasive procedure described by Cousins *et al* [141], which includes reference to the method of access to the body, instrumentation and requirement for operator skill, was created by applying it iteratively to 3946 included articles published over ten years, therefore verifying that the final definition was applicable to all invasive procedures.

### 3.3.3 What is a pilot and/or feasibility study?

As already discussed in Chapter two, there are many definitions of pilot and feasibility studies. Those relevant to clinical trials and the nuances and difficulties with universal acceptance, have been discussed in detail.

For clarity and inclusivity, the term pilot/feasibility study (PFS) was coined for the purposes of this thesis and defined as:

**Any research that is undertaken before a main study and is explicitly intended to inform the design and/or conduct of a future main study** where:

**Main study** is defined as a definitive study (e.g. RCT) of an intervention(s).

Traditionally, so called internal pilot studies, have been included within the wider definitions of pilot and feasibility studies (see also Chapter two). However, expert opinion has grown over the course of this thesis amongst trial methodologists that internal pilots do not meet the 'true' definition of pilot studies [85]. Internal pilots are very distinct from external pilots in their methodology, being designed and funded as the first phase of a main trial with all data generated from this first phase contributing to the final analysis [84]. Internal pilots are, therefore, most often utilised when no substantive changes to key components of the trial, such as the intervention or outcomes, are anticipated. In addition, study protocols of RCTs with an internal pilot phase usually include only limited detail regarding the internal pilot phase itself, such as a list of proposed progression criteria. It was considered, therefore, that a review of trial protocols with an integrated internal pilot phase, would be of limited value for the purposes of this work. Internal pilots were, therefore, not considered to meet the definition of pilot studies in this thesis.

| Sub-classification | Definition | Examples |
|---|---|---|
| Operative | A procedure that involves the clinician gaining access via the patient being cut. | • Laparoscopic procedures (as cuts to the skin are made to gain access).<br>• Laparotomy for bowel resection<br>• Excision of a breast lump<br>• High tie and strip of varicose veins.<br>• Insertion of peritoneal dialysis catheters (adjunctive operative surgical intervention as the intervention to insert the PD catheter is surgical but the Continual Abdominal Peritoneal Dialysis (CAPD) intervention given via the catheter is not).<br>• Some types of dental surgery where cuts in the mouth are made. |
| Endoscopic | A procedure that involves the clinician gaining a view and access within a body cavity via a natural external orifice using any form of endoscope. | • Simple endoscopy procedures with or without biopsy (diagnostic endoscopic surgical interventions).<br>All therapeutic endoscopic procedures:<br>• Oesophagogastroduodenoscopy and Botox injection/stent insertion/PEG placement<br>• Endoscopic retrograde cholangiopancreatography (ERCP) and stone extraction<br>• Colonoscopy and polyp removal<br>• Sigmoidoscopy and transanal endoscopic microsurgery (TEMS)<br>• Proctoscopy and haemorrhoid treatments<br>• Ureteroscopy and stone extraction<br>• Cystoscopy and resection of a bladder tumour<br>• Urethroscopy and urethrotomy<br>• Hysteroscopy and removal of fibroids<br>• Colposcopy and large loop excision of the transformation zone (LLETZ)<br>• Bronchoscopy and removal of foreign body<br>• Rhinoscopy and removal of polyps. |
| Radiological | A procedure that involves the clinician gaining percutaneous access to the body and takes place under the guidance of an imaging technique such as ultrasound, x-ray or computerised tomography (CT). | • Angiogram and angioplasty<br>• Insertion of a drain to treat a collection<br>• Percutaneous transhepatic cholangiography (PTC) and stent insertion<br>• Insertion of a pacemaker<br>• Simple biopsy procedures under image guidance e.g. liver biopsy (diagnostic radiological surgical intervention)<br>• Insertion of tunnelled lines e.g. Hickman or Pick lines under ultrasound guidance, via a percutaneous needle puncture to the skin and tunnelled within the subcutaneous tissues. The lines themselves are neither diagnostic nor therapeutic but are used for the purpose of administering drugs, nutrition or for monitoring. Such lines would therefore be considered as adjunctive radiological surgical interventions.<br>• Aspiration and steroid injection of joints under image guidance would be classified as therapeutic radiological surgical interventions (percutaneous access is gained under image guidance and treatment given). |

***Table 3.1 Sub classification of surgical interventions in terms of method of performance***

## 3.4 Methods for the targeted review and systematic analysis of NIHR funded surgical pilot and feasibility studies

### 3.4.1 Characteristics of information sources

The UK NIHR Health Technology Assessment (HTA) and Research for Patient Benefit (RfPB) programmes were selected to identify PFS of surgical interventions. These programmes are established major national funders of high-quality patient-centred research. The HTA have funded trials for 26 years (since 1993) and up to £10m and RfPB for 13 years (since 2006) and up £350K. Both programmes fund definitive evaluations of the clinical and cost-effectiveness of interventions as well as feasibility studies to inform future definitive trials. They have publicly available and searchable databases of funded studies [142, 143]. Given the scope and longevity of both programmes, it was hypothesised that each would have funded surgical PFS, providing a sample of potentially well-designed work from which to explore the role of PFS in surgery. Other NIHR funding streams were considered but excluded early on, as they fund surgical research far less commonly. The limitations of this approach are discussed in detail in Chapter seven, section 7.4.1.

### 3.4.2 Search strategy and screening

The HTA and RfPB databases [142, 143] were searched for surgical PFS studies, which met the definitions as described above. Titles and abstracts were screened in duplicate by two researchers (Fairhurst/Rowlands – see acknowledgements), with any issues resolved by discussion (Fairhurst/Rowlands) and/or senior input where necessary (Avery/Blazeby – see acknowledgments). Protocols for all included HTA studies were downloaded from the HTA website and those for all included RfPB studies (apart from one available online) were obtained by contacting the chief investigator of each study directly by email. Additional publications relating to included studies were identified by searching for links to published outputs on the NIHR website (HTA only), and using the study title, acronym and chief investigator name to search on two electronic bibliographic databases, PubMed and Google Scholar, and the International Standard

Randomised Controlled Trials Number (ISRCTN) trials registry online. It was important to search for published outputs, to attempt to consider the outcomes of the PFS, in terms of publication of the PFS and main trial feasibility, development and funding.

In addition to searching electronically, chief investigators of the PFS were also contacted by email once the study end date had passed beyond six months, to request PFS outcome information. Finally, at the end of the PhD studentship period, the NIHR were contacted (see acknowledgements) to corroborate published data sources and confirm any main trial funding granted for the included PFS. This was necessary, particularly for the RfPB funded studies, for which there is currently no requirement for public reporting. In addition, main trials do not always retain the same PFS title, study acronym, or chief investigator, making it difficult to search for subsequent main trials electronically.

### 3.4.3 Inclusion and exclusion criteria

Protocols of all surgical PFS funded by the NIHR HTA and RfPB programmes between 01.01.2005 and 31.12.2015 were included. The most recent ten-year period was chosen for inclusion, as it was known that the methodological developments surrounding the design and conduct of PFS for complex interventions had happened over the last ten to fifteen years. In addition, the RfPB only started funding studies in 2006. To reiterate, in the absence of universally adopted definitions of 'surgical interventions' and 'pilot/feasibility studies', for the purposes of this work, pilot/feasibility work was defined as (see section 3.3.4):

**Any research that is undertaken before a main study and is explicitly intended to inform the design and/or conduct of a future main study** where:

**Main study** is defined as a definitive study (e.g. RCT) of an intervention(s).

A surgical intervention was defined as (see section 3.3.2):

**A diagnostic, therapeutic or adjunctive invasive intervention performed by a trained clinician, using hands, instruments and/or devices and includes operative, radiological and endoscopic procedures.**

As described in section 3.3.3 above, internal pilot studies were excluded as they did not meet the definition of a PFS used in this thesis. Funded systematic reviews that did not state any intention to inform a future definitive study (and therefore did not meet the definition of a PFS) were also excluded. In addition, studies that focused on the evaluation of co-interventions to surgery as the main intervention under examination, for example the administration of anaesthetic drugs, and post-operative rehabilitation or enhanced recovery programmes were also excluded. This was because the primary focus of this work was to explore the specific difficulties surrounding studies of surgical interventions (as defined above and section 3.3.2), rather than interventions that have relevance to surgical practice, but are not actually surgical interventions themselves. Table 3.2 below summarises the inclusion and exclusion criteria for the systematic analysis and targeted review.

***Table 3.2 Inclusion/exclusion criteria for screening NIHR HTA and RfPB funded studies***

| Inclusion Criteria | | Exclusion Criteria | |
|---|---|---|---|
| 1 | Pilot/feasibility work of a surgical intervention | 1 | Not a surgical intervention |
| 2 | Funded by the HTA/RfPB between 2005-2015 | 2 | Surgical intervention, not pilot/feasibility work |
| | | 3<br>3a | Surgical intervention is a co-intervention<br>And not pilot/feasibility work |
| | | 4 | Protocol/publication not in English |
| | | 5 | Study not in humans |

## 3.4.4 Data extraction

Data were extracted using a standardised database developed in Microsoft Excel (See Appendix I for details of data extraction), including general study characteristics, available data sources in addition to the study protocol (published papers) and the surgical specialty of the study. Details of the PFS design (randomised or non-randomised; quantitative or qualitative) and conduct, including characteristics of the patient population, were extracted. A framework was developed for capturing the uncertainties and challenges regarding the viability of a future main trial, informed by expert knowledge and previous methodological work regarding the design, definitions and reporting of PFS [71, 72, 74, 75, 79, 89, 91-93, 95, 96], the published MRC guidelines [13] and the IDEAL framework [101, 144].

All possible reasons identified for undertaking PFS were grouped into five key domains: 1) main trial design; 2) logistics; 3) recruitment; 4) intervention, and; 5) outcomes. When developing the domains, the timeline of trial design and conduct was considered and the domains broadly follow this from logistics (domain 2) in terms of trial processes and set-up, to outcomes (domain 5) as the end point of trials. The first domain of 'main trial design' consisted of predominantly overarching uncertainty questions, such as the feasibility or necessity of a main trial, the trial costs, and sample size calculations.

The domains were constructed and ordered as data extraction progressed with cross checking between researchers (Fairhurst/Avery). The framework was modified in an iterative manner as data collection and extraction progressed to be responsive to the emerging data, as further reasons for undertaking PFS were identified. Special consideration was given to uncertainties and challenges considered more specific and/or relevant to surgical trials.

### 3.4.5 Data analyses

Results were analysed within Microsoft Excel, using simple descriptive statistical methods (means, medians, percentages). Data were tabulated and descriptive statistics are reported in Chapter four, with comparison between the HTA and RfPB and randomised and non-randomised cohorts where relevant. Analysing the HTA and RfPB cohorts separately was deemed important to consider any differences between the funding streams, in terms of the types and quality of PFS being funded.

# Phase II: Exploring Perceptions and Experiences of Pilot Work for Surgical Trials: A Qualitative Research Study (PEPSTAR)

## 3.5 Rationale for Phase II of this thesis

The rationale and methods for the qualitative interview study to explore key stakeholder's views, perceptions and experiences of surgical PFS will now be described. In phase I of this thesis, the quantitative targeted review and systematic analysis of NIHR funded surgical PFS protocols, the methods of which are described in sections 3.2-3.4 of this chapter, aimed to identify the type and frequency of surgical PFS being funded and performed in the UK. In addition, the type and frequency of misunderstandings and potential challenges in this area, that may be preventing the optimal design and conduct of PFS and thereby hampering definitive surgical trials were sought.

To improve understanding of why challenges and barriers may exist to optimally designing and conducting PFS in surgery, an in-depth exploration of the views of key stakeholders was considered necessary. The aim of this 'PEPSTAR' study was, therefore, to use qualitative methods to explore in detail the experience, perceptions, attitudes and opinions of surgeons, trialists, methodologists and funders of the role of PFS in surgical trials and identify any barriers and potential solutions. A detailed rationale for using qualitative methods, followed by a description of the specific methods used in the PEPSTAR qualitative interview study are reported below.

### 3.5.1 Rationale for using qualitative methods in Health Services Research

Qualitative research is very simply defined as study of the social world, and has been described as a *"very broad church"* [145] crossing a range of techniques, approaches and research disciplines. Over the last century, the field has been influenced by many disciplines such as sociology and anthropology, and

numerous different schools of thought have emerged resulting in extensive variation regarding the purpose and practice of qualitative research methods [146]. The wider applications of qualitative research are increasingly recognised with the realisation that such techniques may complement and enhance more traditional quantitative methods in answering questions more effectively and efficiently [147].

Qualitative research is rooted within the social sciences and very simply, seeks to understand 'why' and 'how' social phenomena occur, by interpreting data on the perspectives of people involved [148]. It may be performed as standalone research, or in combination with, and complimentary to, quantitative research (such that the combination of these different methods is termed 'mixed-methods' research) [149]. It is now widely recognised that in the context of Health Services Research, the importance of understanding people from the perspective of being both providers and/or recipients of services and treatments, is essential [146]. The widespread inclusion of qualitative research in areas such as Health Services Research and health technology assessment is, therefore, now fairly common practice in the UK [149-151]. However, the robustness of all research depends on the quality of the methods used. As qualitative research is such a broad field it is, therefore, vital that the specific methods used in any project, are described and justified in the context of the research question being asked.

### 3.5.2 Previous qualitative work exploring stakeholders' views towards pilot and feasibility work for surgical trials

No previous qualitative work has specifically explored the experiences and perceptions of surgeons, trialists, methodologists and funders regarding the role of and challenges to completing pilot work for surgical trials. However, qualitative work has been widely utilised in surgical trials more broadly to, for example, explore barriers to recruitment to surgical trials [17, 24, 152-155], explore different surgical trial designs [156], and train surgeons in conveying equipoise and optimising recruitment techniques to surgical trials [157-159]. With the knowledge that qualitative work has been proven to optimise the success of surgical trials [157] such work is therefore, often deemed essential for the design, conduct and

optimisation of surgical trials, though it is often not well reported in the scientific literature [160].

Qualitative work is therefore justified as important to explore why PFS in surgery are not currently optimised, through better understanding key stakeholders' perspectives, experiences and views, and to formulate solutions for how PFS for surgical trials might be pragmatically optimised in the future.

A description and overview of various qualitative research methods that are considered relevant to achieve the aims and objectives of this work now follows. A justification and detailed description of the qualitative methods used for the PEPSTAR qualitative interview study will then be described in detail.

## 3.6 Philosophical approaches to qualitative research

Qualitative and quantitative research methodologies are rooted in philosophy, being influenced by different ontological and epistemological assumptions, often via embedding of such assumptions within a particular research paradigm. Ontology is defined by the Oxford English Dictionary as*: "The science or study of being"*, and epistemology as: *"The theory of knowledge and understanding, especially with regard to its methods, validity and scope, and the distinction between justified belief and opinion"* [136]. Both methodologies focus on defining a research paradigm as a theoretical framework, by which to describe a belief system about knowledge and understanding relating to a given topic. The research paradigm (or theoretical framework) includes the ontological and epistemological assumptions made as part of this belief system [148]. Different research paradigms, such as the posited opposites of positivism and interpretivism, therefore assume different ontological and epistemological approaches to each other (see Table 3.3).

| Paradigm | | Ontology | Epistemology | Methodology |
|---|---|---|---|---|
| Positivism | | Objective reality or truth | People know, accept and can measure this reality | Hypothesis testing<br>Usually quantitative |
| Interpretivism (also termed Constructivism) | | No single reality or truth | Reality needs to be interpreted (or is variable constructed by different individuals) | Hypothesis generating<br>Usually qualitative |
| Postmodernism | | A fragmented ever-changing reality | Many ways of understanding reality (epistemological pluralism) which may include e.g. premodernism (revelation), modernism (science & reason), intuition etc. | Wide range of methods |
| Realism | Critical | Ability to know an objective reality is imperfect | Objective reality is an ideal to strive for through research, but we cannot separate ourselves from our own knowledge | Combination of qualitative and quantitative |
| | Subtle | Can only know an objective reality from own perspective | | |

***Table 3.3 Qualitative research paradigms*** [148, 161, 162]

The paradigms of positivism and interpretivism can be considered as opposing with respect to the view that the former focuses on hypothesis testing using quantitative methods, and the latter hypothesis generating using qualitative methods. The so termed "paradigm debate" encompasses consideration as to whether qualitative and quantitative methodology can be combined [163]. Qualitative research has mostly emerged from the concept of interpretivism, as an alternative to positivism, meaning that there may be many realities and all of these need to be interpreted in the context of social, historical and individual influences (or alternatively reality is 'constructed' by the influence of these contexts; so termed constructivism) [148]. However, positivism and interpretivism should probably not be considered simply as polar opposites [162]. For any research, the approach, or combination of approaches will be determined by what one is trying to achieve [162], and therefore, any single research paradigm may not sit entirely separately from another.

It is suggested that our ontological beliefs, and epistemological understandings are interconnected with the methodology and methods and none can therefore, be viewed in isolation [164]. Many argue that subscribing to a single research paradigm

is unhelpful and restrictive, and that doing so is neither *"necessary or desirable"* for qualitative research [165]. As a result of recognition that different research questions require different approaches, other recognised research paradigms have emerged which blend and build on positivism and interpretivism.

Modern qualitative research, for example, also considers the characteristics and experiences of the researcher(s) and how these may influence the methodology, data collection and analysis (a concept known as reflexivity – see sections 3.7.4 and 3.12) [166]. This is based on the understanding that, at the core of all qualitative research, is the concept that individuals may interpret the world around them differently, depending on the social contexts to which they are subjected [148]. Paradigms such as critical realism [167] (see table 3.3), for example, acknowledge both the way individuals may interpret the world around them, and how the broader social constructs may impact on those interpretations, whilst also retaining a focus on the data itself.

The philosophical basis for research, may also be described as the 'theoretical lens' that underlies the researcher's choices and assumptions, and through which the data collection, analysis and interpretation of the findings is viewed [168]. Primarily, it is key that the researcher identifies if a philosophical paradigm is subscribed to, and clearly describes and justifies the choice (or not) of philosophical paradigm, and how it may affect the interpretation of the results.

## 3.7 Ensuring rigour of qualitative research

There has been much discussion about how to ensure the scientific rigour of qualitative research [162, 164]. Much of this discussion has been around whether it is possible to judge qualitative research by the same quality criteria as quantitative research [150, 169]. This discussion opens a wider epistemological debate about the nature of knowledge produced by qualitative research and if and how the quality of this knowledge and research can feasibly be critiqued. A realist approach, as described in section 3.6 above, assumes that an objective reality or 'truth' is a goal to aim for through research, but might never be fully attained. In the absence of an absolute 'truth', it therefore allows the assessment of the perspectives and values

of different research methods in relation to each other [150]. The various ways of improving and justifying the quality of qualitative research are outlined below.

### 3.7.1 Clear description of qualitative research methods

All research is a product of good (or bad) methodology. A clear account of the process of data collection and analysis is, therefore, essential to determine if and how the research was thoroughly conducted. Unfortunately, it has been observed that many qualitative researchers have neglected to give an adequate description of the assumptions and methods used, which has probably contributed to criticisms of bias and lack of rigour from quantitative researchers [169]. A detailed description of the various qualitative methods available, and how these may be optimally used, is documented below.

### 3.7.2 Qualitative sampling

#### *Qualitative sampling techniques*

The purpose of qualitative research is to consider a more holistic view of the phenomena being investigated [170]. Achieving a statistically representative sample of the population of interest is, therefore, not the aim of qualitative research. Instead, purposive or theoretical sampling strategies, also termed non-probability sampling, are employed to select participants, settings or events with particular characteristics that relate to the research question [169].

**Generalisability** is defined as the extent to which the findings can be applied more widely to other settings or populations [171]. It is a method used in quantitative research to determine the importance or relevance of the research more widely. A criticism of qualitative research has been the difficulty of demonstrating generalisability because of a lack of systematic probability sampling [169, 171].

However, there are many types of purposive sampling, that range from the heterogeneous (sampling to capture a wide range of perspectives relating to the phenomena being studied), to the homogenous (aiming to achieve a sample sharing very similar characteristics). Sampling techniques also have variant terminology, so it is most important to describe what technique was used and

why. Though qualitative samples do not seek to be statistically representative, if a thoughtful, well informed and described sampling strategy is used, it minimises bias from convenience sampling and produces a sample most relevant to the research question. Different purposive sampling techniques include:

**Maximum variation sampling**: Also termed **heterogeneous sampling**, is used to capture a wide range of perspectives of the phenomena being studied. The populations being studied may, therefore, exhibit a wide range of characteristics in terms of attributes, behaviours, experiences, opinions, for example. A maximum variation derived sample should include typical, extreme/deviant and critical cases (see below);

**Typical case sampling:** Refers to selecting the most usual, commonly occurring or ordinary cases of the phenomena being studied;

**Extreme/deviant case sampling:** Aims to sample cases that may offer unusual or rare examples of the phenomena being studied;

**Critical case sampling:** Samples cases with unique perspectives of the phenomena being studied that may provide particularly decisive or important data;

**Homogeneous sampling:** Focuses on sampling cases with very similar attributes, behaviours, characteristics, experiences and so on. Homogenous sampling is the opposite of heterogeneous sampling (maximum variation sampling);

**Theoretical sampling:** First described by Glaser and Strauss [172], this is a form of purposive sampling, where the sample is selected based on the potential contribution to the development of a hypothesis. It involves an iterative approach where an initial sample is selected, the data analysed, and a further sample then selected to refine the emerging theories.

**Snowball sampling:** This sampling technique involves asking current participants to suggest other participants who may usefully contribute to the research through adding additional or different perspectives to the phenomena being studied. Snowballing can also be used to identify extreme/deviant or critical cases, for example.

In practice, a combination of sampling approaches is often utilised. Pragmatic factors such as time and availability of resources may also legitimately contribute to decisions regarding the sampling method [173].

In terms of sample size, for qualitative research this is usually small because phenomena only need to occur once [145]. This means that, as the data are analysed, there will be a point at which **data saturation** is reached and no new themes are emerging and established themes cease to evolve [148]. Data saturation was first coined as a term, in Glaser and Strauss' original publication on grounded theory [172] to signify when data collection and analysis could cease. Since then, extensive methodological discussion in the literature has identified different models of data saturation, which should be operationalised dependent on the research question, theoretical position and data analysis approach taken [174]. In particular, it is advocated that a 'point' of data saturation may not always be appropriate and that defining saturation should be considered as a process, where further data collection becomes *"counterproductive"* [175] or a matter of *"diminishing returns"* [176]. Data saturation may therefore be considered as a *"cumulative judgment"* which considers whether *"sufficient depth of understanding has been achieved in relation to emergent theoretical categories"* [174]. Additional reasons for a moderate sample size include that there is no requirement for statistical analysis and an appropriately powered sample size to determine statistical significance, and that the data are very rich in detail meaning analysis is resource intensive [145].

## 3.7.3 Qualitative data collection

### *Types of qualitative data collection*

The main types of data analysed in qualitative research include direct observations, interviews, and documents.

#### Observations

The systematic study of people and cultures, designed to explore the cultural phenomena from the perspectives of the study subjects, is called ethnography. Predominantly, ethnography involves observation of people and events within their natural or usual setting, either as an immersed participant or as a non-

participant observer. In the former, the researcher immerses oneself within the group, which may allow a unique perspective for gaining deep understanding about the nuances of social interactions [166].

**Interviews**

Individual interviews with participants may be structured or semi-structured. Structured interviews use a questionnaire which delivers the same questions to all participants, whereas semi-structured interviews are usually based around a topic guide, providing a basic structure around which to ask open-ended questions, therefore allowing experiences to be explored and topics of importance to participants to emerge. Regardless of the degree of structure, in-depth interviews allow specific topics or experiences to be explored in more detail.

Focus groups are a form of group interview in which a small group of participants are interviewed together, and their responses to each other, as well as to the interviewer are collected. This group interview technique is useful for exploring shared experiences and, for example, identifying contentious issues and achieving consensus [164].

**Documents**

Ethnographic data collection may also be performed by gathering documents important to the population being studied, such as minutes of meetings, diaries, or photographs [177].

### 3.7.4 Qualitative data analysis

One of the major ways in which qualitative research differs from quantitative research is in the plurality of data analyses and interpretations, whereas in quantitative research the only route from data to conclusions is via statistical or other numerical analysis [171]. Qualitative data analysis begins during data collection and is almost inevitable as the researcher is 'in the field' and naturally beginning to process the data being collected during, for example, one-to-one interviews [178]. The term 'grounded theory' describes an approach first described by Glaser and Strauss [172], which aims to develop theories around phenomena that are 'grounded' in systematic examination of the qualitative data. The key concepts

of grounded theory are that the data are analysed in an iterative and cyclical approach of data collection, using theoretical sampling, and simultaneous analysis. This type of qualitative data analysis may be described as inductive, where data are collected and analysed and theories emerge from the data itself. An inductive approach to data analysis is opposite to a deductive approach, where theories or hypotheses are developed *a priori*, and tested against the analysed data [148]. In the inductive approach consistent with grounded theory, data are collected, analysed, and further participant sampling takes place based on emerging theories, to confirm or challenge these emerging theories [179]. Therefore, a principle of constant comparison is central to grounded theory, where the data are compared and contrasted to emerging theories to allow those theories to develop and evolve into well-grounded theories, induced by the data [179].

Thematic analysis, first named and explicitly described in the 'how to' guidance published in 2006 by Braun and Clarke [180], is a flexible form of qualitative data analysis which is independent of any specific theory or epistemology. Braun and Clarke describe how thematic analysis includes many of the concepts of grounded theory, but without needing researchers to "*subscribe to the implicit theoretical commitments of grounded theory*", though also emphasising that it is still essential to be clear on what philosophical assumptions have been made. Thematic analysis is, therefore, proposed by Braun and Clarke [180] as a more accessible form of analysis for those in the early stages of a qualitative research career. Consequently, thematic analysis is a commonly used approach to analysing qualitative data [177].

The Braun and Clarke guidance [180] describes the six stages of thematic analysis as: 1. Familiarising yourself with the data; 2. Generating initial codes; 3. Searching for themes; 4. Reviewing themes; 5. Defining and naming themes and; 6. Producing the report. The authors describe how thematic analysis can be applied in an inductive or 'bottom up' way, where data are collected specifically for the research (for example, interviews) and analysed without trying to fit it into a pre-existing framework, or the researcher's preconceptions of the subject area. Alternatively, Braun and Clarke describe how thematic analysis can be applied in a deductive or theoretical 'top down' way in which a framework is derived at the start of data analysis and the data coded against this framework. Braun and Clarke describe

how thematic analysis can, therefore, take many different approaches, but a key concept of their thematic approach is that the final account details what approach was taken and why [180].

### *Ensuring validity and reliability of the data analysis*

**Computer software packages** have become increasingly popular to aid qualitative data analysis [169, 178]. They can be used to assist data organisation, sorting and retrieval and also some forms of analysis, for example by using algorithms to identify repeatedly occurring codes [178]. However, it has been widely noted that computer software such as NVivo [181] and ATLAS [182] are not complete methods of analysis in themselves and are, therefore, unable to make the sometimes complex conceptual links between data and emerging theories [178]. It has been proposed that computer software packages may improve the transparency of qualitative data analysis (by producing a well-documented audit trail of the analysis process) and consequently improve reliability [146].

**Multiple coding** has been advocated as another method to improve the reliability of qualitative data analysis. In this approach, independent assessment and coding of the data by more than one researcher occurs, in an attempt to improve inter-rater reliability [169]. However, the purpose of the multiple coding is mostly to identify potentially different explanations for the findings, and therefore refine the emerging themes [183].

The process of **respondent validation** or **member checking** seeks to feed back the research findings to the participants, with the aim of establishing correlations and incorporating any feedback from participants into the final study findings. However, it is thought that this method has limited value in improving qualitative research validity [150]. Researchers seek to summarise and theorise the data, but participants offer a personal and potentially unique view, which may be at odds with the overall study findings. Respondent validation may be more constructively considered as an additional research process to reduce error, which also generates additional data requiring further interpretation [184].

**Triangulation** refers to a process by which data or evidence is sought from two or more independent sources. Whilst often proposed as a method of ensuring

validity it perhaps more accurately describes a method to ensure comprehensiveness and encouraging a reflexive approach to the data [150].

Researchers should also thoroughly examine the data for **deviant or negative cases** which are defined as those data by which the researcher's explanatory theory appears weak or contradicted. This allows emerging theories to be further evolved and refined, or at the very least demonstrate a range of viewpoints [150].

**Reflexivity** refers to being mindful of how both the researcher and research methods have shaped the data collection and analysis. The effects of age, sex, social class, professional status and prior knowledge of the research area can all have an affected on the data collection and analysis (even if unintended). It is vitally important for the validity of the research findings, to explicitly report contextual details around personal or intellectual biases [171].

# 3.8 Qualitative methods used for the PEPSTAR study

## 3.8.1 Ethical approval

Ethical approval was sought for the PEPSTAR study through the University of Bristol Faculty of Health Sciences Research ethics committee (FREC) and granted on 13th October 2016 (Application Number 41001 – See Appendix III to thesis: Favourable opinion letter).

## 3.8.2 PEPSTAR philosophical approach

A critical realist approach was taken to the project conduct and interpretation of the data. A critical realist approach (or theoretical lens), as described in section 3.6 earlier in this chapter, acknowledges that our ability to know an objective reality or 'truth' is imperfect, and that our interpretation of this truth is influenced both by our own knowledge and by the broader social constructs around us. As already established in chapter two, there is no universally accepted truth regarding the definitions, design and conduct of surgical PFS. However, there are published available definitions of PFS and methodological frameworks for designing and

conducting PFS (for example, the IDEAL framework and the MRC complex interventions guidance) in existence, though these are considered imperfect.

In terms of the PEPSTAR interview study, the objective reality, or 'truth', was considered as the available, though imperfectly and variably perceived, definitions and current methodological understanding of PFS as thoroughly described in Chapter 2. A critical realist philosophical approach, or theoretical lens, is therefore justified as appropriate for this work. Specifically, an objective reality currently exists around the design and conduct of PFS, and the perception of this reality may vary and be influenced by the professional roles and experiences of both the participants, and the researcher.

### 3.8.3 PEPSTAR sampling strategy

For PEPSTAR, a list of potential participants was created through considering:

- Trial team members of the studies included in the targeted review and systematic analysis of NIHR funded PFS;
- Team members of other PFS/RCTs of surgical interventions in the UK;
- Senior members of Clinical Trials Units (CTUs) and Royal College of Surgeons Trials Centres in the UK;
- Senior authors of published literature on PFS work methodology;
- Senior editors of surgical journals and those publishing PFS specifically;
- Senior funding panel members of major funding bodies in the UK including:
    - National Institute for Health Research (NIHR)
        - Health Technology Assessment Commissioning Board (HTA CB)
        - Health Technology Assessment Clinical Evaluation and Trials Board (HTA CEAT)
        - Health Technology Assessment General Board (HTA GB)
        - Research for Patient Benefit (RfPB)
        - Efficacy and Mechanism Evaluation (EME)
        - Programme Grants for Applied Research (PGfAR)
    - Cancer Research UK (CRUK)

    - Chief Scientist Office (CSO)
    - Arthritis Research UK (ARUK)
- Senior surgeons with extensive experience of trials research and methodology.

A matrix was developed, including the names, places of work, surgical speciality (for surgeons), areas of expertise, and research roles as a chief investigator, journal editor and/or funding panel member. Seniority was deemed important to ensure extensive experience and involvement with the design, conduct and/or funding of PFS and surgical trials. The matrix of potential interview participants was compiled with the aim of including a broad range of key professional stakeholders involved in designing, funding, conducting and publishing PFS for surgical trials, including surgeons, trial methodologists, funders and journal editors. Potential participants were purposively sampled from the matrix, by considering all the relevant characteristics listed and through discussion with the other researchers involved (Blazeby/Avery/Potter) in order to achieve maximum sample variation across both the clinical and non-clinical participant groups.

Only participants from and/or working in the UK were sampled. International participants were not considered both due to the time and resource limitations of this doctoral thesis, and the relatively different research funding strategies and systems in other countries (see Chapter seven, section 7.4.2 for further discussion of limitations of the sampling strategy). Clinical participants included those working in a range of surgical specialties throughout the UK and those with experience of research as a chief investigator or trial team member and/or a journal editor, and/or a funding panel member. Similarly, non-clinical participants included those working throughout the UK with extensive methodological experience (e.g. as a CTU director or senior statistician) and/or holding a senior position on a funding panel (e.g. NIHR HTA/EME/RfPB, CRUK) and/or a senior editorial position for a surgical journal, and/or experience of designing and conducting surgical trials.

Potential participants were approached by email or face to face at conferences, meetings and other networking opportunities. A copy of the letter sent and/or given to participants can be seen in Appendix IV to this thesis. All participants

received information about the study before the interview (Appendix V: Participant information sheet). In addition, snowballing sampling was employed, by asking participants during the interview, if they had suggestions for other participants who may have relevant or important perspectives for this work. In this way, a variety of sampling strategies were used to ensure that the final sample was heterogeneous and inclusive of a wide range of perspectives and experiences.

### 3.8.4 PEPSTAR data collection

In-depth, semi structured interviews were performed with all consenting participants. The interviews were conducted either face to face or by telephone, at times and locations convenient to the participants. A topic guide was used to inform the interview questions (Appendix VI).

The topic guide was developed by initially mind-mapping a list of questions relevant to the research question. Key concepts and details from the methodological literature and background data and the findings from the quantitative work (Phase I) were considered, to frame questions which explored why PFS are not being optimally designed and conducted. The guide was then ordered to consider how a conversation about these topics would flow sequentially and with natural fluency.

Some introductory questions and gathering of general information were posed at the start of the topic guide (sections one and two). The third section dealt with the different topic areas to be explored. The purpose and importance of PFS more generally and then specific to surgery were considered first, to gauge broad understanding of PFS. For example, participants were asked how they defined PFS and whether they thought their view was the general view. Another question posed was what role participants perceived PFS to have, and what role they thought it should have. This led sequentially into asking participants about their experiences of PFS, in terms of design, conduct and funding. With personal experiences then in mind, questions were posed around the challenges of difficulties of PFS and components considered essential to the design, conduct and funding of PFS.

The topic guide was updated, but not substantially changed after the first eight interviews. The changes mainly related to the order of questions, for example, the impact of PFS was considered with the questions on importance earlier on. In addition, further prompts on topics that had emerged from the first phase of interviews were added (see Appendix VIa).

Whilst the topic guide was used as a tool to frame the interview and provide prompts for the interviewer, the semi-structured approach to the interviews was the chosen method for this research project. This allowed guidance of the conversation to relevant topics of interest around the design, conduct and funding of PFS, whilst also allowing topics of importance to the participants to emerge naturally and be explored further as appropriate. A focus group methodology may also have offered insightful findings but could have hindered exploration of specific experiences and opinions. For example, because a focus of the data collection was exploration of the challenges and difficulties of doing PFS in surgery, a focus group may have led participants to hide contentious or critical issues around funding and funding bodies, for fear of causing offence or jeopardising future research funding. In addition, participants may have hidden their misunderstanding from others for fear of losing face in a room full of experts. It would have been possible to split the sample into clinicians/non clinicians, but with a lot of cross categorisation, this would not have been a simple task. In addition, all participants sampled, were extremely busy professional people working all over the UK. Finding a convenient time to bring groups together at a single location would have, therefore, been logistically challenging.

Written consent was taken either before (if done by telephone) or at the time of the interview in person (Appendix VII: Consent form). The interviews were audio recorded using an encrypted tape recorder, transcribed verbatim in full by a third party (see acknowledgements), and anonymised for the purposes of data storage and anonymity of participants.

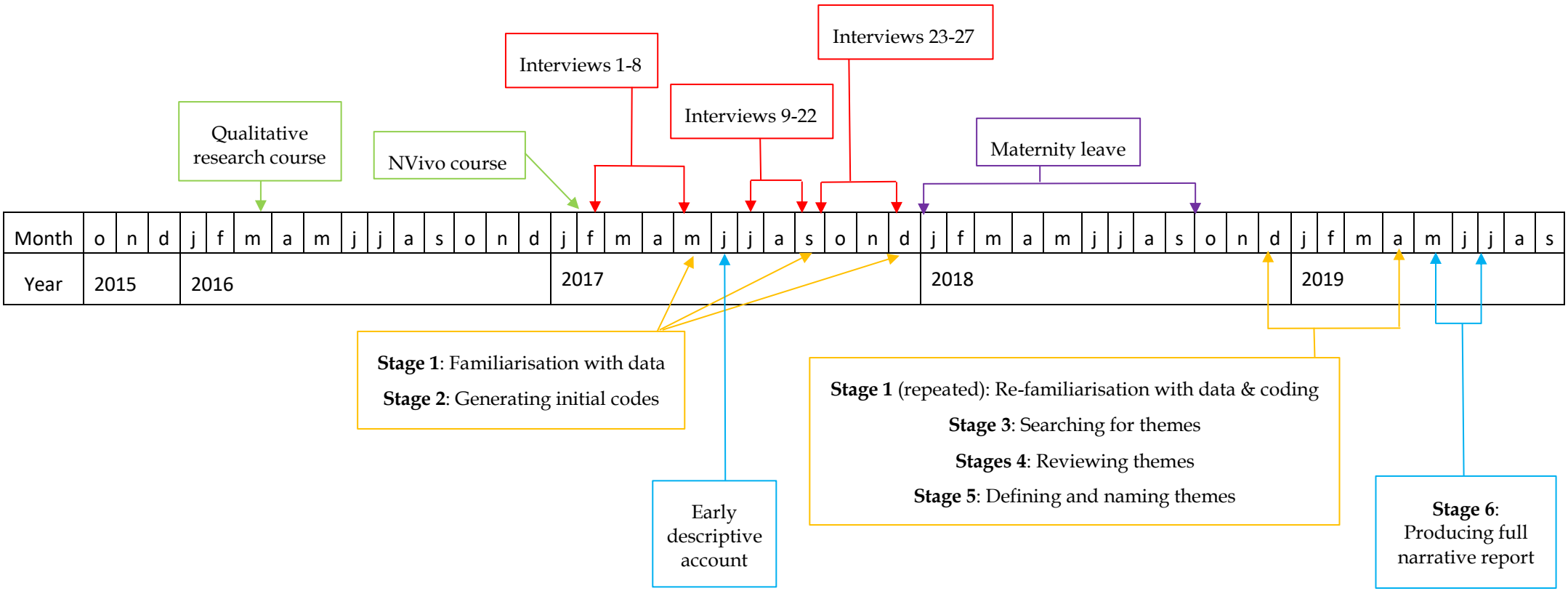
### 3.8.5 PEPSTAR data analysis

The analysis was undertaken using a thematic inductive approach, underpinned by the principles of grounded theory, as described above and in the Braun and

Clarke publication [180]. This method was chosen as it is considered suitable for early career researchers allowing flexibility in its approach for those new to the field [177, 185]. As described in section 3.7.4 in this chapter above, the thematic inductive method offers a step-by-step approach to thoroughly and completely analysing qualitative data, with a reasoned assessment of validity and reliability. As also detailed above, reflexivity is considered a vitally important method to increase the validity and reliability of qualitative research, so reporting these details is, therefore, an essential part of any qualitative data analysis. Reflexive considerations were made throughout the design and conduct of the PEPSTAR study, and also in phase 3 of the work, to synthesise the findings from phases one and two. These considerations are discussed towards the end of the chapter in section 3.12.

The timeline in Figure 3b demonstrates when data were collected and analysed. It should be noted that the author of this thesis took a period of maternity leave from January 2018 to October 2018. At the point the maternity leave was commenced, the interviews had been completed, transcription and coding had occurred, and the process of searching for themes had begun. Following the period of maternity leave, it was necessary to re-familiarise with the full data set, before beginning the process of reviewing, refining themes and subthemes and completing the full narrative account.

***Figure 3b*** *Timeline of qualitative data collection and analysis during the PhD studentship, with reference to the six stages of thematic analysis described by Braun and Clarke* [185]



### *Thematic inductive data analysis*

The thematic inductive data analysis method used in phase two of this work is now described.

Firstly, familiarisation of the data was done through listening to the interview files and making field notes immediately after the interviews, then reading and re-reading the transcripts and field notes alongside listening to the interview files and checking for accuracy. Generation of an initial 'mind map' of numerous codes and groups of codes was produced as familiarisation with the data progressed. The data from the first eight interviews were then collated into the codes identified using NVivo 10 software. The use of NVivo software allowed organisation and coding of the data, to improve visibility and transparency of the data analysis, though only the basic functions of this software were used.

The first eight interviews and field notes were read, listened to, and double coded by a senior (non-surgeon/clinician) researcher (Avery – see acknowledgements), to check understanding of the methodological process, consider reflexive influences, and discuss emerging codes and groups of codes (early potential themes). An early descriptive account was written after the first eight interviews, to further understanding of the methodological process of qualitative research (see Figure 3b above). As further interviews were conducted, the process of coding continued, and codes were checked, amalgamated, renamed as appropriate, with new codes being added as necessary. The process of double coding and discussion with a senior, non-surgical, qualitative researcher, contributed to improving the reliability of the findings, and also assisted with beginning to consider the emerging themes and allowing minor updates to the topic guide used in interviews (See Appendix V and Va)

As the interviews and coding process continued, the data were continually assessed for saturation. The interviews and data analysis continued in three main phases, with the precise order of interviews, mainly determined by participant availability. The coding and thematic diagrams (see Chapter five) evolved over the course of the data collection and analysis, and it was clear that no new codes were emerging towards the end of the third phase of interviews. A few of the

interviews in the final phase were conducted via snowball sampling, but these did not add significantly to the cumulative findings. Satisfaction was therefore reached that participants had been widely sampled for a range of characteristics, and that established codes were no longer evolving, and no new codes were emerging from the data. A judgment was made, both from the final interviews taking place, and the ongoing data collection and analysis, that further interviews were unlikely to add significantly to the data set. Data saturation was therefore judged to be complete, as described above in section 3.8.3.

Once the interviews were complete and transcribed, with the familiarisation and data coding process concluded, the process of searching for, reviewing and refining themes began. This process involved analysing the many codes already applied to the data, considering how different data-containing codes related to each other and how they could be amalgamated to create themes or sub themes. It was possible that further interviews may have become necessary at this point, and time was allowed for this but, reassuringly, this was not deemed necessary as no new or disconnected themes emerged.

At this point, a full narrative account was begun, through organising the data for each theme into a coherent and logical written account. Deviant or negative cases within the data, were particularly sought, to make sure that all perspectives, experiences and opinions were illustrated and considered. The story that each theme told individually was analysed in detail, alongside consideration of how the themes interlinked. Sub themes allowed the larger main themes to be simplified and illustrated more clearly. At the end of this process, the data-containing codes had been fully defined as themes and subthemes with working titles. The final narrative account was then produced (Chapter five) in which the themes and subthemes were organised in a logical, non-repetitive report, with the findings fully supported by embedded participant quotes encapsulating the essence of each point made.

### 3.8.6 Format of the narrative account

The narrative account (Chapter five) was divided into sub-sections relating to the main themes and subthemes that developed and emerged throughout the analysis

of the interviews as described above. Verbatim quotations from the range of participants were provided to illustrate key findings, preceded by a participant identification number. Each quote presented was also preceded by a note of the primary occupation of the participant (e.g. Surgeon or Methodologist) and whether the participant had a role on a funding panel, to clarify the context of each viewpoint.

Ellipses in square brackets ([...]) were used to indicate any omitted words or phrases (for example where two relevant quotes about a topic from the same participant were said at different time points during the interview, or to keep the data anonymised). Words written within square brackets e.g. [xxx] are not direct quotes but used to clarify the context of the quote. Indication of omission was not deemed necessary for: 'Umms' and 'errs'; frequent use of terms such as 'you know', 'sort of', 'kind of'; 'ok'; 'I guess' or 'I mean'; and immediate repetitions of words. These words were always omitted unless their inclusion adds to the meaning of the quote. A sequence of dots (…) not within brackets, was used to indicate a minor pause in conversation.

# Phase III: A synthesis and interpretation of quantitative and qualitative research findings

## 3.9 Rationale for Phase III of this thesis

Mixed methods research describes the use of both quantitative and qualitative methods and the synthesis and interpretation of data from these different sources to answer a research question [179]. Justification for using a mixed methods approach in this thesis and a detailed description of the precise methods used to synthesise the data sources from phases one and two of this work, are described and discussed below.

### 3.9.1 Rationale for using mixed methods research generally

As a now accepted, stand-alone research method, much has been written about the application of various mixed methods research designs [186-188]. In terms of Health Services Research, a mixed methods approach has gained popularity as an appreciation of the impact of the psychosocial elements of health and human nature on medical care has grown [148]. A mixed methods approach should not, however, be applied indiscriminately or without explicitly justifying the rationale for doing so [186].

#### *Conceptual frameworks for using mixed methods*

Various conceptual frameworks have been developed [186, 189, 190] justifying the use of mixed methods research and these overlap considerably. In addition, several mixed methods texts are published which comprehensively discuss the issues around the rationale for mixed-methods research [186, 187, 191] and these are summarised by Hanson *et al* in their 2005 paper [168]. All these works [168, 186, 189, 190] can be broadly summarised into four fundamental justifications for using mixed methods research:

**Reason 1** To generate hypotheses or allow the evolution of hypotheses using one method to generate theories, and another method to test them;

**Reason 2** To identify measurable variables or constructs which may subsequently be measured through the use of existing tools or instigate the development of new measuring instruments;

**Reason 3** To increase the validity of results by using the data and results from one method (quantitative or qualitative) to identify populations to study by the other, therefore expanding on, verifying and/or adding credibility to the results and;

**Reason 4** To overall enhance the breadth and comprehensive understanding of research problems and results by converging quantitative and qualitative data and in particular, allowing the needs or voice of marginalised or under-represented groups to be better considered and understood.

### 3.9.2 Rationale for the mixed methods approach used in this thesis

Three of the four reasons justifying the use of mixed methods research, described above, apply to the approach taken for this thesis and these three reasons are identified and then described below. For clarity, each of the three reasons for using a mixed methods approach is first repeated in italics, with confirmation in brackets ([Xxx]) of which part of the work in this thesis relates to that reason.

> **Reason 2**: *To identify measurable constructs* [Surgical PFS design and misunderstanding and/or challenges faced identified in Phase I] *which may be subsequently measured through the use of existing tools'* [Phase II PEPSTAR qualitative interview study to understand why problems exist].

Phase one of this research was a quantitative targeted review and systematic analysis of NIHR funded PFS protocols to identify the scope and detail of what PFS were being designed and funded in surgery, and to begin to differentiate some of the areas in which PFS in surgery are not currently optimised. The issues identified in phase one, were explored and expanded on in phase two to optimally understand why challenges and barriers exist to completing PFS in

surgery. Using two different methods, therefore allowed further exploration in phase two, of the problems identified in phase one.

> **Reason 3:** *To increase the validity of the results by using data and results from one method* [Phase I findings] *to identify populations to study by the other* [Phase II participants and topic guide for the PEPSTAR qualitative interviews].

The findings from the quantitative part of this thesis (Phase one) informed the qualitative component (Phase two), by identifying some of the people involved in surgical PFS in the UK whom would be usefully included as research participants, and also by contributing to and evolving the semi-structured interview topic guide. Again, the use of two different methods in this thesis increased the validity of the findings overall, as the quantitative findings in phase one informed the qualitative data collection in phase two.

> **Reason 4:** *To overall enhance the breadth and comprehensive understanding of research problems and results by converging* [Phase III] *quantitative* [Phase I] *and qualitative data* [Phase II] *and in particular, allowing the needs or voice of marginalised or under-represented groups to be better considered and understood* [Surgeons and PFS in surgery].

This is the key reason justifying the use of a mixed methods approach in this thesis. Phase three of this thesis is the synthesis and interpretation of the quantitative (Phase one) and qualitative (Phase two) components of this work to overall enhance understanding of the potential for PFS to optimise trials in surgery.

The next section will describe the key methodological approaches underpinning mixed methods research generally. This will be followed by description and justification for the methods used in this thesis, to synthesise and interpret the quantitative and qualitative data to produce recommendations for research practice from this work.

## 3.10 Methodological approaches for using mixed methods

Three basic steps have been defined when designing a mixed methods study by Hanson *et al.* [168] as: 1) Whether and which philosophical paradigm will be used to underpin the methodological basis of the study; 2) Deciding how the data collection will be sequenced (concurrent or sequential) and prioritised (equal, or either method prioritised) and; 3) How and when data analysis and integration will occur. Each of these three steps will now be discussed.

### *Step one: Philosophical paradigm*

Because the philosophical approaches of quantitative and qualitative research are traditionally opposed (see section 3.6 earlier), there has been much debate about the philosophical assumptions that are required for mixed methods research. Some argue that using competing research paradigms results in contradictory and challenged ideas and theories, which can positively impact on the research [168]. Others argue that pragmatism is the best research paradigm for mixed methods research, meaning both methods can be used in a single study, and that the research question is of primary importance, rather than the method, or philosophical research paradigm that underlies it [187]. As described earlier (see sections 3.6), the philosophical basis for research, may also be described as the 'theoretical lens' that underlies the researcher's choices and assumptions, and through which the data collection, analysis and interpretation is viewed [168]. Most importantly it is key that the researcher identifies if a philosophical paradigm is subscribed to, and clearly describes and justifies the choice (or not) of philosophical paradigm, and how it may affect the interpretation of the results.

### *Step two: Data collection*

Data collection in a mixed methods project can occur in different ways. The implementation of data collection of different parts of the research may be concurrent (collected at the same time) or sequential (collected at different times). Priority may be given to one form of data over another (unequal) or the data sets may have equal emphasis on the findings. Again, the important part of this step

is that the process used is described, though several authors have produced typologies to formally classify mixed methods research designs [186, 188]

### *Step three: Data analysis and integration*

Data analysis and integration may be completed separately, or using transformation (where for example, qualitative data is transformed into a quantitative format such as counts or ratings), or they may be connected, where the analysis of one type of data, informs the analysis of the other. When data analysis and interpretation will be performed must also be defined and may occur during the data collection, analysis or interpretation stages.

Three techniques for integrating data from mixed methods studies have been identified and described, with particular reference to Health Services Research, by O'Cathain *et al.* [192]. These are termed: 1) Following a thread; 2) Developing a mixed methods matrix; and 3) Triangulation. The latter two methods occur at the analysis stage, whereas the triangulation method would be followed at the interpretation stage. Each of these techniques are described in turn.

The technique of following a thread involves, first performing initial analysis of each data set, then identifying key themes that need further exploration and 'following' these themes from one data set to the other. O'Cathain offers an example of using this technique in Health Services Research (primary care) [193], but describes following a thread as a technique that is infrequently described in the published literature [192].

Developing a mixed methods matrix, is a data integration method which can be used when there is data from two or more sources available for the same populations (e.g. individuals or settings). A physical matrix can then be created that displays, for example, quantitative and qualitative data available for each population in rows, and columns illustrating the different data categories collected. The mixed methods matrix then allows direct comparison of the different data sets collected across the same populations, and themes and/or inconsistencies identified, which allows for further exploration [192, 194].

The term triangulation can have different meanings in research methodology. It can be used to describe degrees of corroboration between two sets of findings, therefore reflecting the validity of the findings [150]. However, it can also mean the process of studying a problem using different methods to gain a more comprehensive answer, and this is the meaning most often used in mixed methods studies [192]. Several techniques for triangulating findings have been described, but in summary they involve assessing to see where findings from each method might converge fully or partially (convergence), offer complementary information (complementarity) or contradict (discrepancy or dissonance) [192]. A fourth outcome of 'silence' has also been described by some [195], albeit in relation to two types of qualitative data, which fits within the 'contradictory' category, and describes a situation where findings are seen in one set of results but not another. Farmer *et al*., describe silences as potentially expected due to differences in the purpose and nature of the data sets [195]. This can be extrapolated to mixed methods studies, as one purpose of these studies is to strengthen the findings by using different methods to explore different elements of the same problem. However, O'Cathain *et al*. emphasize that finding unexpected silences in one data set might help to increase understanding of the problem and lead to further avenues of research [192].

## 3.11 Methodological approach for a synthesis and interpretation of quantitative and qualitative data in this thesis

The methods used to synthesise the data from phases one and two of the research conducted in this thesis will now be described in light of the theory underpinning different mixed methods approaches just described. Each of the three steps necessary when designing a mixed methods study will now be described in relation to the mixed methods approach used in this thesis.

### *Step one: Philosophical paradigm*

Earlier in this chapter (see section 3.8.2), the rationale for using a critical realist approach for analysing the qualitative data in phase two, was presented and

justified. A critical realist theoretical lens was also applied during phase three of the research to synthesise the findings from phases one and two. This critical realist theoretical lens was considered appropriate because as with analysing the qualitative data from the PEPSTAR study, the available definitions and methodological work in the area of PFS, were perceived as the imperfect truth (or objective reality), against which all data (both quantitative and qualitative) were compared.

### *Step two: Data collection*

The published literature, current guidance documents relevant to designing and conducting PFS [9, 13, 102, 144], and the methodological conceptual work on PFS design and conduct described by others [71, 72, 74, 75, 79, 89, 91-93, 95, 96], was extensively described, discussed and illustrated in Chapter two. Some of the evolution of thinking and publications in the area of PFS design and conduct, occurred during the process of this thesis (October 2015 – October 2019). The staggered availability of these information sources was, therefore, taken into consideration during the synthesis and interpretation of findings. All three phases of the research in this thesis were, therefore, conducted in light of the background literature and methodological knowledge around PFS, using a critical realist approach, as described in step one above (See Figure 3c).

The quantitative targeted review and systematic analysis of the NIHR funded PFS protocols, was performed and analysed before the PEPSTAR qualitative interviews (as described in section 3.4 of this chapter). The quantitative data analysis, therefore, informed the qualitative data collection and analysis. This is a sequential data collection model, with equal priority of components. Data analyses were, therefore, connected, as the quantitative findings, impacted on what data were collected in the qualitative work, and how these qualitative data were analysed and interpreted.

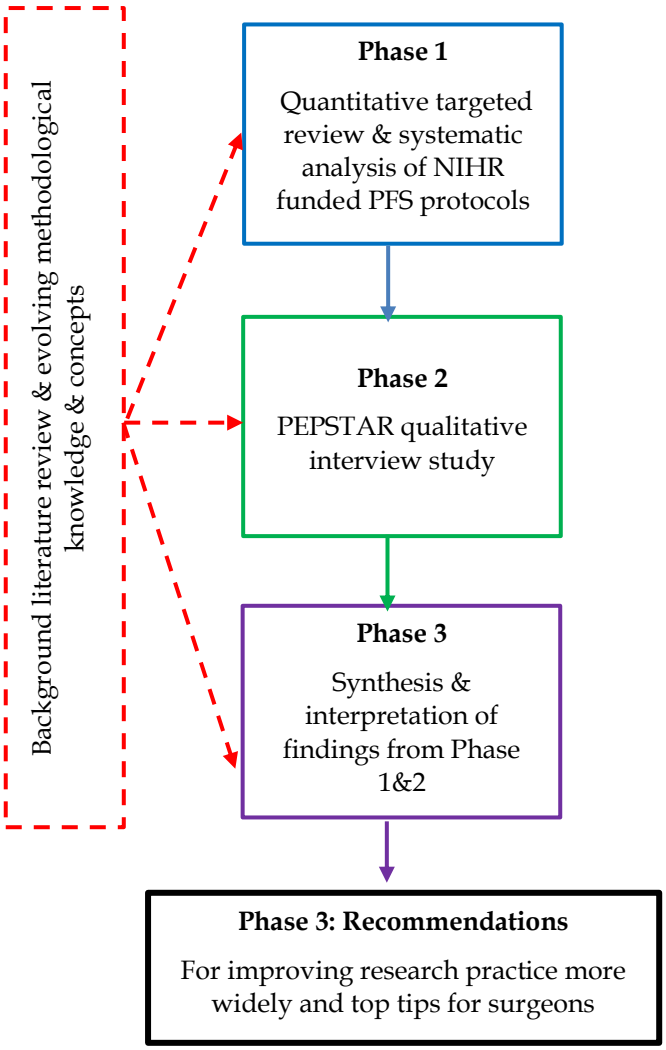



***Figure 3c*** *Process of data collection and analysis in this thesis*

### *Step three: Data analysis and integration*

Whilst the quantitative and qualitative data were analysed separately initially (see Results; Chapters four and five), the process of interpretation was one of data triangulation to synthesise the findings from phases one and two subsequently (see Results; Chapter six), in light of the background literature review and methodological work in this area (described in detail in Chapter two). This process is illustrated in Figure 3c above.

The overall aim of the work in this thesis, was to develop a detailed understanding of the potential for pilot and feasibility work to optimise surgical trials and provide clear recommendations for surgeons and study teams for how

to improve research practice. As described at the end of chapter two, what is currently lacking is guidance tailored specifically to optimally design and undertake PFS. The fourth objective of this thesis (see section 3.1) was, therefore:

> **Objective 4** To develop recommendations from this work for all key stakeholders and, *specifically, accessible and practical top tips for surgeons* to optimise the design and conduct of pilot and feasibility studies to inform main trials in surgery.

Phase three of this work, therefore, results in recommendations for improving research practice more widely amongst all key stakeholder groups. In addition, a brief practical 'Top Tips' tool to operationalise these recommendations for use by surgeons designing and conducting PFS in surgery was produced (as illustrated in Figure 3c above).

The methodological approach employed in phase three of this work, utilised a process of identifying the key findings from each of the data sources, and representing these diagrammatically. Through the triangulation approach, it was sought to define current guidance and theoretical understanding of PFS, define what the potential issues for surgical PFS are, and explore why these barriers exist and what solutions may be needed for optimisation of future PFS in surgery, the learnings from which would inform the development of the recommendations and 'Top Tips' tool for surgeons. A process of seeking evidence of any convergence, complementarity and dissonance between the data sets was undertaken to produce a detailed written narrative account of recommendations from this work in Chapter six.

## 3.12 Reflexive considerations for this thesis

In practice, the process of data interpretation for phase three of this work began, once the qualitative data collection started in phase two, as it is impossible to separate oneself from one's own knowledge [180]. Reflexive considerations, which had influence over both phases two and three of the work in this thesis, are discussed below.

### *Reflexive considerations*

The author of this thesis (KF) is a female general surgical trainee, with a subspecialty interest in oncoplastic breast surgery. KF completed a BSC in physiology in 2000, then qualified from medical school with an MBChB in 2004 and worked as an NHS doctor for 11 years before starting this PhD. Whilst KF had been involved in a variety of small research projects prior to starting this PhD, including a systematic review of the resection of breast cancer liver metastases and publishing work on surgical simulation, she had no previous involvement in or experience of qualitative research. To address this, KF attended several structured qualitative research courses at the University of Bristol in the early stages of her PhD Studentship (Introduction to Qualitative Research Methods, School of Social and Community Medicine, May 2016; NVivo Training, February 2017), and was supervised by two (of four) researchers with extensive experience of using mixed methods and qualitative methods in Health Services Research.

As a surgical trainee, KF may have been perceived differently by clinical and non-clinical participants. Her preconceptions of the knowledge and understanding of the surgical community may have been influenced by her previous NHS work and experiences. In addition, the thesis was supervised within the Centre of Surgical Research at the University of Bristol, which has a longstanding role in using qualitative methods to understand the role of surgeons in surgical trial recruitment, and in the education of surgeons in trials methodology. Her supervisors may, therefore, also have had preconceptions about the relative depths of knowledge and understanding of PFS amongst surgeons and methodologists.

As the author of this thesis is a surgical trainee, when reflective influences are considered, this work was perhaps inevitably done with surgeons and research practice amongst surgeons in mind. It would be fair to surmise that as a surgical trainee, KF had witnessed first-hand, many of the difficulties of performing surgical research without formal methodological training, and the consequences

of not working with a robust evidence base underpinning surgical practice. These experiences certainly influenced the reasons for undertaking of this PhD.

Reflexivity was therefore employed throughout this work, where KF recognised: a) The impact of both her knowledge and experiences of working as an NHS surgical trainee; b) How the quantitative results of the systematic analysis of NIHR funded PFS protocols exploring the use of PFS in surgery impacted on and related to the PEPSTAR interviews, c) The impact of working within the Centre for Surgical Research at the University of Bristol, which has strong methodological expertise in surgical trials methodology and the training of surgeons in trial design and recruitment, and developing surgical trials for research questions that have previously been considered impossible and; d) The evolving methodological literature and guidance documents relevant to surgical PFS, such as the IDEAL guidance. These considerations were kept in mind throughout the process of data collection, analysis and interpretation.

The next chapter (Chapter four) will describe the results of the targeted review and systematic analysis of NIHR funded PFS protocols. This will be followed by the results of the PEPSTAR qualitative interview study (Chapter five), and finally the synthesis and interpretation of the findings from the first two phases to produce recommendations for improving research practice from this work (Chapter six).

# CHAPTER 4

*Phase I*

*A targeted review and systematic analysis of surgical pilot and feasibility work funded by the National Institute for Health Research*

## 4.1 Introduction

In order to consider how to optimise future PFS for surgical trials, it is important to understand, what and how PFS have been designed and conducted in recent years. This chapter describes the results from phase one of this work, which was a comprehensive quantitative targeted review and systematic analysis of NIHR funded PFS performed over a recent 10-year period. The results of this work illustrate and describe the quantity, type and quality of PFS being funded and performed nationally, within a surgical context.

## 4.2 Results of a targeted review and systematic analysis of surgical pilot and feasibility studies funded by the National Institute for Health Research

***Screening***

Over the 10-year period (2005 to 2015), 1341 funded studies were identified from the RfPB (n= 638, 48%) and HTA (n=703, 52%) databases (see Figure 4a). Of these, 1265 (93.7%) studies (RfPB n=610/638, 95.6%; HTA n=655/703, 93.2%)

were excluded primarily because the study interventions were not surgical (n=1115, 88.1%). Other reasons for exclusion included: Surgical intervention but not a PFS (n=65, 5.1%), surgical intervention being studied was a co-intervention (n=28, 2.2%), surgical intervention being studies is a co-intervention and also not a PFS (n=57, 4.5%).

Of all clinical research funded by the NIHR HTA and RfPB programmes in the last 10 years, 16.7% (225/1341) had some association with surgery (n=75 surgical studies assessed for eligibility [76 studies – 1 study with no surgical intervention] + n= 150 surgical studies excluded for reasons 2/3/3a: 225/1341]. Of these studies, surgery was a co-intervention in more than a third (studies excluded for reasons 3 + 3a = 85/225, 37.8%), with only 10.4% of the funded research in this cohort examining surgery as the main intervention in one or more of the study groups (n=75 surgical studies assessed for eligibility + n=65 surgical studies excluded for reason 2: 140/1341) (data all from Figure 4a). Just a quarter of this work was pilot/feasibility work for surgical interventions (35/140, 25%).

### *Eligibility*

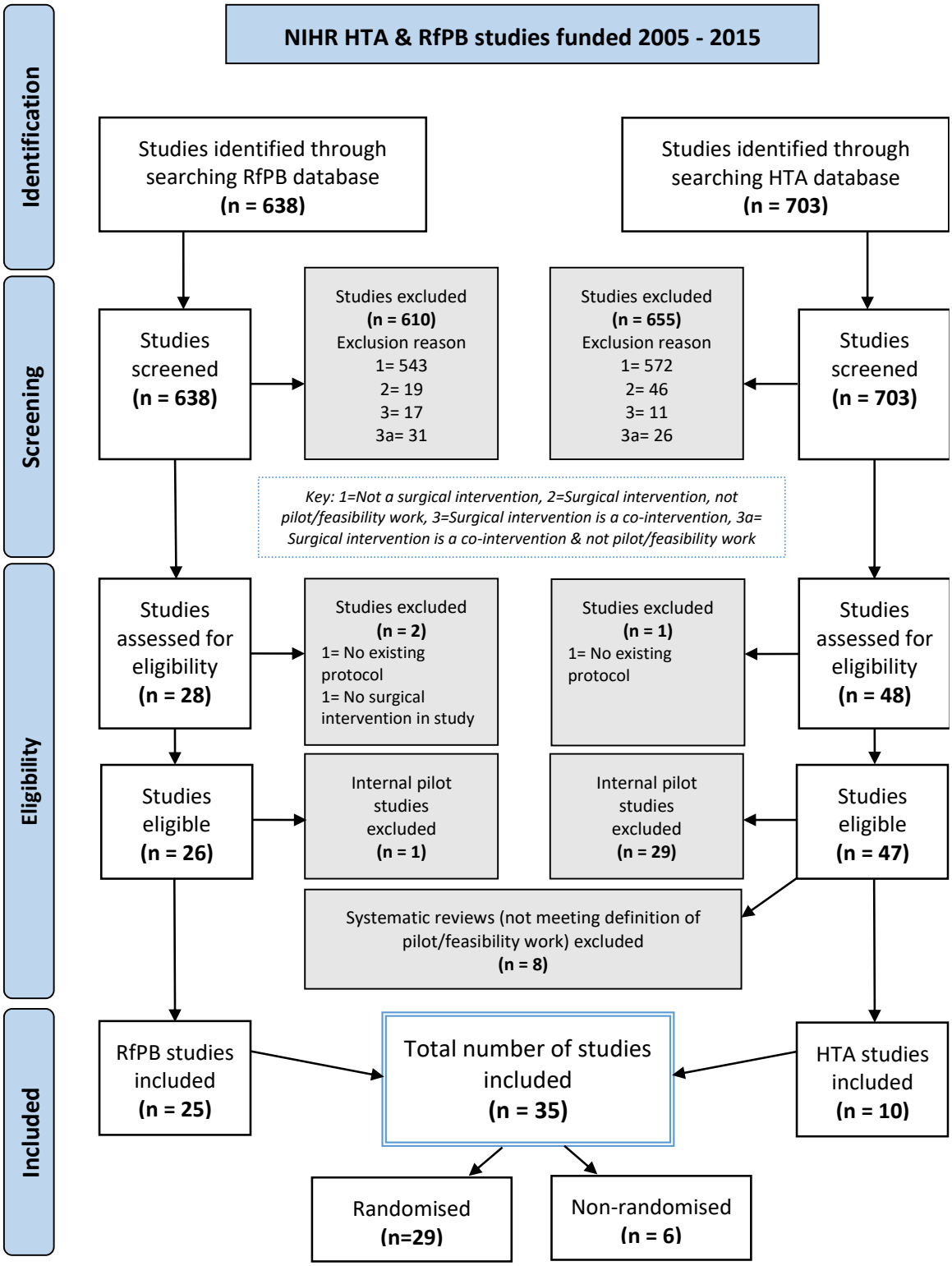
The remaining 76 studies (RfPB 28, 36.8%; HTA 48, 63.2%) were assessed for eligibility. Of these, 30 (RfPB n=1; HTA n=29) were internal pilot studies; 8 (RfPB n=0, HTA n=8) were systematic reviews, 2 (RfPB n=1; HTA n=1) had no available study protocol and one did not include a surgical intervention. These studies were excluded. In total, 35 study protocols (RfPB n=25, HTA n=10) were included in the analysis. Details of these studies are available in Appendix II.

### *Agreement on included studies*

Of the 1341 studies screened in duplicate, agreement was reached on inclusion or exclusion for 1283 (95.7%) studies at the first attempt. The remaining n=58 (4.3%) studies where there was disagreement were discussed (Fairhurst/Rowlands) and agreement reached in all cases. For example, in 25/58 (43.1%), the studies were excluded by the second screener (Rowlands) as being an internal pilot or systematic review and not meeting the definition of pilot/feasibility work; these studies were later excluded anyway, as per the PRISMA diagram in Figure 4a.

Difficulties determining the type of surgical intervention and whether the study was pilot/feasibility work were the main reasons for discordance in the remaining 33/58 (56.9%) studies where disagreement regarding inclusion/exclusion occurred.

***Figure 4a*** *PRISMA flow diagram of studies identified and selected for a systematic analysis of NIHR HTA and RfPB pilot and feasibility studies of surgical interventions.*



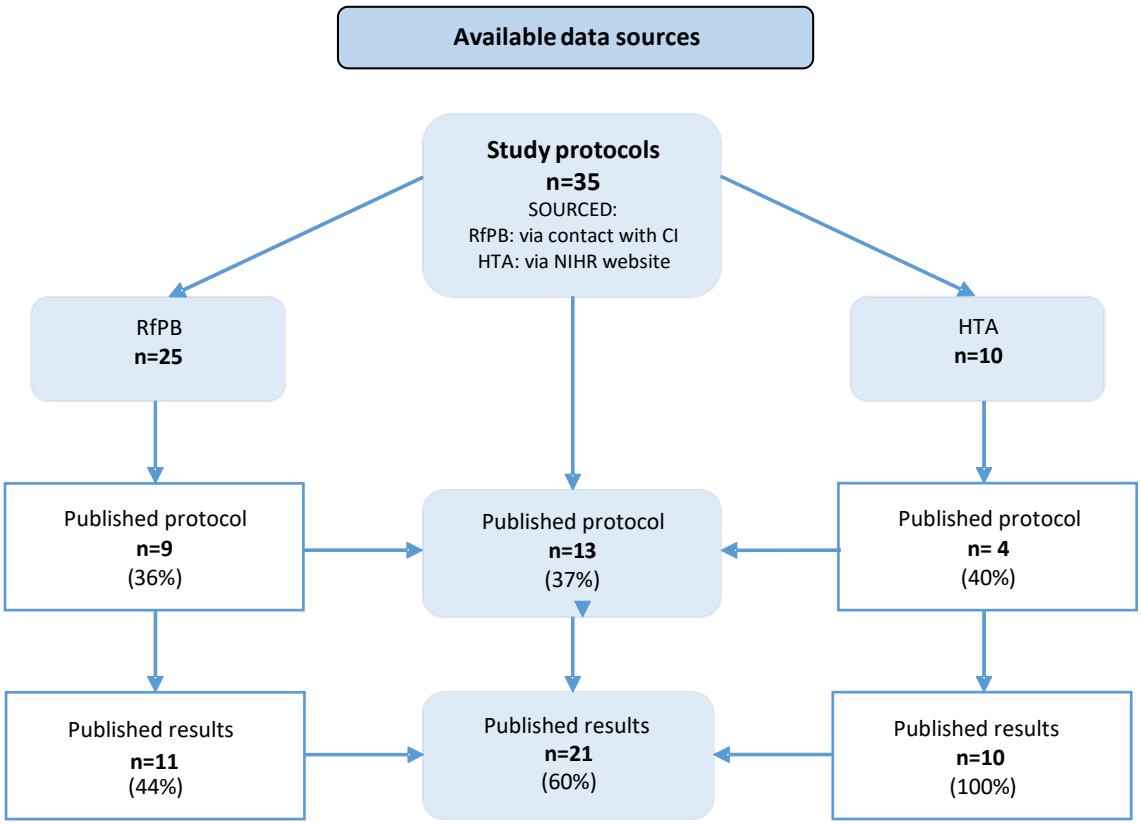

**KEY:**

**NIHR:** National Institute of Health Research
**HTA:** Health Technology Assessment
**RfPB:** Research for Patient Benefit

## 4.3 Data sources available for included studies

All 35 included study protocols were available either through contact with the chief investigator or via the HTA website as detailed in the methods (section 3.4.2). Additional data sources were available for 26/35 (71.4%) studies; this included a published protocol paper for 13/35 (37%) studies and a paper (or abstract n=1 only) reporting the study findings for 21/35 (60%) studies (Figure 4b). Studies funded by the HTA were more likely to publish the results (HTA 10/10, 100% vs RfPB 11/25, 44%), than those funded by the RfPB, but the study protocols had a similar publication rate between funding streams (published protocol: HTA 4/10, 40% vs RfPB 9/25, 36%)

***Figure 4b*** *Data sources available for the PFS included in the analysis.*




**KEY:**

**CI:** Chief Investigator
**HTA:** Health Technology Assessment
**NIHR:** National Institute of Health Research
**RfPB:** Research for Patient Benefit
**Published results:**
HTA = HTA report +/- other paper(s) in peer reviewed journal
RfPB = published in peer reviewed journal (NB n=1 abstract only)

## 4.4 Characteristics of included studies

### *Randomised pilot/feasibility studies*

Most (29/35, 83%) of the 35 included studies used a randomised design (Figure 4a), the design characteristics of which are detailed in Table 4.1. These 29 randomised PFS covered a breadth of surgical specialties, with the majority (22/29, 75.9%) planned as multi-centre and a quarter (7/29, 24.1%) as single centre studies. The proposed median sample size of these studies was 60 (range 30-200), with most planned in adult populations (24/29, 82.8%) of both sexes (24/29, 82.8%) and predominantly in the UK only (27/29, 93.1%). A minority of studies also had collaborative centres in Europe (2/29, 6.9%, both RfPB). The median number of personnel proposed in the trial team was nine, with a range of 0-24 (some studies did not state the proposed study personnel).

Operative therapeutic interventions (19/29, 65.5%) were investigated most commonly, followed by endoscopic therapeutic interventions (6/29, 20.7%). The most common comparator was an operative therapeutic intervention (9/32, 28.1%), followed by a pharmacological comparator (5/32, 15.6%). Most studies had two study groups (25/29, 86.2%) and a single comparator (26/29, 89.7%). While primarily randomised in design, over half (15/29, 52%) of studies also included other types of pre-trial work, such as qualitative interviews (with staff and participants), surveys/questionnaires (with surgeons and/or participating centres), and economic modelling.

### *Non-randomised pilot/feasibility studies*

Like the randomised PFS, the 6/35 (17%) non-randomised studies were conducted in a variety of surgical specialties (Table 4.2). Their design varied considerably, including a national audit (n=1), a non-randomised cohort study (n=1), a systematic review (n=1) and questionnaire surveys and/or qualitative work (for example, interviews/focus groups) to explore stakeholders' opinions (n=9). Unlike the randomised studies, most were planned as single centre (n=4/6, 67%), with fewer multi-centre studies (n=2/6, 33%). Most were planned in the UK only (n=5/6, 83%) and in adults (n=4/6, 67%) of both sexes (n=4/6,

67%). The study teams of the non-randomised studies comprised fewer personnel than the randomised studies, with a median of six (range of 0-25). Because all six studies were, in comparison to the randomised studies, earlier on in the research process, only two focused on a specific intervention and none specified any comparator intervention(s).

***Table 4.1*** *Summary of the design characteristics of the 29 randomised pilot/feasibility studies included in the analysis.*

| Study design characteristic | | RfPB (n=22) | HTA (n=7) | TOTAL (n=29) |
|---|---|---|---|---|
| **Surgical specialty of study, n (%)** | | | | |
| Gastrointestinal | | 8 (36.4) | 1 (14.3) | 9 (31.0) |
| Urology | | 3 (13.6) | 2 (28.6) | 5 (17.2) |
| Cardiothoracic | | 3 (13.6) | 0 (0) | 3 (10.3) |
| Orthopaedic | | 2 (9.1) | 2 (28.6) | 4 (13.8) |
| Obstetrics/gynaecology | | 2 (9.1) | 0 (0) | 2 (6.9) |
| Maxillofacial/ENT | | 2 (9.1) | 1 (14.3) | 3 (10.3) |
| Plastics | | 1 (4.5) | 0 (0) | 1 (3.4) |
| Paediatrics | | 1 (4.5) | 1 (14.3 | 2 (6.9) |
| **Number of centres, n (%)** | | | | |
| 1 | | 7 (31.8) | 0 (0) | 7 (24.1) |
| 2-20 | | 14 (63.6) | 7 (100) | 21 (72.4) |
| >20 | | 1 (4.6) | 0 (0) | 1 (3.5) |
| Median (range) | | 3 (1-23) | 4 (2-10) | 3 (1-23) |
| **Proposed number of participants in study** | | | | |
| Median (range) | | 50 (30-200) | 70 (60-144) | 60 (30-200) |
| **Patient population characteristics, n (%)** | | | | |
| **Age:** Adults/Children/Both | | 19 (86.4)/2 (9.1)/1 (4.5) | 5 (71.4)/1 (14.3)/1 (14.3) | 24 (82.8)/3 (10.3)/2 (6.9) |
| **Sex:** Male/Female/Both | | 1 (4.5)/3 (13.6)/18 (81.9) | 1 (14.3)/0 (0)/6 (85.7) | 2 (6.9)/3 (10.3)/24 (82.8) |
| **Country:** UK/Europe/Worldwide | | 20 (90.9)/2 (9.1)/0 (0) | 7 (100)/0 (0)/0 (0) | 27 (93.1)/2 (6.9)/0 (0) |
| **Number of personnel in trial team** | | | | |
| Median (range) | | 6.5 (4-19) | 9.5 (0-24) | 9 (0-24) |
| **Type of surgical intervention under investigation/exploration, n** | | | | |
| Operative | therapeutic | 14 | 5 | 19 |
| | diagnostic | 0 | 0 | 0 |
| | adjunctive | 1 | 0 | 1 |
| Radiological | therapeutic | 1 | 1 | 2 |
| | diagnostic | 0 | 0 | 0 |
| | adjunctive | 0 | 1 | 1 |
| Endoscopic | therapeutic | 5 | 1 | 6 |
| | diagnostic | 1 | 1 | 2 |
| | adjunctive | 0 | 0 | 0 |
| NB HTA cohort adds up to n=9 as one intervention is both endoscopic & radiological adjunctive and another is both endoscopic & radiological therapeutic. | | | | |
| **Total number of study groups, n (%)** | | | | |
| 2 groups | | 20 (90.9) | 5 (71.4) | 25 (86.2) |
| 3 groups | | 2 (9.1) | 2 (28.6) | 4 (13.8) |
| **Number of comparator groups, n (%)** | | | | |
| 1 group | | 21 (95.5) | 6 (85.7) | 26 (89.7) |
| 2 groups | | 1 (4.5) | 1 (14.3) | 3 (10.3) |
| **Type of comparator intervention, n (%)** | | | | |
| | | **NB for n=23 comparators** | **NB for n=8 comparators** | **NB for n=32 comparators** |
| Pharmacological | | 5 | 0 | 5 |
| Usual/standard care | | 3 | 0 | 3 |
| Other invasive procedure | | 2 | 2 | 4 |
| Expectant management | | 1 | 2 | 3 |
| Best medical therapy | | 1 | 1 | 2 |
| Operative | therapeutic | 6 | 2 | 9 |
| | diagnostic | 0 | 1 | 1 |
| | adjunctive | 1 | 0 | 1 |
| Radiological | therapeutic | 1 | 0 | 1 |
| | diagnostic | 0 | 0 | 0 |
| | adjunctive | 0 | 0 | 0 |
| Endoscopic | therapeutic | 2 | 0 | 2 |
| | diagnostic | 1 | 0 | 1 |
| | adjunctive | 0 | 0 | 0 |
| **Types of non-randomised pre-trials work for n=15 studies, n** | | | | |
| | | **NB for n= 9 studies** | **NB for n= 6 studies*** | **NB for n=15 studies*** |
| Qualitative interviews | | 8 | 6 | 14 |
| Participant/researcher survey | | 1 | 1 | 2 |
| Economic modelling | | - | 1 | 1 |
| * Some studies planned more than one type of non-randomised work | | | | |

***Table 4.2:*** *Summary of the design characteristics of the six non-randomised pilot/feasibility studies included in the analysis.*

| Study design characteristic | HTA (n=3) | RfPB (n=3) | TOTAL (n=6) |
|---|---|---|---|
| **Surgical specialty of study, n** | | | |
| Gastrointestinal | 2 | - | 2 |
| Obstetrics/gynaecology | - | 1 | 1 |
| Maxillofacial/ENT | - | 1 | 1 |
| Paediatrics | - | 1 | 1 |
| Breast | 1 | - | 1 |
| **Number of centres, n** | | | |
| 1 | 3 | 1 | 4 |
| 2-20 | - | 1 | 1 |
| >20 | - | 1 | 1 |
| Median (range) | 1 (1) | 2 (1-65) | 1 (1-65) |
| **Number of personnel in trial team** | | | |
| Median (range) | 1 (0-12) | 7 (2-25) | 5.5 (0-25) |
| **Patient population characteristics, n** | | | |
| **Age:** Adults/Children/Both | 2/1/0 | 2/0/1 | 4/1/1 |
| **Sex:** Male/Female/Both | 0/1/2 | 0/1/2 | 0/2/4 |
| **Country**: UK/Europe/Worldwide | 2/0/1 | 3/0/0 | 5/0/1 |
| **Types of non-randomised pre-trial work for each study, n*** | | | |
| Qualitative interviews | 2 | 2 | 4 |
| Participant/researcher survey | 2 | 3 | 5 |
| Systematic review | 1 | - | 1 |
| National audit | - | 1 | 1 |
| Cohort study | - | 1 | 1 |
| * Most studies planned more than one type of non-randomised work. | | | |

# 4.5 Rationale for conducting surgical pilot and feasibility studies

The rationale described in the protocols, for conducting the 35 PFS included in the systematic analysis are summarised in Table 4.3. The work identified 43 reasons for conducting PFS, and these reasons were grouped into five key domains (or themes) of: 1) main trial design; 2) logistics; 3) recruitment; 4) intervention, and 5) outcomes, as described in chapter 3, section 3.4.4.

### *Uncertainty surrounding main trial design*

The overarching aim to determine whether a main trial was possible or necessary (Total: 27/35, 77.1%; RfPB: 17/25, 68%; HTA: 10/10, 100%) was the second most commonly cited as a reason for performing PFS, which is perhaps unsurprising given that this is generally the overall purpose of PFS. Around half of all studies also considered issues regarding the sample size for the main trial (Total: 19/35,

54.3%; RfPB: 17/25, 68 %; HTA: 2/10, 20%) and costs/funding for the main trial (Total: 16/35, 47.7%; RfPB: 14/25, 56%; HTA: 2/10, 20%). Despite being termed pilot/feasibility work, one third of studies (n=11, 31.4%, all RfPB funded) aimed to collect data regarding the safety or effectiveness of an intervention to inform the main trial (Table 4.3). Of these, almost three quarters (8/11, 72.7%) specified plans for formal hypothesis testing by statistically comparing the intervention(s) and/or control groups to test effectiveness and/or safety.

### *Uncertainty surrounding logistics*

Uncertainties surrounding logistics were considered in two thirds of studies (Total: 23/35, 65.7%; RfPB: 16/25, 64%; HTA: 7/10, 70%) (Table 4.3). The most common logistical uncertainty given was to develop/test data collection forms/methods (Total: 19/35, 54.3%; RfPB: 16/25, 64%; HTA: 6/10, 60%), followed by to test the logistics of multicentre studies (Total: 6/35, 17.1%; RfPB: 5/25, 20%; HTA: 1/10, 10%), and to develop/test questionnaires/surveys (Total: 6/35, 17.1%; RfPB: 5/25, 20%; HTA: 1/10, 10%). Less commonly explored logistical uncertainties were to test response rates to questionnaires/surveys, to develop a research network as a resource for a future main trial and to assess the logistics of delivering an intervention as part of a trial in the NHS.

### *Uncertainty surrounding recruitment*

Addressing uncertainties around trial recruitment, was cited as the most common reason overall (Total: 32/35, 91.4%; RfPB: 23/25, 92%; HTA: 9/10, 90%) for undertaking PFS (Table 4.3). Specifically, two thirds of studies (Total: 22/35, 62.9%; RfPB: 9/25, 36%; HTA: 5/10, 50%) considered assessing the numbers/rates of recruitment and consent. Other commonly explored recruitment uncertainties included testing the acceptability of randomisation/trial design (Total: 17/35; 48.6%; RfPB: 12/25, 48%; HTA: 5/10, 50%), and determining the acceptability of the intervention to clinicians and patients (Total: 16/35; 45.7%; RfPB: 12/25, 48%; HTA: 4/10, 40%). More rarely explored issues around recruitment were to test/modify

inclusion/exclusion/eligibility criteria, and to estimate the expected prevalence or rate of incident cases in the population.

### *Uncertainty surrounding the intervention*

Only a quarter (10, 28.6%) of PFS sought to explore uncertainties around the surgical intervention itself, such as intervention development, stability, delivery or the surgical learning curve (RfPB: 6/25, 24%; HTA: 4/10, 40%) (Table 4.3). The ten studies that did plan to explore uncertainties around the intervention, most commonly considered developing and testing the implementation and delivery of the intervention (Total: 4/35, 11.4%; RfPB: 1/25, 4%; HTA: 3/10, 30%), followed by assessing and monitoring the development of an intervention and/or it's stability (Total: 3/35, 8.6%; RfPB: 2/25, 8%; HTA: 1/10, 10%). Less frequently explored uncertainties around the intervention included developing pathways and protocols for co-interventions, testing rates of crossover, training staff in delivery and assessment procedures.

Further details of the ten studies specifically examining aspects of the intervention are summarised in Table 4.4. Of these, six studies were comparing surgery with no surgery, three studies were comparing a novel/new surgical technique with a standard surgical procedure and one study was a non-randomised study examining novel/new surgical techniques.

Of the 25 studies not stating plans to explore uncertainties around the intervention, a third (n=9) were comparing surgery with no surgery, a quarter (n=6) were evaluating novel surgical interventions and one study aimed to compare surgery to a placebo and no surgery (three arm study) (Table 4.5).

When comparing the study type used in those studies stating examination of the intervention versus those that did not, the commonest study type for both groups was surgery versus no surgery, followed by novel/new surgical technique versus standard surgery (see Table 4.5). However, the studies not stating examination of the intervention were more likely to be surgery versus no surgery studies where both interventions were an established technique (4/25, 16% vs 0/10, 0%), or non randomised PFS (5/25, 20% vs 0/10, 0%) (See Table 4.5)

### *Uncertainty surrounding outcomes*

Less than half of studies considered uncertainties around outcomes (Total: 15/35, 42.9%; RfPB: 11/25, 44%; HTA: 4/10, 40%). Specifically, those studies stating exploration of outcomes planned to use PFS to select the most appropriate primary outcome measure (Total: 9/35, 25.7%; RfPB: 9/25, 36%; HTA: 0/10, 0%) and/or determine appropriate/important/suitability of outcome measures for patients/clinicians (Total: 7/35, 20%; RfPB: 3/25, 12%; HTA: 4/10, 40%).

***Table 4.3*** *Reasons cited in the study protocols for conducting PFS for the 35 included studies.*

| Area examined | | Rationale | Number of studies stating each rationale in the study protocol, n | | | Proportion of studies stating examination of each area in the study protocol, % (n) | | |
|---|---|---|---|---|---|---|---|---|
| | | | RfPB n=25 [NR n=3] | HTA n=10 [NR n=3] | TOTAL n=35 [NR n=6] | RfPB n=25 | HTA n=10 | TOTAL n=35 |
| Main trial design | Main trial possible +/- necessary | To examine and test whether a main trial is possible | 14 [2] | 8 [1] | 22 [3] | 68% (17) | 100% (10) | 77.1% (27) |
| | | To assess whether main trial is needed and/or produce a protocol | 3 [0] | 0 [2] | 3 [2] | | | |
| | | To test whether the protocol can be adhered to and modify it as necessary | 2 [0] | 2 [0] | 4 [0] | | | |
| | Sample size | To estimate the variability in outcomes to help determine a sample size for the main trial | 15 [1] | 2 [0] | 17 [1] | 68% (17) | 20% (2) | 54.3% (19) |
| | | To determine a sample size for the main trial | 3 [0] | 0 [0] | 3 [0] | | | |
| | Costs / funding | To assess/gather information on costs of performing the trial (direct and indirect) | 2 [0] | 0 [0] | 2 [0] | 56% (14) | 20% (2) | 45.7% (16) |
| | | To perform/prepare for a cost effectiveness analysis of the intervention(s) | 13 [1] | 2 [0] | 15 [1] | | | |
| | | To provide information/evidence to funders | 1 [0] | 0 [0] | 1 [0] | | | |
| | Safety and effectiveness data | Preliminary data on safety to inform a main trial | 2 [0] | 0 [0] | 2 [0] | 44% (11) | 0% (0) | 31.4% (11) |
| | | Information on adverse events | 4 [0] | 0 [0] | 4 [0] | | | |
| | | Planned formal hypothesis testing of safety outcomes * | 3 [1] | 0 [0] | 3 [1] | | | |
| | | Preliminary data on effectiveness to inform a main trial | 0 [0] | 0 [0] | 0 [0] | | | |
| | | Planned formal hypothesis testing of effectiveness outcomes * | 7 [1] | 0 [0] | 7 [1] | | | |
| Logistics | | To test the logistics of multicentre studies | 5 [0] | 1 [0] | 6 [0] | 64% (16) | 70% (7) | 65.7% (23) |
| | | To develop a research network as a resource for a future main trial | 1 [1] | 0 [0] | 1 [1] | | | |
| | | To develop/test patient information content/forms/methods of delivery | 1 [0] | 3 [0] | 4 [0] | | | |
| | | To develop/test data collection forms/methods | 13 [1] | 6 [1] | 19 [2] | | | |
| | | To develop/test questionnaires/surveys | 5 [0] | 1 [0] | 6 [0] | | | |
| | | To test response rates to questionnaires/surveys | 0 [0] | 0 [0] | 0 [0] | | | |
| | | To prepare/plan/assess monitoring procedures | 0 [0] | 1 [0] | 1 [0] | | | |
| | | To determine what resources are needed for a main trial (funding/staff) | 3 [0] | 0 [0] | 3 [0] | | | |
| | | To assess the logistics of delivering an intervention as part of a trial in the NHS | 1 [0] | 0 [0] | 1 [0] | | | |
| | | To test (novel) methods of blinding | 1 [0] | 1 [0] | 2 [0] | | | |
| | | To assess proposed data analysis techniques | 1 [0] | 1 [0] | 2 [0] | | | |
| | | To lean about the day-to-day running of a trial | 1 [0] | 0 [0] | 1 [0] | | | |
| Recruitment | | To test/modify inclusion/exclusion/eligibility criteria | 2 [0] | 0 [1] | 2 [1] | 92% (23) | 90% (9) | 91.4 % (32) |
| | | To estimate the expected prevalence or rate of incident cases in the population | 1 [1] | 1 [0] | 2 [1] | | | |
| | | To estimate the number to be screened and proportions of eligible patients | 9 [0] | 3 [0] | 12 [0] | | | |
| | | To assess numbers/rates of recruitment and consent | 17 [0] | 5 [0] | 22 [0] | | | |
| | | To test the randomisation procedure | 5 [0] | 3 [0] | 8 [0] | | | |
| | | To test the acceptability of randomisation/trial design | 12 [1] | 5 [2] | 17 [3] | | | |
| | | To determine the acceptability of the intervention to clinicians and patients | 12 [1] | 4 [2] | 16 [3] | | | |
| | | To assess rates of retention in the study | 11 [0] | 2 [0] | 13 [0] | | | |
| Intervention | | To assess and monitor the development of an intervention and/or it's stability | 2 [1] | 1 [0] | 3 [1] | 24% (6) | 40% (4) | 28.6% (10) |
| | | To develop and test the implementation and delivery of the intervention | 1 [0] | 3 [0] | 4 [0] | | | |
| | | To train staff in delivery and assessment procedures | 1 [0] | 0 [0] | 1 [0] | | | |
| | | To monitor the surgical learning curve | 2 [1] | 0 [0] | 2 [1] | | | |
| | | To test rates of crossover | 0 [0] | 1 [0] | 1 [0] | | | |
| | | To examine reasons for non-adherence/cross-over for the main trial | 2 [0] | 0 [0] | 2 [0] | | | |
| | | To develop pathways and protocols for co-interventions | 0 [0] | 0 [0] | 0 [0] | | | |
| Outcome | | To select the most appropriate primary outcome measure | 9 [1] | 0 [0] | 9 [1] | 44% (11) | 40% (4) | 42.9% (15) |
| | | To develop and test a new outcome measure | 0 [0] | 0 [0] | 0 [0] | | | |
| | | To determine appropriate/important/suitability of outcome measures for patients/clinicians | 3 [0] | 4 [2] | 7 [2] | | | |
| | | * Formal hypothesis testing to demonstrate the safety and/or effectiveness of an intervention is generally not recommended for PFS because of the underpowered sample size | NR = Non-randomised study | | | | | |

***Table 4.4*** *Details of the ten studies that examined elements of the intervention*

| Evaluation | INTERVENTION | Categorised type | COMPARATOR | Categorised type |
|---|---|---|---|---|
| Surgery vs no surgery | | | | |
| Management of childhood intermittent distance exotropia | Eye muscle surgery | Operative therapeutic | Active monitoring of children with intermittent exotropia | Expectant management |
| Management of perineal wound healing following vaginal delivery complicated by wound dehiscence | Re-suturing of dehisced perineal wounds | Operative therapeutic | Healing by expectancy (secondary intention) | Expectant management |
| Management of idiopathic overactive bladder in children | Single administration of an intravesical injection of Botox® into the bladder, in a dosage of 5 IU/Kg (maximum dose of 150 IU) under general anaesthesia | Endoscopic therapeutic | Tolterodine XL, 4 mg orally once daily, for the duration of the trial | Pharmacological |
| Management of femoro-acetabular impingement | Hip arthroscopy | Operative therapeutic | Personalised hip therapy: simple analgesia, physiotherapy, hip corticosteroid injection, postural adaptation, exercise, acupuncture, manual therapy techniques & lifestyle advice, gait modification. | Other invasive intervention |
| Nutritional management of patients undergoing chemoradiation for head and neck cancer | Pre-chemo/radiotherapy gastrostomy (may be either endoscopic or radiologically guided) | Radiological adjunctive/endoscopic adjunctive | Oral feeding unless/until unable to swallow, then NGT (invasive intervention) | Other invasive intervention |
| Management of peripheral stage 1 non-small cell lung cancer in patients considered higher risk of complication from surgical resection | Surgery - Thoracotomy or VATS (Video Assisted Thoracoscopic Surgery) | Operative therapeutic | Stereotactic ablative radiotherapy | Radiological therapeutic |
| Novel/new surgical technique vs standard surgical procedure | | | | |
| Techniques to surgically treat oesophageal carcinoma | 1. laparoscopic gastric mobilisation and right thoracotomy 2. totally minimally invasive surgery | Operative therapeutic | Open gastric mobilisation and right thoracotomy | Operative therapeutic |
| Techniques for portal vein control in liver resection for colorectal liver metastases | Portal vein clamping (during liver resection for colorectal metastases). The portal vein will be isolated and occluded with a clamp. | Operative adjunctive | Pringle manoeuvre (during liver resection for colorectal liver metastases to prevent bleeding. The entire hepatico-duodenal ligament containing both hepatic artery and portal vein, will be occluded using standard vascular clamp. | Operative adjunctive |
| Techniques for detection of oesophageal dysplasia during Barrett's surveillance | Acetic acid endoscopy of Barrett's oesophagus (targeted biopsy) | Endoscopic diagnostic | Mapping biopsy endoscopy of Barrett's oesophagus (standard care) | Radiological diagnostic (standard technique) |
| Techniques for implant-based breast reconstruction | Immediate implant-based breast reconstruction | Operative therapeutic | None | N/A |

***Table 4.5*** *Type of trial for studies that examined details of the surgical intervention and those that did not*

| Type of study | Studies examining the intervention (n=10) n (%) | Studies not examining the intervention (n=25) n (%) | Total (n=35) |
|---|---|---|---|
| Surgery vs no surgery | 6 (60) | 9 (36) | 15 (43) |
| Novel/new surgical technique vs surgery | 4 (40) | 6 (24) | 10 (29) |
| Non-randomised pilot/feasibility work | 0 (0) | 5 (20) | 5 (14) |
| Surgery vs surgery (both established techniques) | 0 (0) | 4 (16) | 4 (11) |
| Surgery vs placebo & no surgery (2 arms) | 0 (0) | 1 (4) | 1 (3) |

### *Comparison between randomised and non-randomised pilot/feasibility studies*

Some differences between the rationale for undertaking randomised and non-randomised PFS were observed, as detailed in Table 4.6, though recruitment issues and trial feasibility were the most common reasons for undertaking both randomised and non-randomised PFS. Non-randomised studies were more likely than randomised studies to be conducted to determine whether a main trial was possible or necessary (n=5, 83.3% vs n=22, 75.9%) and to inform outcome selection (n=3, 50% vs n=12, 41.4%). However, randomised studies were more likely to explore logistical issues (n=21, 72.4% vs n=2, 33.3%), collect data to inform sample size calculations (n=18, 62.1% vs n=1, 16.7%) and address issues around costs and funding (n=15, 51.7% vs n=1, 16.7%) than non-randomised PFS. When randomised studies are compared to non-randomised studies, uncertainties around the intervention were still the least explored area of all in both cohorts (non-randomised: n=1/6, 16.7% vs randomised: n=8/29, 27.6%) (See Table 4.6).

### *Consideration of the number of uncertainties addressed*

When the number of reasons for conducting each study was considered (Table 4.7), a minority of studies stated the intention to address more than ten areas of uncertainty (Total: 4/35, 11.4%; RfPB: 3/25, 12%; HTA 1/10, 10%). Almost a third planned to address fewer than five uncertainties (Total: 10/35, 28.6%, RfPB 5/25, 20%; HTA 5/10, 50%) (Table 4.7).

The number of uncertainties addressed in the PFS did not vary considerably between the HTA and RfPB cohorts (see Table 4.7). However, non-randomised studies tended to address far fewer uncertainties than randomised studies (n=5/6,

83.3% non-randomised addressed 1-5 uncertainties vs 5/29, 17.2% randomised) (Table 4.7).

***Table 4.6*** *Comparison of proportions of randomised and non-randomised studies stating examination of each area in the study protocol.*

| Area examined | | Proportion of studies stating examination of each area in the study protocol, n (%) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Randomised RfPB (n=22) | Non-randomised RfPB (n=3) | Randomised HTA (n=7) | Non-randomised HTA (n=3) | Randomised TOTAL (n=29) | Non-randomised TOTAL (n=6) |
| **Main trial design** | **Main trial possible +/- necessary** | 15 (68.2%) | 2 (66.7%) | 7 (100%) | 3 (100%) | 22 (75.9%) | 5 (83.3%) |
| | **Sample Size** | 16 (72.7%) | 1 (33%) | 2 (28.6%) | 0 (0%) | 18 (62.1%) | 1 (16.7%) |
| | **Costs/funding** | 13 (59.1%) | 1 (33.3%) | 2 (28.6%) | 0 (0%) | 15 (51.7%) | 1 (16.7%) |
| **Logistics** | | 15 (68.2%) | 1 (33.3%) | 6 (85.7%) | 1 (33%) | 21 (72.4%) | 2 (33.3%) |
| **Recruitment** | | 21 (95.5%) | 2 (66.7%) | 5 (71.4%) | 2 (66.7%) | 26 (89.7%) | 4 (66.7%) |
| **Intervention** | | 5 (22.7%) | 1 (33.3%) | 4 (57.1%) | 0 (0%) | 9 (31.0%) | 1 (16.7%) |
| | **Hypothesis testing** | 10 (45.5%) | 1 (33.3%) | 0 (0%) | 0 (0%) | 10 (34.5%) | 1 (16.7%) |
| **Outcome** | | 10 (45.5%) | 1 (33.3%) | 2 (28.6%) | 2 (66.7%) | 12 (41.4%) | 3 (50%) |

***Table 4.7*** *Number of uncertainties considered as stated in 35 study protocols.*

| **Number of uncertainties considered in each study** | **TOTAL** (n=35) | **RfPB** (n=25) | **HTA** (n=10) | **Randomised** (n=29) | **Non-randomised** (n=6) |
|---|---|---|---|---|---|
| **1-5** | 10 (28.6%) | 5 (20%) | 5 (50%) | 5 (17.2%) | 5 (83.3%) |
| **6-10** | 21 (60.0%) | 17 (68%) | 4 (40%) | 20 (69.0%) | 1 (16.7%) |
| **11-15** | 4 (11.4%) | 3 (12%) | 1 (10%) | 4 (13.8%) | 0 |

### *Consideration of the PFS study outcomes*

In order to assess the true value of PFS, it would be helpful to evaluate the outcomes of if and/or how PFS informed the design and successful funding of a subsequent main trial. This was certainly considered as part of the study plan but was not entirely possible due to a lack of accessible and complete data. Whilst all HTA studies submit and publish a full report via the NIHR journals infrastructure, the same is not true of RfPB studies. For these studies, a report is submitted to the funder, but is not currently made publicly available. It was hoped that all studies could be followed up to categorise the outcomes of the PFS, as shown in table 4.8 below.

***Table 4.8*** *Outcomes of surgical PFS where data publicly available*

| Outcome of PFS | | Proportion of studies | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | HTA, n=10 | | RfPB, n=25 | | TOTAL, n=35 | | |
| | | R (n=7) | NR (n=3) | R (n=22) | NR (n=3) | R (n=29) | NR (n=6) | All Studies |
| Main trial not feasible | | 2 (20%) | 0 (0%) | 3 (12%) | 0 (0%) | 5 (14.3%) | 0 (0%) | 5<br>14.3% |
| Main trial feasibility uncertain and | More work needed | 0 (0%) | 1 (10%) | 3 (12%) | 2 (8%) | 3 (8.6%) | 3 (8.6%) | 6<br>17.1% |
| | Another trial already in progress | 0 (0%) | 0 (0%) | 1 (4%) | 0 (0%) | 1 (2.9%) | 0 (0%) | 1<br>2.9% |
| Main trial feasible | Funded/In progress | 2 (20%) | 0 (0%) | 1 (4%) | 0 (0%) | 3 (8.6%) | 0 (0%) | 3<br>8.6% |
| | Completed | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0<br>0% |
| | Unable to get funding | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0<br>0% |
| Main trial feasible but no further studies planned (e.g. RCT no longer needed/done by others) | | 1 (10%) | 0 (0%) | 0 (0%) | 0 (0%) | 1 (2.9%) | 0 (0%) | 1<br>2.9% |
| Main trial feasible but uncertain if further work is planned (lack of data) | | 2 (20%) | 2 (20%) | 2 (8%) | 1 (4%) | 4 (11.4%) | 3 (8.6%) | 7<br>20% |
| No data on PFS outcome | | 0 (0%) | 0 (0%) | 10 (40%) | 1 (4%) | 10 (28.6%) | 1 (2.9%) | 11<br>31.4% |

For 11/35 (31.4%) studies (all RfPB funded), there was no available data for the outcome of the PFS, despite searching online (see methods) and contacting the NIHR and the chief investigators and/or study team directly. The NIHR were very helpfully able to confirm (see acknowledgements) where studies had been published and where studies had been granted further funding for a main trial, as

this data is in the public domain. For a further 7/35 (20%) of studies, whilst the report or published results stated that a main trial was feasible, it was unclear whether plans to design this trial and apply for funding were planned, started or in progress. There were 5/35 (14.3%) studies concluding a main trial was not feasible, 7/35 (20%) showing the feasibility of a main trial was still uncertain and either more work was needed before a definitive trial (6/35, 17.1%), or another separate trial was already in progress (1/35, 2.9%). A further 3/35 (8.6%) studies had progressed to securing main trial funding and/or beginning the trial. A single study (1/35, 2.6%), whilst demonstrating feasibility of the main trial, was not planning to proceed to a main trial, as the intervention in question was already falling out of use as a treatment option.

## 4.6 Summary

This review and analysis of NIHR funded PFS in surgery, has examined the type and scope of PFS conducted in surgery. Firstly, this work has demonstrated that surgical studies were rarely funded relative to all work funded by the NIHR in the ten-year period reviewed: only one in ten studies were examining surgery as the main intervention and only a quarter of these were PFS.

A result of this work was to group and label domains of uncertainty that can be explored in PFS. The five identified domains were 1) main trial design; 2) logistics; 3) recruitment; 4) intervention, and 5) outcomes. Within each of these, frequently explored specific uncertainties existed. For example, with regard to main trial design (domain 1), the overarching reason of examining and testing whether a main trial was possible was the predominant reason mentioned in many PFS protocols, which is unsurprising given that this is the principal rationale for performing PFS. The two other most prevalent reasons for performing PFS identified with regard to main trial design, were estimating the variability in outcomes to help determine main trial sample size, and performing or preparing for a cost effectiveness analysis of an intervention. With regard to logistics (domain 2) developing or testing questionnaires of surveys was by far the most commonly identified reason for performing PFS, and within recruitment (domain 3), assessing the numbers/rates of recruitment and consent was easily the

favoured uncertainty to explore. The intervention (domain 4) as described above was the least commonly explored area of uncertainty overall. Exploring uncertainty around outcomes (domain 5), was also not frequently planned for in comparison to issues around recruitment or logistics, and this is surprising, particularly given the need for patient and public involvement in this area of trial design. Secondly, these results suggest that PFS design is centred around addressing a favoured few uncertainties and that the full potential of PFS in surgery is, therefore, yet to be realised, despite the wider range of uncertainties identified, which could be specifically considered when designing surgical PFS.

Thirdly, whilst the expectation is not that every study should examine every uncertainty, it is apparent that surgical pilot work could be further optimised, with particular emphasis on exploration of the intervention. More than two thirds of all studies (25/35, 71.4%) did not plan to explore uncertainties around the intervention. Undoubted uncertainty existed around the intervention in two thirds (15/25, 60%) of these studies, simply by way of the trial design being to either compare surgery with no surgery or to evaluate a novel surgical intervention. The expectation would be that all PFS studies of surgical interventions, would explore some element of the intervention.

In addition, it is clear that the purpose of PFS in surgery is still misunderstood, with the continued inappropriate testing of the safety and effectiveness of interventions using PFS data. Poor dissemination of PFS results may also be contributing to widespread misunderstanding of the purpose, breadth and value of PFS in optimising surgical trials, with less than two thirds of studies (21/35, 61%) having published their results to date.

The next chapter (Chapter five) describes the results of the PEPSTAR qualitative interview study, which sought to explore in depth, key stakeholders' opinions, perceptions and experiences regarding why there are challenges and barriers to designing and conducting PFS for surgical trials, and what solutions might be needed to optimise surgical PFS in the future.

# CHAPTER 5

## *Phase II*

## *EXPLORING PERCEPTIONS AND EXPERIENCES OF PILOT WORK FOR SURGICAL TRIALS: A QUALITATIVE RESEARCH STUDY (PEPSTAR)*

## 5.1 Introduction

The targeted review and systematic analysis of NIHR funded PFS protocols (Chapter four) highlighted that surgical PFS are relatively rare, their design and conduct not optimised and that they continue to be poorly reported and disseminated. This chapter describes the findings from an in-depth qualitative interview study, which sought to explore the perspectives and experiences of key stakeholders around the challenges and barriers to undertaking PFS in surgery. In addition, the interview study considered what solutions might be needed to improve current research practice, in order to inform development of recommendations to improve research practice in this area.

## 5.2 Demographics of the sample

A total of 33 participants were invited to participate in an interview by the research team (see Chapter three, Methods, section 3.8.3 for details of how the participants were selected). Of these, 28 (85%) expressed an interest and 27 (81%) consented and were interviewed. These included 18 (67%) males and 9 (33%) females. The interviews were undertaken over a period of ten months, in three

phases relating to the iterative phases of the analyses (first phase: participants 1-8; second phase: participants 9-20; third phase: participants 21-27). Most interviews were conducted by telephone (17/27) with the remainder (10/27) face-to face, mostly in university office locations. The length of interviews varied from 27 to 101 minutes (mean 58 minutes). Participants were working throughout the UK including Scotland (2/27), Northern England (6/27), the Midlands (5/27) and Southern England (14/27). Further demographic characteristics and research experience of the 27 participants interviewed are described in Table 5.1.

### *Surgeons*

Ten participants held a current clinical role as a surgeon, and one further participant had previously held a clinical role as a surgeon but had recently retired from practice. Surgeon participants were predominantly male (10/11) and were qualified in a broad spectrum of surgical specialties including hepatopancreaticobiliary (HPB), oesophagogastric, cardiovascular, colorectal, breast, vascular and orthopaedic surgery. All the surgeon participants interviewed were currently directly involved in surgical research (for example, as a chief investigator or co-applicant on a surgical research study and/or currently recruiting patients into surgical research studies) but experience of involvement in pilot and feasibility research studies varied widely. All had experience of being the chief investigator of trials: two thirds (8/11) had experience of clinical trials of surgical interventions, with a quarter (3/11) being experts in this area (primarily university-employed to do research, and holding an honorary NHS contract); one quarter (3/11) had experience of predominantly translational and pharmaceutical studies (within the context of surgery), and; one surgeon had extensive experience of studies of novel (innovative) devices and technology in surgery.

### *Methodologists*

The 16/27 participants who did not identify as surgeons, identified primarily as trial methodologists. Half (8/16) of the methodologist participants were females and half males. All reported experience in designing and conducting trials of complex interventions and some had a specific experience in surgical trials (3/16). All had experience of designing and conducting both internal and external pilot

studies. Of the methodologist participants, 9/16 were Clinical Trials Unit (CTU) directors, 4/16 were trial statisticians and one was a trial methodologist. Of the 16 methodologist participants, 4/16 held or had previously held a role as a doctor in a clinical specialty other than surgery.

### *Funders and editors*

More than two thirds (20/27) of participants interviewed had current or recent experience of membership on a UK research funding body panel. Of these, 6 had a position as a panel chair, 3 as a deputy chair and 12 as a panel member, though the length of individual experience varied considerably. More than half (15/27) were part of journal editorial teams and most (24/27) also currently held a professorial position at a UK university.

### *Groups of participants*

During data analysis, it became apparent there were three distinct groups amongst the participants, based on the level of trials methodology experience they had. These groups are defined and termed in the narrative account as: Surgeons (with little or no methodological experience), methodological experienced surgeons, and methodologists. The methodologically experienced surgeons all held a formal academic position at a UK University (not an honorary position) and all had experience of major funding panel membership and/or chair/deputy chair positions. As described in the methods, participants were asked at the start of the interviews about their roles relating to PFS, which enabled identification of those with significant methodological experience. It was considered important to distinguish between these groups because it was hypothesised that the level of methodological experience held by participants would influence their perceptions and opinions. In addition, providing further individual detail in Table 5.1 of specific methodological expertise, would have risked participant anonymity.

***Table 5.1*** *PEPSTAR participant demographic characteristics and surgical research/trials experience (n=27)*

| Participant ID | Gender | Date of interview | Length of interview (HH:MM) | Mode of interview | Clinical role | Roles held currently/recently relating to research and/or trials | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Trial involvement | Funding panel member | | | Editor | | University academic position held |
| | | | | | | | Current | Position | Previous | Current | Previous | |
| 001 | Male | 04/02/17 | 00:29 | Face-to-face | S | CI | | | | x | | Professor |
| 002 | Male | 04/02/17 | 00:27 | Face-to-face | S | CI | | | | | | Professor |
| 003 | Male | 17/02/17 | 00:57 | Face-to-face | S | CI | | | | | x | Professor |
| 004 | Female | 23/03/17 | 00:30 | Face-to-face | S | CI | | | Deputy chair | x | | Professor |
| 005 | Male | 10/04/17 | 01:01 | Face-to-face | S | CI | | | | x | | Professor |
| 006 | Male | 10/04/17 | 01:41 | Face-to-face | S (previous) | CI | | | Member | | x | Professor |
| 007 | Male | 11/04/17 | 01:16 | Face-to-face | S | CI | | | | x | | Professor |
| 008 | Male | 18/05/17 | 01:22 | Face-to-face | S | CI | x | Member | | x | | Professor |
| 009 | Male | 31/07/17 | 01:08 | Telephone | O (previous) | CTU director | | | Member | x | | Professor |
| 010 | Female | 20/07/17 | 00:56 | Telephone | | Statistician | x | Member | | x | | Professor |
| 011 | Male | 01/08/17 | 00:42 | Telephone | O | CI | x | Chair | | x | | Professor |
| 012 | Female | 02/08/17 | 00:57 | Telephone | O (previous) | CTU director | | | Member | | | Professor |
| 013 | Female | 03/08/17 | 00:40 | Telephone | | Statistician | x | Member | | | | Principal Research Associate |
| 014 | Female | 08/08/17 | 00:50 | Telephone | | Statistician | | | | x | | Professor |
| 015 | Male | 08/08/17 | 01:12 | Telephone | | Methodologist | x | Deputy chair | | x | | Professor |
| 016 | Male | 09/08/17 | 00:55 | Telephone | | Statistician | x | Member | Chair | x | | Professor |
| 017 | Female | 10/08/17 | 00:57 | Telephone | | CTU Director | x | Chair | | | | Professor |
| 018 | Male | 11/08/17 | 01:25 | Telephone | | CTU Director | | | Member | x | | Professor |
| 019 | Male | 14/08/17 | 01:28 | Telephone | | CTU Director | x | Member | | | | Fellow |
| 020 | Female | 15/08/17 | 01:23 | Face-to-face | | CTU Director | x | Member | | | | Professor |
| 021 | Male | 04/09/17 | 00:41 | Telephone | S | CI | | | | x | | None |
| 022 | Male | 11/09/17 | 00:47 | Face-to-face | S | CI | x | Chair | | | | Professor |
| 023 | Male | 14/09/17 | 00:44 | Telephone | S | CI | | | Member | x | | Professor |
| 024 | Male | 18/09/17 | 00:36 | Telephone | O (previous) | CTU Director | x | Chair | | | | Professor |
| 025 | Female | 09/10/17 | 00:58 | Telephone | | CTU Director | x | Member | | x | | Professor |
| 026 | Female | 11/10/17 | 01:01 | Telephone | | CTU Director | x | Deputy chair | | | | Professor |
| 027 | Male | 05/12/17 | 00:52 | Telephone | O (previous) | CI | x | Chair | | | | Professor |

***KEY:*** *CI = Chief Investigator, CTU = Clinical Trials Unit, S = surgeon, O = other clinical specialty, (previous) = no longer practicing in a clinical role*

## 5.3 Results from a thematic inductive analysis of the PEPSTAR qualitative interview data

### *Derivation of themes*

As described in the methods (Chapter 3, section 3.8.5), analysis of the PEPSTAR interview data was undertaken using a thematic inductive approach as described by Braun and Clarke [180]. This approach includes six stages named as: 1. Familiarising yourself with the data; 2. Generating initial codes; 3. Searching for themes; 4. Reviewing themes, 5. Defining and naming themes and; 6. Producing the report.

### *Evolving themes*

An initial mind map of numerous codes and groups of codes was produced as familiarisation of the data progressed (Figure 5a). As further interviews were completed codes were checked, amalgamated, renamed as appropriate and new codes added as necessary, leading to an evolving coding map as illustrated in Figure 5b. Once all the interviews were complete, the process of searching for, reviewing and refining themes began leading to the thematic diagram illustrated in Figure 5c. This figure illustrates how the different data containing codes began to relate to each other and how they began being amalgamated to produce themes and subthemes. Once the narrative account was begun, it became possible to determine the story that each theme told individually and consider further how these themes interlinked (Figure 5d). Ultimately, this process led to the themes and subthemes being fully defined with working titles as illustrated in Figure 5e.

### *Final thematic framework*

Figure 5e illustrated the final thematic framework, with three distinct but related key themes of: A) Differential understanding of PFS; B) Challenges of PFS for surgical trials and; C) Solutions to optimise the design and conduct of PFS for surgical trials. This diagram (Figure 5e), illustrates the format of the narrative account in this chapter to follow, where each of these themes and sub themes will be discussed in turn.

***Figure 5a*** *Coding map determined from early interviews (n=8) with surgeons*

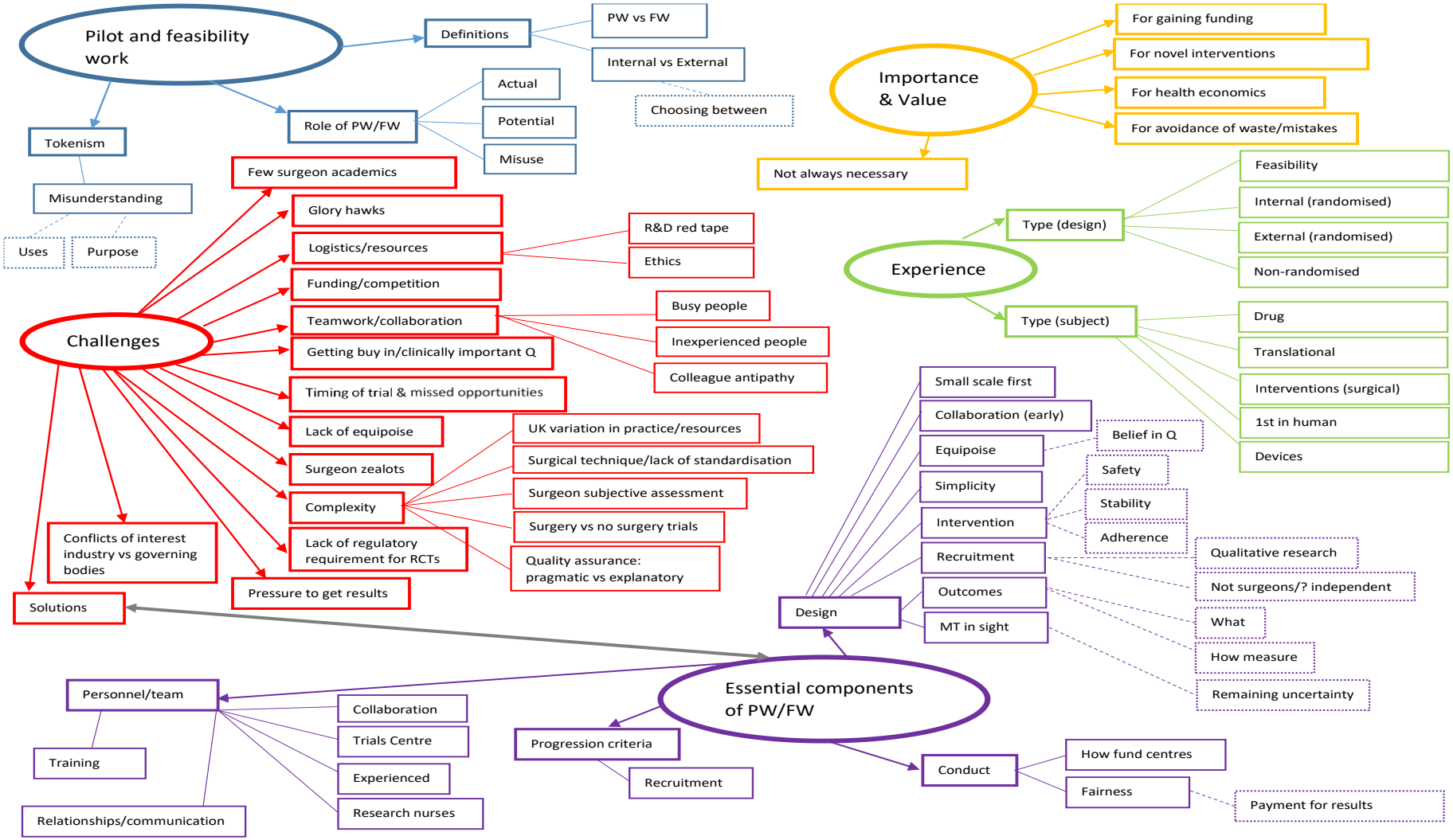


***Figure 5b*** *Coding map illustrating the ongoing organization of codes generated from the first 20 interviews.*

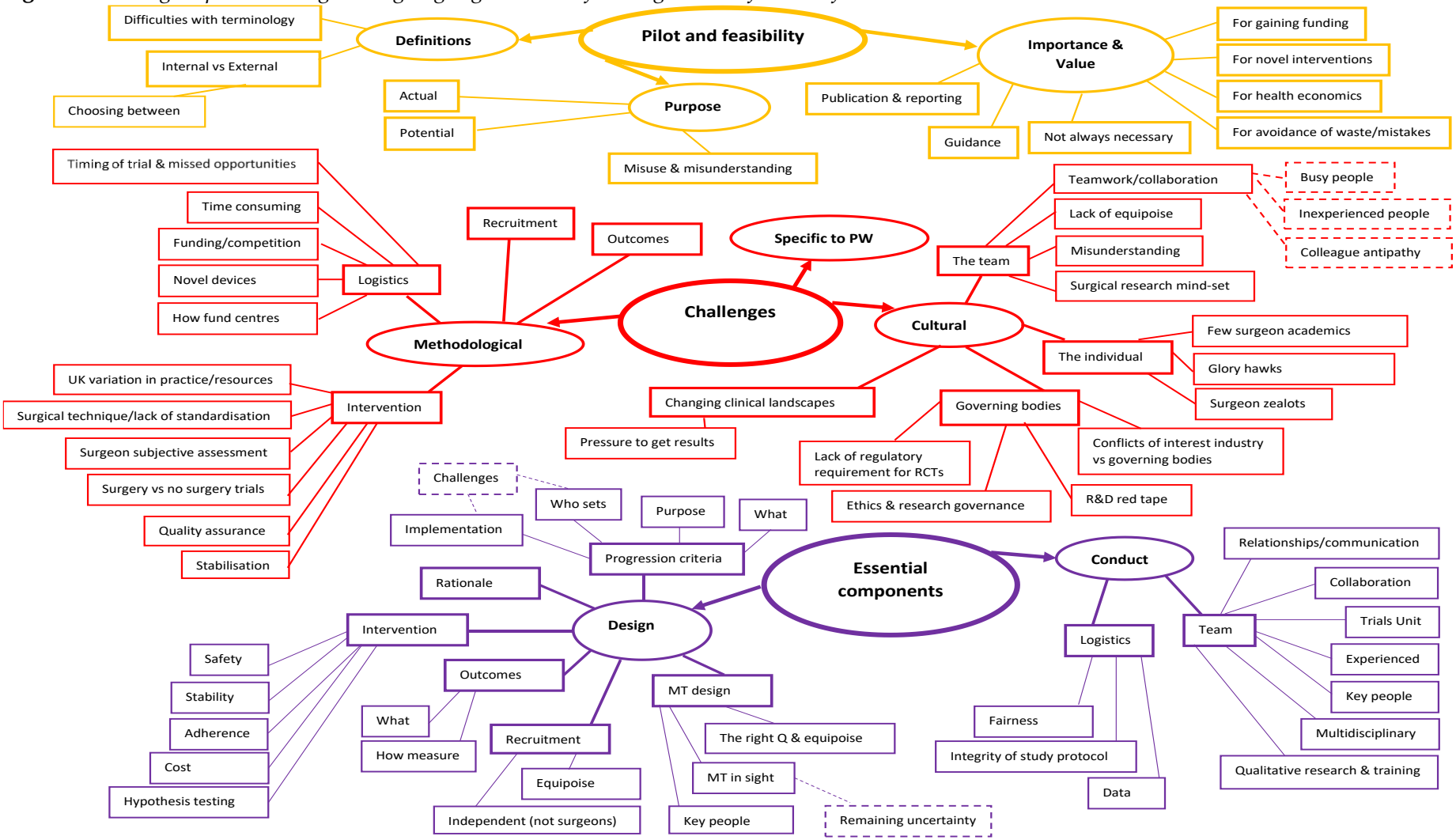


***Figure 5c*** *Searching for and developing potential themes and subthemes from the final data set*

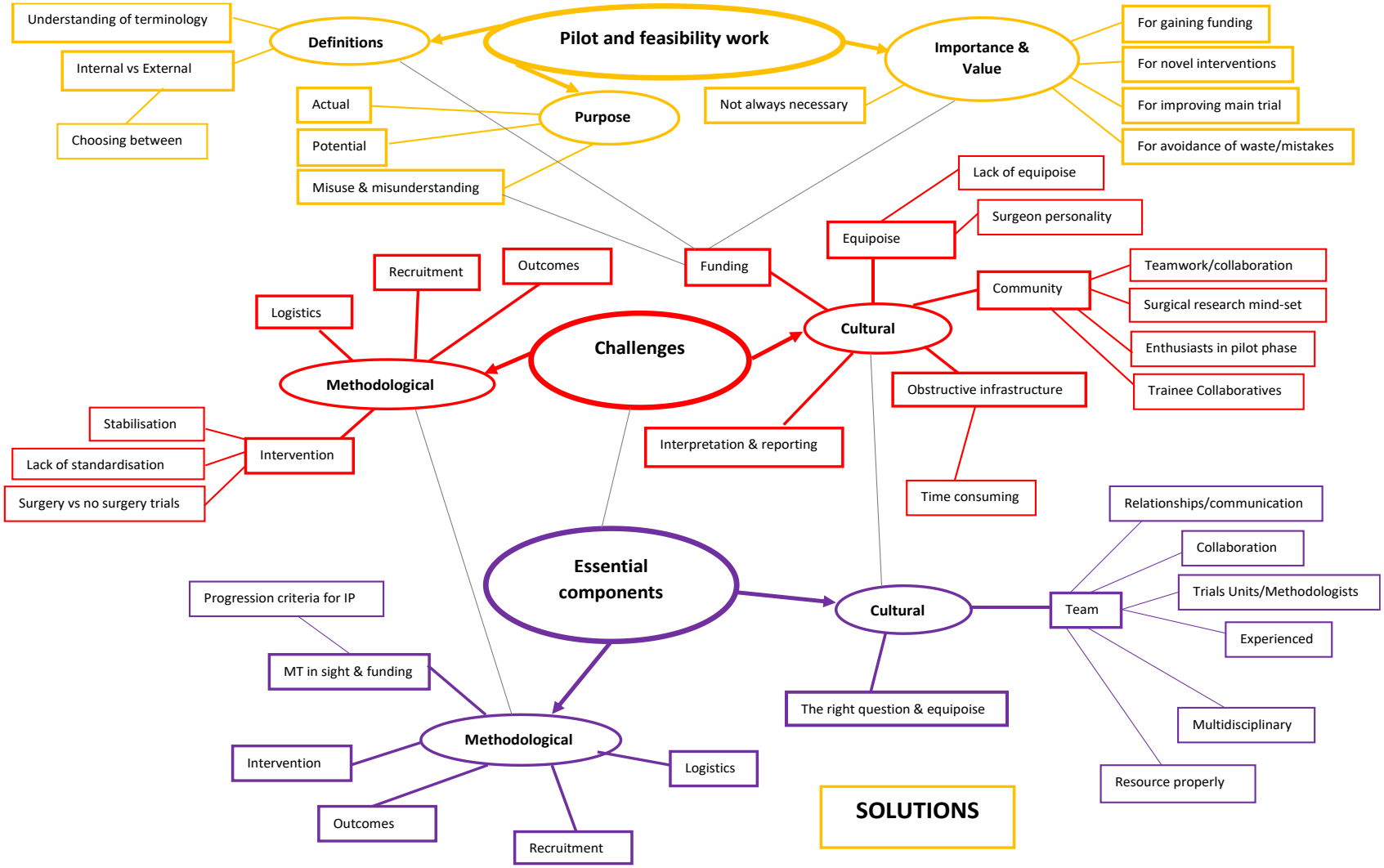
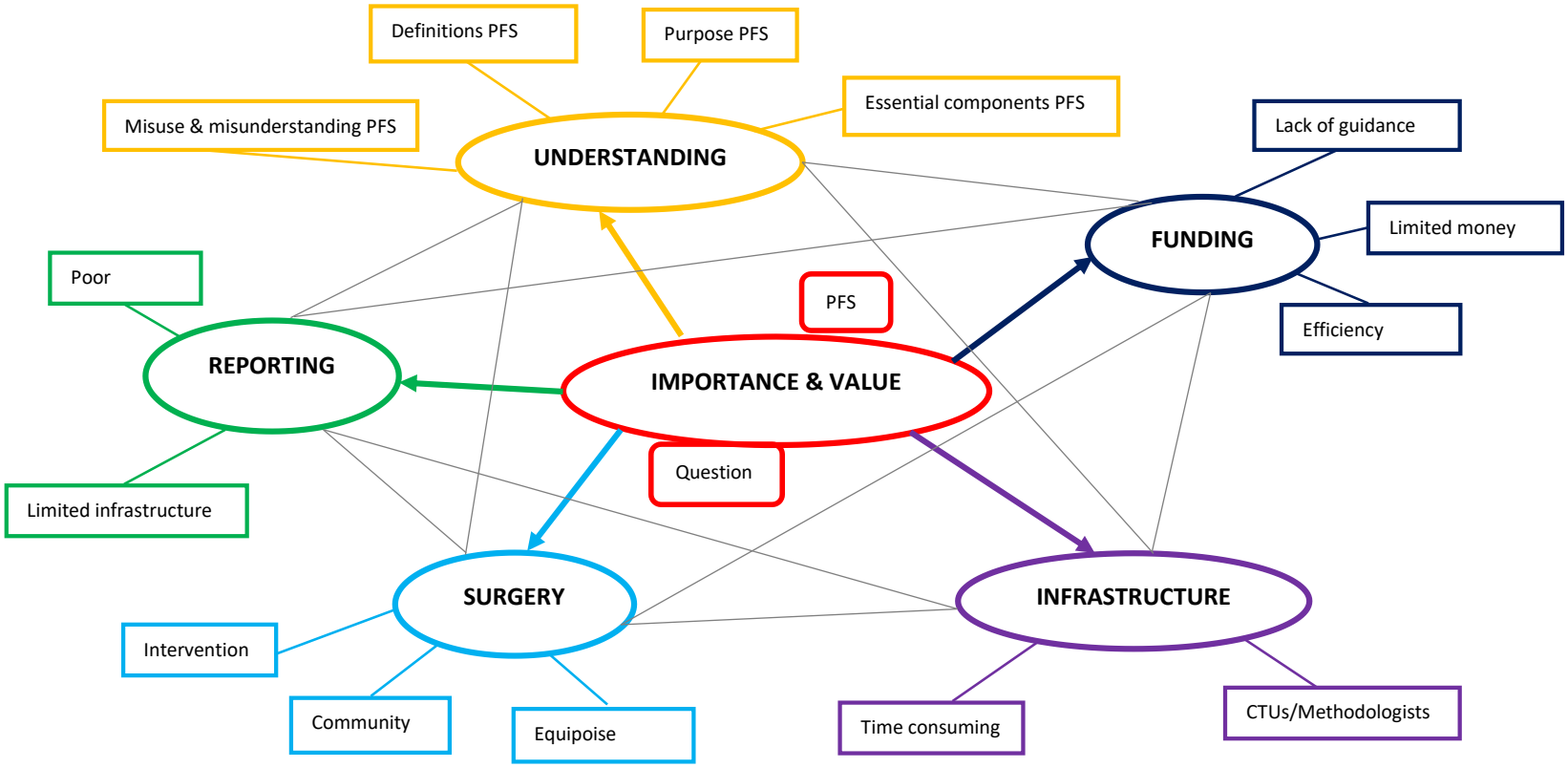
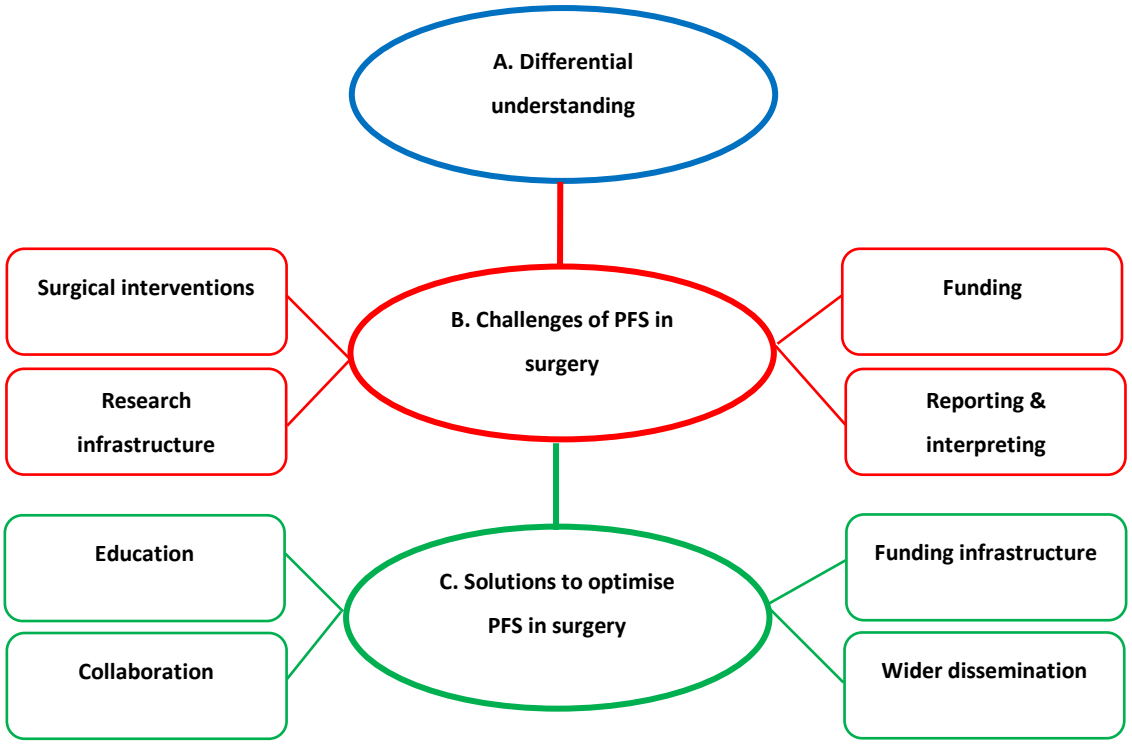


***Figure 5d*** *Reviewing and refining themes and subthemes from the final data set*




***KEY****: PFS = Pilot/feasibility studies, CTU = Clinical Trials Unit*

*Figure 5e* ***Final thematic framework illustrating the key themes and sub themes that emerged from the in-depth interview study.***




***KEY****: PFS = Pilot and feasibility studies*

= Main theme = Subtheme

## 5.4 A) Differential understanding of pilot and feasibility studies

*008 (Surgeon/Funder): "I think the truth is most, many surgeons even those involved in trials don't actually understand what feasibility and pilot work is and confuse it with a… it's just a smaller trial, well no it's not, it has a completely different role, maybe it's stepping stone role but a very different role […] so, it's not universal, but I think actually if you ask me what the average surgeon understood by that, they wouldn't have a clue really…"*

### 5.4.1 Differential understanding of purpose

A key finding throughout the data was a differential understanding between methodologists and surgeons of the purpose of PFS. Methodologists demonstrated a more complete and accurate understanding of the principal purpose of PFS in optimising the design and conduct of a future main trial.

*014 (Methodologist): "Well you hope it's going to improve it; stop the trial going ahead if it's too problematic. I just think there's so many things to think about, certainly in complex interventions, that I just can't see that you can think about them all until you've had one dry run."*

*010 (Methodologist/Funder): "You've got questions and objectives that you need to answer that will inform your running of the main study. They'll either let you move forward with confidence that the processes that you've got in place are gonna work, or it will demonstrate that you need to alter them in some way."*

In contrast, surgeons generally had a much narrower appreciation of what PFS are and how they can be utilised. Throughout, surgeons tended to focus on reasons that were more pragmatic for performing PFS, for example as a pre-requisite or necessary step to getting a main trial funded.

*005 (Surgeon): 'I think pilot work is very important but mainly to demonstrate value for money that you can actually recruit to time and target, because pilot work in my opinion it's more to benefit the funders […] I think whether we do a pilot or not really depends on money; that's what it boils down to. Every trial if possible, should be set up as a pilot to full […] in the real world that's not always the case…it boils down to the funds; sums that you can actually raise. So, if you can only raise enough for a pilot or feasibility, standalone, no matter how you call it, then that has to be how it is."*

*003 (Surgeon): "I think it's essential for big studies and also its essential cos you won't get it funded otherwise."*

Other surgeons seemed to recognise only the face value of PFS by considering the feasibility of the main trial in very broad terms only.

*007 (Surgeon): "The main role of the feasibility trial is to decide whether it's worthwhile doing a clinical trial."*

Some surgeons also regarded PFS as playing an important role in evaluating the safety and effectiveness of new surgical interventions not widely used in current clinical practice. This illustrates misunderstanding amongst surgeons about what PFS are and how they should be used.

> *003 (Surgeon): "I think it's extremely valuable if you're looking at a surgical technique or a new surgical bit of kit, because I do not think we should introduce things willy nilly, which we're notorious for doing."*

> *006 (Surgeon): "I suppose increasingly, people will be less and less able to start doing something different, so instead of: 'We happen to have pilot data cos we've all been having a go at it', we will have to translate this to: 'If you start this, you've got to document'."*

Furthermore, some surgeons incorrectly perceived that the main purpose of PFS was to act as a formal mechanism to discard unfeasible new technologies and to speed up the process of evaluating multiple new interventions and devices simultaneously.

> *023 (Surgeon): "I think that [PFS are] absolutely fundamentally important, so I think it's very very important, not just important. And the reason I say that is, we have to start designing processes that sort the wheat from the chaff […] We need to be evaluating 10 technologies at once and decide that we're taking two forward very quickly […] I see [PFS] as a mechanism for discarding things that don't work, and in a way, it's an early phase trial in terms of the selection process."*

It was not always appreciated by surgeons that the fundamental function of PFS in optimising main trials was, as an essential precursor to generating high quality randomised evidence. A perceived external pressure for PFS to provide findings related to intervention effectiveness compounded this.

> *003 (Surgeon): "I wouldn't be a fan of doing a pilot study to see whether the design of the study worked…otherwise people are gonna go 'oh what', if you've not given them an answer; surgeons are pretty binary."*

Funders, by contrast, perceived the importance and purpose of PFS as a way of managing risk. PFS were regarded as particularly important in surgery, where there is often significant uncertainty relating to either the pertinence of the research question, and/or the ability of the surgical trial team to deliver a main study.

> *011 (Funder): "…if it's an important question, then we should be trying to answer it and we should be prepared to take a bit more risk and spend a bit more money trying to answer it. So that's why pilot and feasibility work is so important, because it gives us a way of managing that risk."*

> *017 (Methodologist/Funder): "I think surgery absolutely is up there with the kind of things you think, are they ready to launch into the main expensive study yet?"*

Some funders, therefore, emphasised how PFS served as an opportunity for the surgical team to prove that they could deliver a project within a certain time, thereby reducing their uncertainty about funding a future main trial.

> *008 (Surgeon/Funder): "They will take a chance on new investigators if they've got the right support, but they want to see that they're not brand new. You can't have somebody who's got no experience so if you've done that standalone pilot funded by industry, RfPB, charity whatever, and you've taken that through to completion and then that informs a bigger study, then actually that is pilot work but not about the interventions, outcome measures, participation. It's about the team, the team showing they can deliver this project on a limited scale. Now they'll take a chance, and we'll deliver it on a big scale."*

### 5.4.2 Differential understanding of scope

A differential understanding between methodologists and surgeons of the scope of PFS was also identified. Methodologists, for example, understood the full scope of PFS to explore all areas of uncertainty about a potential main trial.

> *025 (CTU Director/Funder): "I think if you've not got uncertainty then no you don't need to do any [PFS], but I think in most cases you would have some kind of uncertainty, so some sort of pilot and feasibility work, however small, would be necessary."*

All methodologists therefore perceived PFS as having a particularly important role in surgical trials where there is often significant uncertainty in multiple areas.

> *018 (CTU Director): "The more uncertainty you've got, the more you need a pilot, and if you're uncertain about surgeon equipoise, uncertain about patient equipoise, uncertain about the actual intervention then, compliance whatever, then that increases the probability you need a pilot."*

The surgeons were generally less specific about the breadth of uncertainty that could be explored in PFS, tending to focus on recruitment as a primary area of uncertainty.

> *002 (Surgeon): "Well it's all about recruiting isn't it, at the end of the day? And the trial's got to be attractive to clinicians and the staff who are addressing the patients. So, anything that obstructs recruitment has to be addressed in a pilot study, I would have thought…"*

> *021 (Surgeon): "So it's good for telling you about recruitment and it's good for telling you about compliance because sometimes you can be as clever as you like, but if people won't consent into a trial then they're never gonna work so…there's lots you can learn from them."*

A smaller number of surgeons with more methodological experience, however, demonstrated a greater depth of understanding about the range of uncertainties that PFS could address before a main trial.

> *004 (Surgeon): "I think it's done for lots of reasons. One is to improve the design of the main trial in terms of, to work out how to recruit, best methods to recruit, how to find patients, how to establish eligibility criteria. How to get the interventions standardised, how to work out which outcome, how to measure the outcome, how to learn how to teamwork, all those things."*

These more methodologically experienced surgeons also emphasised the importance of using PFS as a vehicle for teaching and training surgeons how to perform trials.

*004 (Surgeon): "Just to practice it cos that's the problem surgeons haven't done it before, so to cut their teeth and actually faffing around with all the organisation and paperwork that's involved in running a trial. I think that's a really good place to do it, in the pilot phase or the feasibility work. So, it's all just about preparation for hopefully running a seamless main trial."*

*008 (Surgeon/Funder): "Sometimes we do pilot work as vehicles for training […] it becomes a training vehicle to run a standalone pilot for someone and then the trainee or whoever is leading that would then, or sometimes a new investigator […] so you're trying to provide support to them in a much more contained environment without the [funder] level overview."*

### 5.4.3 Differential understanding of pilot and feasibility studies for limiting research waste

Exploring all areas of uncertainty to optimise the main trial, enabling better trials and better evidence with which to change practice, was identified as the main objective of PFS by methodologists. However, the methodologists also understood that in doing so, this avoids failed trials, thereby limiting research waste.

*027 (Methodologist/Funder): "I think the problem is, we've wasted a lot of money by not having adequate development and fidelity work before heading off into the deep blue yonder of the main trial."*

Some surgeons were also aware of the value of PFS in reducing research waste by demonstrating when a main trial would not be possible.

*001 (Surgeon): "I'm sort of quite conscious of the amount of money in the pot to fund things, and I think there needs to be some evidence to support a major investment otherwise the money could be spent better somewhere else."*

*022 (Surgeon/Funder): "Many people think that a successful pilot or feasibility study is one that shows you can do the trial. I think some of the most successful ones show that you can't do the trial that you were sure you could do. So, it saves years and years of work and sometimes millions of pounds worth of money. A lot of people are looking 'so I've failed', but actually they've succeeded."*

However, despite recognition that PFS may help in limiting research waste, some felt that PFS had also contributed to it. Several methodologists noted that historically, PFS had often been funded wastefully, without any intention (of the investigators, the funders, or both) of progressing to a main trial. Whilst PFS may show that a main trial is viable, if it does not progress to a definitive study, then this could be considered as adding to research waste.

*009 (CTU Director): "Sometimes you see pilot feasibility work, is seen as an end in itself […] [Funders have] funded masses of feasibility studies, most of which have gone nowhere and*

*produced nothing of any interest […] I look at these things and say well, why would we go for these things? They're three years of work, they're usually very thin on resource, they're a lot of hard work, and at the end of it we get nothing out of it that is of any use to us […] You can write up your feasibility work, it will go into the Ruritanian Journal of Unreproducible Results, nobody will ever read it again and it ticks a box for our masters that we're generating income. but it doesn't generate us any really top-class research output, so most of it goes nowhere."*

In contrast, funders perceived that increasing attention was being paid through greater consideration in recent years of the plausibility and likelihood of a main trial. Scrutiny of the relative importance of the research question when applicants proposed a PFS was regarded as crucial to this.

*019 (CTU Director/Funder): "That's the process at [Funder] these days, for an opinion on it from [Funder] about whether this is interesting and sometimes that comes back as: 'We can see no way that the [Funder] would fund this study, even if the pilot was successful', and that's often, it's usually the kiss of death. Because the [Funder] sees, if there isn't anyone prepared to fund the main trial, what's the point of doing the pilot."*

*016 (Methodologist/Funder): "I think that a funding panel could agree that a study was well designed, it had sensible objectives and it had methodology to address those objectives. But if it was asking a clinical question that ultimately wasn't going to be of enough priority to [Funder], then it probably shouldn't go ahead, because, you know that's then wasteful."*

### 5.4.4 Reasons for differential understanding

A number of reasons for differential understanding of the scope and purpose of PFS amongst surgeons and methodologists were identified, but central to these was the perceived inaccessibility of guidance regarding the design and conduct of PFS.

#### *Inaccessibility of theoretical/conceptual guidance*

Most methodologists acknowledged and described key theoretical and/or conceptual literature (for example, scientific publications) on the optimal design and conduct of PFS but many felt that these publications were inaccessible to the wider trials community being too technically complex. Rather than help, many methodologists perceived this literature might actually add to the confusion surrounding the scope and purpose of PFS.

*015 (Methodologist/Funder): "There are some nice papers actually on pilot and feasibility studies, particularly to do with sample size and I think they're a bit… I mean they are very useful if you're a methods person, I think they're a bit tricky if you're not […] so they're good, but it's obvious from the stuff we get from the [name of funder], that that message is not getting across through those articles in a way that is understandable."*

In contrast, the surgeons either did not reference the conceptual literature at all, or if they did, focused on the literature that related to definitions and differences between types of pilot and feasibility studies. Surgeons therefore perceived the conceptual work as irrelevant in practical terms and, again, only important for a successful funding application.

> *008 (Surgeon/Funder): "I struggle with that division between feasibility and pilot despite having read and re-read all the crap describing the differences, actually I don't think the differences are particularly meaningful […] that division is probably not important but only exists, yes in funding scheme applications.'*

Some methodologists specifically referenced conceptual tools such as the MRC guidance for designing and evaluating complex interventions as being helpful in the design of PFS. However, none of the surgeons mentioned use or awareness of this guidance document.

> *025 (Methodologist/Funder): "I think I would go much more on the MRC framework [for complex interventions] which I think is very helpful actually […] They talk about uncertainty, that everything should be based on uncertainty, and I do think that's true in pilot/feasibility work as well as in the main clinical trials that we do. So, I think that what people ought to be doing is to think about where their uncertainty is and try and build the pilot/feasibility work around that uncertainty."*

Despite its unique relevance to surgery, participants rarely referred specifically to the IDEAL framework when discussing PFS. One methodologist, who acknowledged the IDEAL framework as a positive step towards changing the way surgical research is done, felt that most surgeons did not really understand the value and importance of PFS, and were not commonly using the IDEAL framework to help design such studies.

> *014 (Methodologist): "In my opinion most surgeons haven't got as far as pilot and feasibility studies, there wasn't very much awareness of them […] I got the impression that not many people were using the IDEAL framework and weren't aware of it […] They seem to be getting away with doing mini randomised controlled trials and, as I've described, they shouldn't really be doing that […] I felt that they hadn't got as far as doing pilot and feasibility studies, or embracing them in their work."*

### *Inaccessibility of funder endorsed guidance*

Like the conceptual and theoretical guidance, most participants perceived funder guidance as generic, and offering limited direction in how to operationalise the guidance in practice. Funder guidance was, therefore, perceived as another potential barrier to widespread and comprehensive understanding around the design and conduct of high quality PFS. Whilst most funders offer some guidance

on their websites, this is mainly limited to brief definitions of PFS, which vary between funders.

Most methodologists, for example, perceived that the definitions of PFS endorsed by funders, such as the NIHR, encouraged applicants to mould their PFS study design to match a specific definition.

> *009 (CTU Director): "I think people get too bogged down sometimes in exactly what the words and definitions mean."*

Some methodologists expanded on this viewpoint, explaining that definitions incorporating specific reasons for undertaking PFS (for example, to explore recruitment) may restrict applicants from thinking more widely, and encourage them to ignore other important uncertainties around their specific research question.

> *019 (CTU Director/Funder): "I think they [the definitions] push people into particular boxes, whereas different questions are better answered using different approaches. As for almost any research question you care to name there are pros and cons to different approaches but here, the question you're answering is 'what's stopping me doing the main trial'."*

> *025 (Methodologist/Funder): "I think the key list of things is also a bit of a stumbling block, and that's in the NIHR feasibility definition, they have this list of things, and I think people think they've got to do that."*

Similarly, surgeons generally perceived the defined methodological differences between 'pilot' and 'feasibility' studies as only relevant in terms of funding applications and felt that confusion existed even amongst the funding panel members.

> *005 (Surgeon): "It's an interesting question you've asked and the distinction between feasibility and pilot is not very clear in the funding body's perspective."*

Most methodologists agreed, in line with the NIHR definitions that the major difference between an internal and external pilot was that an internal pilot was integral to the main study, with the data contributing to the final analysis, and that an external pilot was a standalone piece of work. However, many methodologists expressed concern over the unhelpfulness of these terms, feeling they may be contributing to general confusion.

> *019 (CTU Director/Funder): "But the internal pilot, in the last couple of years I've come to believe an internal pilot just shouldn't exist. In [that] it's not particularly a useful term, and it's a bit artificial in terms of how you then handle it."*

> *015 (Methodologist/Funder): "I wouldn't call it an internal pilot. I think we're treating [an internal pilot] more like monitoring. I know it's still, of course, a possibility that in the three*

*sites you did it in, your external pilot things went fine and as soon as you go to some sites that were not involved, things do not go fine. But I wouldn't call that a pilot, we're just now running the trial, and we're trying to make it work."*

The terms internal and external pilot are used almost exclusively by the NIHR. Representatives from other funding bodies also found the distinction to be mostly irrelevant to what really happens in practice.

*024 (CTU Director/Funder): "I would think that if we were doing a pilot study, what we're doing is, well it is proof of concept, is the concept warranting further study. So it's pilot data to put into your grant application. So, that's not the way the NIHR think about it, I don't really know how they think about it quite honestly […] we never talk about pilot studies in that sense, we don't talk about internal pilots, and I know NIHR does, but I don't know what they mean about it. So, there is clearly a distinction between what funders, how funders use these terms."*

Perhaps unsurprisingly, the surgeons demonstrated difficulty understanding the predominantly funder-coined nomenclature of 'internal' and 'external' pilot studies.

*005 (Surgeon): "No I'm not too familiar. What I understand from an internal pilot is that you get a set number of either one centre or set number of centres first and then followed by… or the external component is people who join in after the progression goals have been met."*

*001 (Surgeon): "I think that, when you say internal do you mean you do a pilot study and then it continues to become a larger trial?"*

Some methodologists also felt that a lack of universally accepted terms and explanations of types of PFS across funding bodies added to applicants' confusion.

*026 (CTU Director/Funder): "I think although it is well defined, the terms are used interchangeably by researchers, and perhaps may even be used differently by other funding bodies […] From a funders point of view […] when we see the terms, we expect a certain set of questions to be, a range of questions to be answered, depending on the term that's given, and also we're expecting a certain kind of discussion around those terms. So, if the terms are not used appropriately it does cause confusion […] It would be better, yes it would be better if we all had a terminology that we understood more clearly."*

Both surgeons and methodologists felt that funders' definitions and descriptions of PFS drive what is written in funding applications. However, some methodologists also felt that reviewers and journal editors' misunderstandings of the purpose and scope of PFS potentially caused further confusion.

*025 (CTU Director/Funder): "What the funders say drives a lot of what people write in their application, so obviously they're going to try and match what they say in their funding applications with the definitions that the funders have produced, and they think that's going to be more likely to get them funding. I think the second thing […] is that there's a lot of confusion about what people should be doing in these studies, and I think a lot of that comes from misunderstandings of journal editors and reviewers and trying to push people in a certain direction."*

## 5.5 B) Challenges of pilot and feasibility studies for surgical trials

### 5.5.1 Challenges with designing and conducting pilot and feasibility studies in general

Methodologists and surgeons with methodological experience identified several challenges with the design and conduct of PFS more generally, which were not necessarily related to surgical trials specifically. When considering a main trial, challenges were perceived as predominantly managerial and logistical relating to running the trial on a larger scale. In contrast, the challenges of PFS were perceived as more fundamental, relating to the many potential uncertainties about how the research question should best be answered.

> *022 (Surgeon/Funder): "There are different challenges so… when you go into the main trial it's upscaling it all so your feasibility and your pilot work […] actually the fundamentals of how you run your trial. You could have made some decisions around that that are basically wrong or you…there were things you didn't know that you had to find out whilst you're running a multi-centre thing, about logistics, understanding of a health care setting for each individual one […] You might have 6 or 7 different outcome measures, you wanna decide which one you wanna use but you're not gonna be faced with that in the main trial […] So they're different kinds of problems. It's much more managerial [in the main trial] as to… you've got the basic machine that can run now, it's now making that machine run in all these various settings and trying to make it run the same way in all the various settings, when every hospital is unique in the way it runs things."*

#### *Challenges with identifying uncertainties*

Some methodologists noted that PFS are inherently challenging specifically because their objective is to focus on addressing uncertainty or 'unknowns', some of which are anticipated (known unknowns) and some of which are not (unknown unknowns).

> *020 (CTU Director/Funder): "I think what's so difficult and challenging with feasibility studies is that it is a walk into the unknown. You might not know what you don't know, and actually just even costing them, planning for them is really challenging because you feel like you're going down a certain path with your sat nav and then suddenly, you're kind of going off in a different direction"*

> *017 (CTU Director/Funder): "They're inherently difficult because by definition you're doing them because you're not quite sure how best to do it, so they're not off the shelf things but yea that's also the fun of them."*

One participant referred to a quote from Tolstoy in Anna Karenina: *"All happy families are alike; each unhappy family is unhappy in its own way"*. The participant, a

surgeon, associated this philosophy with the common identifiable challenges that might be applicable to most PFS, and other more specific challenges that are individual to the study and may, therefore, be more difficult to identify *a priori*.

> *022 (Surgeon/Funder): "What does Tolstoy say?[…] All happy families are happy in the same way but not every unhappy family is unhappy in their own unique way (laugh)[…] there'll be some commonalities but there are always some things that you never thought about."*

Following on from this, methodologists and surgeons with more methodological experience perceived that identifying which uncertainties need exploring was more challenging than designing an appropriate PFS study to explore those uncertainties.

> *008 (Surgeon/Funder): "More people fall down because they've not addressed the gaps where the uncertainty is, rather than they've not designed the study well enough to address the gaps they have identified."*

### *Uncertainties to address in PFS*

All participants perceived recruitment as an important area of uncertainty to explore in PFS. Indeed, most surgeons, focused on recruitment as the key uncertainty to explore in PFS.

> *004 (Surgeon): "I think for surgical trials, I think recruitment still remains to be… that's the biggest reason why trials run into trouble. So, recruitment I think is really important."*

However, methodologists and surgeons with more methodological experience identified several other additional uncertainties that were of key importance to address in PFS. In addition to recruitment, these included exploring the intervention and examining outcome measures.

> *017 (CTU Director/Funder): "The thing we worry about as funders, and you'll hear me keep saying it, recruitment, recruitment, recruitment. Another one on my list is implementation of the intervention would be a general thing, that and can you collect the outcome measures."*

> *022 (Surgeon/Funder): "So, you know we can recruit the patients, we can do the intervention and we can follow them up and get an appropriate primary outcome measure. I think those are the 3 most important things."*

In contrast to surgeons though, most methodologists and funders, placed particular emphasis on using PFS to address uncertainties surrounding the surgical intervention itself.

> *027 (Funder): 'I guess the main, the first thing is, you want to know that your intervention is feasible, that it is acceptable, so the feasibility study will often involve exploring issues of the acceptability of the study intervention.'*

Whilst the surgeons also described some elements of the intervention that may be explored in PFS, they were generally less explicit in defining the intervention as a key parameter to explore for surgical studies.

In addition to the three key areas of recruitment, the intervention and outcome measures, other logistical areas of uncertainty to explore in PFS were also mentioned, predominantly by the methodologists. These included administration (for example, paperwork), randomisation pathways, sample size calculations and formalising consent processes.

> *012 (CTU Director): "Testing the feasibility of data collection, because some people sometimes, are being a bit ambitious, and actually it's just not possible to collect certain data in certain ways."*

> *020 (CTU Director/Funder): "Do you randomise at the start of the surgery […] How does that impact their behaviour in the surgery or do you randomise at the end of the operation […] Is it feasible to randomise at different time points, all those sorts of things. There were lots and lots of things. We did randomise at two different time points, each proved feasible to actually do it, but there were lots of unknowns so there was no way that study was ready to go to a trial, because it was just a string of question marks."*

> *010 (Methodologist/Funder): "Obviously issues around sample size calculations, consent processes are a big one, and it allows you to get more information than you could do from say, a hypothetical study."*

### 5.5.2 Challenges unique to trials of surgical interventions

As demonstrated above, all groups of participants regarded PFS as an opportunity to address uncertainties around the intervention. However, several specific uncertainties related to the intervention that were unique to the context of surgical trials were identified.

#### *Methodological challenges related to delivery of a surgical intervention*

Most methodologists identified how exploring uncertainties around the intervention in PFS was critical for surgical trials.

> *011 (Funder): "The area that's probably more neglected, is the feasibility of the intervention and again there's often a lack of appreciation of the complexity of interventions, lack of awareness of guidance in relation to evaluation of complex interventions, and often a naïve assumption that an intervention that's been used by an expert in a specialist centre can simply be taken off the shelf and implemented throughout the NHS. And you really need to know about implementation before you start a full-scale pragmatic evaluation."*

Whilst surgeons did not necessarily name the intervention as a key parameter to explore in PFS, some were aware of why surgical interventions specifically may be challenging, giving examples from their own experience of how the intervention had been explored in PFS.

> *023 (Surgeon): "So, in the feasibility study there, we had to actually look at stabilising the technique to see if we could actually reproduce it and train other people to do it. And after we ran a feasibility study on 20 patients, which we ran across five sites, after that we had a stable teaching rollout and a stable technique that we could then test."*

However, whilst some surgeons could give reasons and examples why surgical interventions were specifically challenging, they did not always make the link as to how and why uncertainties could and should be explored in PFS, tending to focus instead on needing to obtain results about the effectiveness of surgical interventions quickly.

> *003 (Surgeon): "Now, when we were asked whether we could do a randomised prospective study we said 'no we can't do a pilot study because we don't have enough numbers, it'll take 10 years to get the results' […] It will be unethical cos it was NICE approved in 2007 […] I don't think for that sort of thing, that pilot studies are any good."*

When elaborating on the issue of why surgical interventions were challenging, both methodologists and surgeons identified several unique challenges related to the development and delivery of surgical interventions. However, methodologists were seemingly much more aware than surgeons were, of the effect the unique nature of surgical interventions had on the design and conduct of trials and why PFS were particularly important to this context.

### *The surgical learning curve*

The influence of the learning curve on trial results was identified by methodologists as one of the intervention-specific challenges unique to surgery that is necessary to address before beginning a randomised evaluation in a main trial. Sometimes, this was linked to needing to employ statistical methods to adjust for a possible learning curve effect.

> *017 (CTU Director/Funder): "I think there's things to do with clustering surgeon effect, they talk about learning effect. Now that can be individual surgeons, that can be teams, hospital systems and I think some of those are true […] But the RCT's that have grown up from the pharmaceutical industry just don't have to grapple with those kind of things, so I think although surgery isn't unique, it's got some really special things that if you're not aware of, you're gonna get badly tripped up."*

*010 (Methodologist/Funder): "When you're designing a surgical trial you've got issues around clustering and around the learning curve and in a lot of surgical trials there is very little mention of how they're handling the learning curve or the clustering in the design. A limitation of pilot work would be that even if you were to do a pilot, and you were say looking at trying to get an estimate of clustering with the size of pilots, that would still be difficult to achieve."*

Some surgeons were also keenly aware of the potential impact of the learning curve on trial results and the need to adjust or allow for this, particularly in trials of new or novel interventions.

*003 (Surgeon): "Learning curve is crucial isn't it, especially for a new technique […] It is very difficult to analyse something where there's a learning curve especially against an established procedure […] I think the learning curve issue is absolutely crucial and especially with the things you're talking about introducing, new surgical techniques, new surgical instruments; it's crucial."*

### *Standardisation*

A second intervention-related challenge specific to surgical trials, was the lack of standardisation of many surgical procedures and practices. Methodologists perceived that the standardisation of certain elements of procedures between different surgeons and hospitals in order that they can be evaluated, as particularly important.

*011 (Funder): "Heterogeneity of surgical techniques. You may think that an operation is an operation and all surgeons do it the same way but when you actually look at what they're doing they could be doing it in different ways, different degrees of competency. How similar do these surgical techniques need to be before you can say 'well they're all pretty much doing the same thing?' How different do they have to be before you say; 'well actually these two surgeons are doing something completely different so we can't evaluate them together.'"*

Methodologists cited pragmatic trials as a method to overcome needing to standardise interventions explicitly, by allowing all variations of practice to be included. However, it was acknowledged that this might lead to subsequent problems with convincing the surgical community of the findings of such trials.

*012 (CTU Director): "I'd be really fascinated by that, to know whether we can standardise surgery because […] I don't know whether you end up doing a trial and the sites are really selected because you get the people who yes, who are more team players, or wise with that sort of thing, but then it becomes difficult to translate that back into practice exactly. So, you're almost better off doing a pragmatic trial, describing what the range of practices were, and then just… take the flak really, because there will be flak you know."*

Some surgeons and methodologists perceived the need for quality assurance of surgical interventions in trials as important. These participants felt ensuring the delivery of sufficiently high-quality surgical interventions within a trial was to

enable acceptance of the trial results (whether positive or negative) by the wider surgical community.

*005 (Surgeon): "The main difficulty with surgical research is, that in my opinion when you… say an operation doesn't work you need to demonstrate to the world that the operation was done by competent, experienced surgeons […] Is a trial explanatory or pragmatic, and so the question is does the operation work or does the operation work only in expert's hands?"*

*020 (CTU Director/Funder): "It's got a surgery arm and a non-surgery arm, because a previous trial was criticised with the quality of the surgery. So they want, therefore, to prove, well hopefully to show, that the quality of the surgery for whatever; if they see an effect that it's not attributable to poor surgical quality in some sense."*

### *Surgical expertise*

Surgeons identified studies comparing more than one surgical technique as potentially problematic because surgeons may not want to take part if they don't themselves use both the techniques in question, or because of issues with equipoise (discussed in further detail below). Several surgeons described difficulties with randomised PFS they had been involved with, where recruitment was affected and/or the pilot work had failed because surgeons would not or could not be randomised to performing one or other of the interventions.

*003 (Surgeon): "I agreed to take part in the study because I thought it was a valuable question, but I had trouble with the study because I didn't think the [procedures] were equivalent so we didn't enter very many so that's… And we didn't enter very many because we found reasons not to enter very many."*

Again, surgeons focused on the more practical elements of trialling surgical interventions and appeared to be less aware of the importance of PFS to explore these issues. However, some more methodologically experienced surgeons described surgery versus no surgery trials as particularly challenging in terms of the acceptability of randomisation, and the sorts of trials for which PFS were, therefore, absolutely necessary.

*004 (Surgeon): "There'll still be some surgical versus non-surgical trials where really to have a whole pilot on: 'Is it possible to randomise someone to surgery or no surgery', I think is worth it…It is incredibly difficult and a lot of training and investment is… I think external pilots are good for those difficult, surgery/no surgery trials."*

*008 (Surgeon/Funder): "I think a big element of any feasibility is can you get the surgeons to deliver the interventions in the context of a randomised trial, absolutely."*

### *Cultural challenges related to the intervention*

A fourth key challenge unique to trials of surgical interventions concerns the surgical culture itself. There was widespread acknowledgement among all groups of participants that the culture inherent to surgery presented difficulties to undertaking trials in this area. This was regarded as further justification for undertaking PFS.

> *012 (CTU Director): "Surgeons are such, they're really interesting people, and I think some are very obedient aren't they… they'll adapt their practice to fit in with the greater good, and others just don't see it that way."*

> *002 (Surgeon): "It's a mindset, isn't it? And surgeons aren't historically that good at that. We've just heard a talk on how useless our RCTs are […] But that doesn't mean to say that we shouldn't be trying to do them. I suspect that was what you were trying to say to him."*

All groups of participants frequently talked about challenges with community and individual equipoise. While these challenges are not unique to surgical trials, problems with equipoise appear to be compounded by issues relevant specifically to a surgical context.

> *003 (Surgeon): "I think we missed the chance because the two [devices] developed at about the same time; we should have trialled them at the beginning…people, they weren't in equipoise, they were really really wedded to one, because each one you had to fiddle around. It took years to get used to it, so that you could get it to work. As I say if the, if somebody said to me what operation do you think you spent most time fiddling around with, it would be those [devices] and I did spend years, and eventually I thought I've cracked it I really… And I'm sure other people with the other [device] found a whole load of ways to get theirs. So, we missed, we needed to do it at the beginning.'*

Whilst the surgeons understood that a lack of equipoise often affected the success of trials, they were less able to articulate reasons why, and tended to accept the inevitable consequences of lacking equipoise, such as 'impossible' trials or failed trials. The methodologists, however, were much clearer on the reasons for equipoise complexities amongst surgeons. Some discussed that, whilst true equipoise may be difficult to achieve, there is capacity for all surgeons to have at least some uncertainty sufficient enough, to enable them to commit to answering the research question.

> *019 (CTU Director/Funder): "If you get away from equipoise with the connotation of equals. If you're talking uncertainty, even surgeons can often admit to uncertainty. Some are better at it than others. It doesn't help that some of the old school surgeons have often been trained never to portray uncertainty but they… well you might not get equipoise, but I think in even surgeons, interventional cardiologists and people like that, you can get into the realm of uncertainty.'*

In addition, methodologists understood that the interventions being studied in surgical trials can be vastly different, making equipoise more challenging. The

methodologists illustrated how surgery is perceived to be a much more personal treatment for both the delivering surgeon and the receiving patient.

> *016 (Methodologist/Funder): "The really obvious one to me is trying to persuade both clinicians who are taking part in the study, or recruiting patients to the study and delivering the intervention in the study, so persuading them, and the patients to take part in the study when they could, by chance, they could receive one of two or more quite different treatments. So, I think that is a particular difficulty in surgery […] The difference between intervention and control can be more extreme, I think, than in other studies."*

With surgery having the potential to cure or kill, confidence and self-belief that the treatments surgeons deliver are predominantly beneficial rather than harmful, is a natural expectation and requirement of a surgeon's role. Methodologists understood that equipoise for surgeons can, therefore, be at direct odds with this well-developed confidence and belief such that being in equipoise, requires challenging not only personal beliefs, but also doing so very publicly to patients, peers and the wider trials community.

> *020 (CTU Director/Funder): "I think possibly in all areas of medicine but particularly in surgery, surgeons have to make decisions, and they have to make decisions quickly. Very often, they've got people's lives in their hand. They may be doing potentially life threatening surgery […] They have to make decisions, and they have to make decisions quickly […] The whole idea of equipoise; equipoise is about uncertainty, equipoise is about not knowing and I think that is not what surgeons typically are about. They have to have faith in their… I mean I'm surmising I don't know, but I think that they have to have faith and belief in their ability because of the nature of what they do."*

> *013 (Methodologist/Funder): "It's partly to do with the fact that surgeons are people who have to have a belief that the new intervention that they're testing is better. I think there is less equipoise […] They wouldn't be taking the knife to someone, unless you felt sure you were doing something good for them. It's a different psychological phenotype of clinician, I think, and obviously they perform a very important task, but I think they're slightly different from people who dispense drugs, which can be stopped quickly, and suddenly, if something goes wrong."*

Methodologists described the centrality of the surgeon to delivering the intervention and the irreversibility of many surgical interventions as another reason why conveying uncertainty is more difficult in surgical trials. Many methodologists also perceived the unique nature of surgery as a craft specialty, with potentially incapacitating and often-irreversible invasive interventions being delivered by the hands of people who have undergone thousands of dedicated hours of training, all further explaining why surgery is a more personal intervention.

> *015 (Methodologist/Funder): "One of the big differences with surgical trials, I think, is that the surgeon is such an important part in intervention delivery and it's often irreversible. So, you can stop taking the pill but once you've removed something or changed something or whatever,*

*it's difficult or impossible to go back. So, I think there's something there about the irreversibility and the absolute centrality of the surgeon to the intervention, which is different for some other type of interventions."*

Surgeons, whilst perhaps less clear on reasons for lacking equipoise, did discuss how elements of the 'surgical personality', such as being expected to demonstrate leadership and innovating, may negatively influence trials. For example, surgeons perceived that the desire to make progress quickly in a treatment area may mean the chance to develop or assess an intervention in the context of formal research is seen as a distraction.

*003 (Surgeon): "Not too late to do a trial probably, because I think it should be done, but probably too late to convince the zealots to do a trial. You've given them… it's a bit like running after somebody if you've given too much of a head start. You'll never catch you'll never get them. You have to grab them at the beginning and go 'whoa, whoa, whoa don't run off'…"*

In addition, some surgeons perceived that driving progress forward too quickly without adequate thought and involvement of trials expertise, might result in trials running into difficulty because they have been designed around specific surgeon beliefs.

*006 (Surgeon): "The zealots amongst the surgeons have worked their way round by persuasion, muscle, conviction and all the usual things we do when we want to win our case into an explanatory mechanistic study and only if you do it the right way […] I think they may have evolved a trial to prove their own surgical machismo, and I would like to evolve a trial to see which benefits patients… They just regard me as a traitor […] You don't want to lose too many friends, but who wants friends like that who can't see the purpose of the trial is to get health gains for individual patients, patients in general, and for society."*

### *An evolving surgical research culture*

Methodologists described how the historical lack of emphasis on surgical research has resulted in surgical trials not being embedded in routine practice. This was observed to contrast with other fields of research, such as oncology pharmaceutical trials, for example. A relative lack of regulation in the way that new surgical procedures can be introduced into clinical practice was regarded as a contributing factor.

*017 (CTU Director/Funder): "I think one of the issues is that there has been much less of a culture of research in surgery; there's some good reasons for that. Trials in medicine are a highly regulated environment and so there's a whole infrastructure of how you do a study that's evolved because of that. Whereas in surgery, there is not. There is then intrinsic problems that genuinely, when do you study something. So, I think there are challenges in doing any kind of studies in surgery. I mean, there are some extremely good groups but there's an awful lot where it's a foreign language - they don't have the culture, the training and so understanding about different sorts of studies like pilot and feasibility, might be hard to explain. But that also makes it an area where pilot and feasibility studies are probably needed for that kind of reason."*

Both surgeons and methodologists talked about observing a cultural shift in recent years in the standing of surgical trials within the surgical community. This was noted to be illustrated with the creation of certain surgical research initiatives, such as the RCSEng Surgical Trials Initiative [196] and the nationwide surgical trainee collaboratives, which have grown a network of surgeons interested in conducting trials.

> *023 (Surgeon): "So I'm not saying that everything is roses, and everything is fine, but there has been a change in attitude in the last five years that I don't think anyone would have believed would have occurred. And everyone knows what we do, and everyone knows what we're trying to do and it's almost, it's on a par with LBG riots, in terms of, you can't speak out against it. So, it's a good place we're in, and now it's up to us to make sure that we maximise the patient benefits on it […] Your trainee collaboratives have been fantastic and inspirational to all of us."*

> *015 (Methodologist/Funder): "I think there is a slow realisation that not only is it possible to do trials in surgery but that it ought to be done. It's unacceptable to not rigorously evaluate something just because it's surgical, so I think there's a realisation from the surgeons I talk with […] I think it's more acceptable, more desirable to be involved with trials and doing trials in surgery, than it is perhaps in an older generation. The culture is different, so I think it… things are changing and I think again, the UK is a good place to be for those sorts of change. I think we're ahead of the curve."*

Some surgeons described how trainee collaboratives have helped to deliver PFS more quickly and efficiently, with less need for infrastructure investment in terms of trial personnel, recognising again the practical challenges of PFS.

> *021 (Surgeon): "The trouble is, it becomes tricky as a researcher unless you, if you're actually paying for a research team, if you're buying a research team to put together and hiring people that's tricky if they're… you don't know whether it's going to go beyond one year […] The trial that [name of professor] was doing with the […] because it's run by trainees, run by a trainee so it's fantastic…it doesn't matter whether it stops or starts or keeps going, because there's nobody who needs to be laid off so, incredibly attractive of that sort of trial."*

However, some methodologists were more sceptical about the current evidence for trainee collaboratives to deliver projects successfully, citing that not all projects initiated by these groups had succeeded.

> *011 (Funder): "I know there have been some very successful trials delivered in surgery using trainee networks. But equally I think they're having some unsuccessful ones, or some that have struggled. There's been over reliance on the surgical network and that does make it quite difficult for funders to judge."*

### 5.5.3 Research infrastructure barriers

Whilst surgical interventions were felt to be specifically challenging, further necessitating exploration within PFS before a main trial, participants also

identified barriers within the current research infrastructure that further made surgical PFS challenging to instigate and complete.

***Inaccessible methodological expertise***

Methodology input to the design and conduct of PFS was perceived as important by most methodologists, funders and some more methodologically experienced surgeons.

> *015 (Methodologist/Funder): "We'd like to see methods expertise in it. So a good way to get rejected is to be doing a pilot study, and in no way is there a statistician, mention of a statistician or a methods person, and often it's very clear in the proposal that they haven't, and they end up getting rejected. The one this morning, it was really obvious that what they need is methods input."*

> *022 (Surgeon/Funder): "The surgeons were talking about just doing a little bit of statistical analysis themselves on the data, he said: 'I'll tell you what, you can do some statistical analysis and I'll go do the operation', and he said: 'I think I know more about surgery and I can probably do a better fistula operation than you could do the stats. What do you think?' So, you could say if you are going to run a trial and you aren't qualified to do it, should you really be doing it. We wouldn't let the trialists come and run labour ward or… that sounds ridiculous to us, but we can have a go at it and it's all right? Expertise is important."*

The surgeons, however, recognised significant barriers in being able to access Clinical Trials Units (CTUs) and methodological expertise. Surgeons described these barriers firstly as inaccessibility if you were not employed within and/or collaboratively working directly with a CTU already, and secondly the expense of CTU input, which is not usually affordable within a PFS budget.

> *008 (Surgeon/Funder): "I mean working within a trials unit to run these sorts of trials, you think is absolutely essential really, you can't imagine working outside of it. I mean I can't, but I know most, most clinicians do work outside and they have to. There's a huge barrier between getting access to the trials units, getting CTUs to talk to you. Even well-established people are having trouble engaging with their CTU."*

> *007 (Surgeon): "The major stumbling block is the fact that a lot of bodies require you to have a clinical trials unit, and the clinical trials units are often too expensive."*

> *015 (Methodologist/Funder): "We can't pay for a CTU to run it and nor will that ever be the case."*

In addition, whilst methodologists and surgeons recognised that the funding package for PFS does not generally allow for full CTU involvement, some methodologists stated that it could be challenging to have only a peripheral role with PFS because they so heavily centre on potential uncertainties.

> *020 (CTU Director/Funder): "I've done some advisory stuff that has been more peripheral to some studies, and to a CTU, it's not a preferred model. I suppose because we more naturally like*

*to engage with the team throughout the study, and I think part of the issue with pilot and feasibility is that they do throw up, well every trial throws up things that you don't expect every trial presents challenges, even the ones where you think you know what you're doing."*

A few methodologists were however aware that they, and the available guidance on PFS, may be relatively inaccessible to those working outside of CTUs, and that CTUs producing clearer guidance and becoming more approachable to potential surgeons needing advice, would be helpful.

*015 (Methodologist/Funder): "I think there's a bit of a failure on our part, meaning methods people like me, to translate for want of a better word, our stuff into a format that people who really have got better things to do, to use. They've got better things to do, than read through half a dozen papers. What they really want is to distil the key things that they really, really need, so they can build it into their idea."*

***Inefficient research infrastructure***

Participants unanimously agreed that the overall length of time it took to seek funding and to complete PFS and main trials was a significant barrier to engaging with PFS at the start.

*023 (Surgeon): "So if you do it in the linear way, the way you're supposed to, we'd all be dead before you finished the main trial, which again comes back to my point about the present structure, is just too inefficient."*

*010 (Methodologist/Funder): "I think one of the issues is how long it takes. So, somebody comes up with the question and then it can still take, you're still looking at a good 18 months before; you've months to write, draft an application, send it in, go through a two stage process, get the funding, get the contract and get in a position to even recruit a patient. So, in some areas of need, you're just not willing to wait for that kind of length of time."*

The pressure to produce definitive results was perceived as particularly pertinent in surgical studies, where the clinical landscape can move relatively quickly, and where different or new techniques can be adopted before trials are completed. The research question can therefore, become obsolete during the time it takes to conduct the PFS and the subsequent main trial. In a sense PFS were perceived by all as potentially part of the problem, rather than the solution, because they prolong the time it takes to get definitive results.

*011 (Funder): "The concern with surgical trials is that you're going to start evaluating something and it's going to change while you're evaluating it. So, the learning curve is obviously the surgeons learning the technique, and that's an issue, but then there's the technology itself that might move on while they're doing the trial."*

*009 (CTU Director): "In your world of surgical trials, this is dreadful, because whatever you find in a surgical trial, by the time you've set up and run a full HTA trial over five years from inception to paper, the surgical technique has moved on five years. So, what you're saying is*

*well what we did five years ago did or didn't work, but what we do now may or not work, because we're doing something different."*

*012 (CTU Director): "Things will move on, and it's pointless to do a whole bunch of small pilot or feasibility studies, and then actually the question has moved on by the time we've worked out whether you can (laugh)…"*

External standalone PFS, were regarded as particularly problematic as they extend the time for answering the research question to an even greater extent because of the hiatus between completing the PFS and applying and getting funding for the main trial.

*015 (Methodologist/Funder): "You could get up to £300,000 for a feasibility study and after that you're looking at NIHR money, so if you wanted £2 million, the only place you have to go is NIHR. HTA really, which means there's a delay then of at least a year, and probably more than that. So, it really stretches out the development of that trial."*

### 5.5.4 Funding barriers

Both surgeons and methodologists perceived that difficulties with funding were a major barrier to completing PFS. Funders explained their rationale for careful funding decisions around PFS, perceiving that the importance of any surgical research question is determined through contextually considering its importance to patients, to surgeons, to the NHS and, through consideration of these groups, to funders. Funders felt that any pre-trial work must therefore contribute to ultimately answering a question (through a definitive trial) deemed to be important for all stakeholder groups.

*026 (Methodologist/Funder): "I mean as a funder, we want the main question to have been considered before the feasibility work is undertaken. So, unless it's a question of real importance that, we want to answer, then the feasibility work shouldn't be funded. So, I suppose we want people to engage with the main funder so, up front, look at the package, rather than individual feasibility."*

*011 (Funder): "Deciding whether a trial is worthwhile involves judging value for money and that value for money judgement has to be made from the point of view of the health service and the funder. There are one million-pound questions that are worth one million pounds, and there are five million-pound questions, or even ten million pound questions and we will look at the question, we will judge what it is. But we then need to know how much will a definitive trial cost. Because if a definitive trial is going to cost ten million pounds, and it's only a one-million-pound question, then there's no point funding the standalone pilot and, again, I think this is often underappreciated."*

Funders further emphasised the limited money available overall to fund clinical research, reinforcing why the judged importance of the research question was key in deciding which PFS are funded.

*010 (Methodologist/Funder): "My experience on the board would probably say that cost is a significant issue… it doesn't really matter whether it's a pilot, or a main trial, whatever, it's all about the research question."*

*015 (Methodologist/Funder): "It's definitely the case that funding committees look at the headline figure of how much does this thing cost, and if you are doing a feasibility and it's £300,000, you better have a really compelling case."*

Funders perceived, therefore, that the main reason surgical PFS do not get funded is that the overall question is not considered valuable enough to warrant significant funding investment.

*011 (Funder): "I think if feasibility trials aren't getting funded, it's probably because they haven't passed that value for money test […] We have to be very parsimonious in what we fund and there's lots of good questions that we simply can't fund because it'll be too expensive to answer them."*

In contrast, surgeons perceived a lack of surgical representation on funding panels, and competition with translational science and experienced research teams for funding, as significant barriers to fair funding opportunities.

*006 (Surgeon): "You'll have one token surgeon with a grant giving body, who I should think, often, is not terribly diplomatic or very experienced in that sort of thing and it's so competitive. Not just at the grant giving money for trials stage, but at the next level up where we're going to spend our money, and translational studies and other types of sexy sounding personalised medicines and the humdrum randomised trial is hard to do. If they put their money into knock out mice, teams who know what they're up to and have got to mould the work, crank that handle, they get the money and out comes a 'Nature' paper and we're… it's very hard for us to compete with that."*

*007 (Surgeon): "In terms of funding there are all sorts of funding streams that I've used in the past, including industry, and the major stumbling block is the disparity between what the funding bodies actually tell you, how they're all interested in surgical research and how the minority of surgical research gets funded, and therefore, we're all mobilising ourselves to make sure that that's reversed. The fact that that's not the case, and they're completely disinterested in surgical research… I can say that safely across the board. I think it's fair to say that a lot of charities are completely disinterested in anything that's of clinical value, or that involves surgical research."*

Surgeons also perceived other barriers to funding related to the lack of regulatory requirement for formal evaluation of new surgical procedures and surgical devices. This was observed to have led to a lack of research infrastructure within industry, resulting in fewer avenues for funding surgical research, when compared to for example, pharmaceutical research.

*008 (Surgeon/Funder): "Speaking candidly, the big companies have got no interest in… in fact it's almost a disincentive to do head to head comparisons of their technology against whatever other interventions are out there, and because there was no requirement for them to do so and they never actually developed… what you discover is they've got no infrastructure to do that."*

*007 (Surgeon): "The difference is that the pharmaceutical industry, not only has more funding, but has the requirement to carry out the work, whereas the device industry has got less money, but has plenty of money, but has no requirement to carry out the work. There's a lack of a regulatory requirement."*

### 5.5.5 Challenges of reporting & interpreting pilot and feasibility studies

Several issues with the interpretation and reporting of PFS were recognised, predominantly by methodologist participants.

***Underpowered RCTs masquerading as PFS***

Methodologists highlighted the ongoing problem of underpowered RCTs being published as PFS in the surgical literature.

*017 (Methodologist/Funder): "My pet hate is something described as a pilot study. When you read it, all of the objectives and everything else reads as if it's a definitive study but with a massive, thumping great effect size and actually it's nothing of the sort. It means, we wanted to do the full study but we couldn't afford it, or couldn't find enough patients so we've done something, we've called it pilot, and yes it's neither fish nor fowl."*

*014 (Statistician): "They seem to be getting away with doing mini randomised controlled trials and, as I've described, they shouldn't really be doing that."*

Both surgeons and methodologists with editorial roles recognised the practice of publishing underpowered RCTs masquerading as PFS to be a phenomenon particularly common in surgery. This was understood to contribute to the widespread misunderstanding and devaluation of PFS.

*010 (Methodologist/Funder): "I think, that if you're talking about an external pilot I think that there is, or at least there genuinely has been, the perception that these are small version clinical trials that you can do without a sample size calculation and can get passed onto a registrar, or somebody in need of a project."*

*021 (Surgeon): "I would say between 5 and 10% of the RCTs we see are badged as pilot trials or feasibility studies. We reject the vast majority of them […] The worst ones are what I said to you, they're underpowered RCTs and what I usually do is write back to them and say; 'if you do the full RCT we'd like to see that […] but we don't want a pilot trial thank you' […] I'm sure they get it published somewhere else, yes, but to my knowledge, no one has ever come back with the full trial, after we've rejected their pilot."*

In contrast, participants gave examples of the reverse situation, where pilot studies were performed, but then badged as RCTs in an attempt to achieve publication.

*008 (Surgeon/Funder): "With my journal editing hat on, I spend a lot of the time trying to persuade investigation groups so, they've actually done a pilot study and please report it as such. Even though they try and sell it to you as a definitive project with 20 patients in each group for a complex intervention."*

Some methodologists felt that this misunderstanding was cyclical, in that journal editors only want to publish definitive results, meaning study teams are tempted to offer these to achieve publication.

> *025 (CTU Director/Funder): "There's a lot of confusion about what people should be doing in these studies, and I think a lot of that comes from misunderstandings of journal editors and reviewers and trying to push people in a certain direction."*

Some surgeons with editorial roles also felt that the publication of underpowered RCTs was acceptable practice, rationalising that such studies could be used in a meta-analysis.

> *021 (Surgeon): "Sometimes we do get randomised trials that are underpowered cos a pilot randomised trial goes seamlessly into a grossly underpowered randomised trial, and I don't feel so bad about publishing those because at least if they're well done, and well written up, they can go into a meta-analysis… those sort of trials are never wasted, because they can then be used in meta-analyses. As long as they're well described, I think that's ok."*

Views amongst funders towards the appropriateness of PFS objectives written in funding applications varied somewhat. Some funders explained that funding applications were commonly written with appropriate feasibility objectives but devoted the entire analysis section to planned hypothesis testing of effectiveness.

> *015 (Methodologist/Funder): "When it comes to the analysis what they're saying is: 'We want to test out our procedures, testing recruitment, willingness to be randomised, willingness to fill in the data blahblahblah', and all of the analysis section is about some multi-level logistic modelling that they're doing on the outcome data. I actually saw one this morning…It's in a grant application. So, we see a lot of that and then we bang our heads together and think 'oh it's a shame', but they're not telling us anything about what they're going to do on the things that the whole pilot is really being done for. They can do some outcome analysis if you're looking for promise too, but to have that entire focus; very, very common."*

Other funders, however, emphasised that such applications were now in the minority, with underpowered RCTs more commonly observed only in the published literature.

> *017 (Methodologist/Funder): "At least within [name of funder] I'm not saying every last nuance is worked out but there is now a good understanding of what they do, what they're for and so on. I think the one's, my pet hates, are the ones if you were doing a literature review it would come up, but actually it's not true, it's some sort of relatively kind of lone researcher (type of study)."*

### *A place for publication of PFS*

Methodologists unanimously agreed that PFS should be published and/or the results made publicly available, even if the results show that a main trial is not feasible. Publication of PFS was regarded as important for preventing the research

being repeated unnecessarily thereby preventing research waste, for allowing others to learn from previous work, and for developing ideas for future research.

> *019 (CTU Director/Funder): "They absolutely should be published. The two reasons are, one it stops somebody else wasting their time unnecessarily, repeating what you've done and the second one is, if someone can spot a way in which they can improve on what you've done, it gives them an opportunity to build on it, so they really should be published."*

Methodologists also felt that if the work had been publicly funded, publication was obligatory from a standpoint of good research practice.

> *017 (CTU Director/Funder): "They've been publicly funded so why wouldn't they [be published] …that's the slightly bureaucratic reason for doing it, it's important but the purpose behind it is that other people can learn from it."*

However, most methodologists noted that despite many PFS being funded from the public purse, public reporting of PFS is not a requirement of all funders.

> *016 (Methodologist/Funder): "I suppose there's probably a cost argument to it, because I'm sure that the HTA's staff incur a cost for the editorial process so there's the back and forth and there's obviously the big unpaid reviewing that goes on of reports as well for HTA. So, there needs to be a whole mechanism behind it to make peer reviewed [articles] publicly available from the reports, for RfPB in the same way as the HTA. But I would have thought that that was worth investing in because… not everyone will publish their feasibility study in 'Trials' or somewhere else in 'BMC' or 'BMJ Open' or something like that. Some just will never see the light of day."*

Some funders defended this position, stating that the infrastructure for the smaller funding streams did not lend itself to peer-review publication, despite agreeing that there should be a principle of publication.

> *017 (CTU Director/Funder): "The HTA is a large well-funded programme, so there was a requirement if we're gonna spend a million, two million on a trial, it jolly well ought to be out there and that's it, that's where publishing is brilliant. But RfPB was always was light touch… but in general you want at least some kind of report somewhere…"*

In addition to funders not always requiring publication, methodologists also noted challenges with journal editors and academic institutions not prioritising or valuing PFS, as they perceived them to have a low impact overall in not offering definitive practice-changing results.

> *009 (CTU Director/Funder): "And regrettably the university won't see this as being an important paper because, nobody's going to see it as being three or four star […] it ends up in a low key journal, they look at it, you haven't collected any data, you haven't got really hard outcomes other than saying it can't be done or it can be done, we're not interested, where's the main data?"*

Some methodologists further illustrated how the value of PFS had been undermined because of many funded projects not progressing beyond the pilot phase historically and subsequently not being reported. These PFS were felt by

methodologists to be widely perceived as standalone projects with little purpose or value, which for inexperienced trialists, may diminish reasons for performing them.

> *019 (CTU Director): I do think there are probably some people out there who spend their career doing pilots and don't move forward into doing the definitive trial which is harder'*

However, some methodologists also lamented that it was difficult for trial teams to be motivated to write up the work for publication, particularly if it was 'negative' in that it demonstrated a definitive study was not possible.

> *009 (CTU Director): 'Well actually the biggest problem, biggest bias in getting this stuff published is getting the team to get their backsides in gear to write the paper. If you do the feasibility study and you demonstrate, it's not feasible, that's a really important finding. And there's a whole lot of experience to go into that. You do the pilot study and you decide you can't do it that should be written up. But they lose enthusiasm."*

> *019 (CTU Director): "It's less exciting, I guess, to write up a study where the pilot showed something couldn't be done. I guess some people might view it as embarrassing I wouldn't […] we saved the HTA programme 1½ million pounds, which they can now spend on something else, I can see it's slightly embarrassing, but I don't see there's any good reason not to report it."*

In addition, several methodologists perceived that whilst PFS contained much of the methodologically important material transferable to other studies, if neither study teams, journal editors or universities valued this, the cycle of misunderstanding and devaluation would continue.

> *009 (CTU Director): "All the science, all the clever stuff is in the protocol paper, which counts for nothing. And actually, in the feasibility work, because it's all the positive feasibility work that got you to the point that you could do the main study. It's where all the clever stuff is. The wonderfully concise paper in 'The Lancet' says, we tested it, we found a load of people with whatever it is they've got, we gave them whatever these two interventions were and either it worked, or it didn't. I mean there's nothing scientific about that."*

Considering all these barriers to the publication of PFS, several methodologists welcomed and cited the introduction of *The Journal of Pilot and Feasibility Studies* [133] stating how important it was for there to be a place where PFS were accepted and made widely available in the published literature

> *020 (CTU Director/Funder): "The new journal of pilot and feasibility studies is really, really, really important because it gives an opportunity, an outlet to get these things published and they may well be very highly cited ultimately, because they're telling us things that don't come out through the big all singing, all dancing study."*

## 5.6 C) Solutions to optimise the design and conduct of pilot and feasibility studies for surgical trials

As detailed above, participants identified many challenges around the design, conduct and completion of PFS. Adding to problems with variable understanding of the purpose and scope of PFS amongst the wider surgical community, these challenges were perceived to result in confusion, undervaluation, and under-utilisation of PFS in surgery. Several solutions to optimise future PFS for surgical trials were, however, proposed.

### 5.6.1 Education

Several elements of education were perceived by participants as important solutions to improving PFS. These included; the education of surgeons more broadly in trials methodology (both undergraduate and postgraduate research training); the development of accessible guidance for surgeons; complementary (not contrasting) endorsed guidance from funding and other regulatory bodies; and finally 'on-the-job' education through surgeons working collaboratively with methodologists within trial teams.

#### 5.6.1.1 Education of surgeons in trial methodology

The education of surgeons in trials methodology was proposed by methodologists as necessary to improving understanding of the nature and value of PFS amongst the surgical community.

> *025 (CTU Director/Funder): "I think it's just by slow drip, drip process of taking it out there to people."*

> *020 (CTU Director/Funder): "Because of the nature of their work being quick decision making and trials are very pedantic and take a long time to set up, so I think there's an element of frustration sometimes in appreciating all of the regulation around trials."*

Methodologists also felt that, whilst the education of surgeons was important to improve understanding, journal editors and reviewers might also benefit from such education to improve the perceived value of PFS by breaking the cycle of publishing small underpowered RCTs labelled as PFS in surgical journals.

> *014 (Methodologist): "Show them some of these small mini-randomised controlled trials that they do, show them the bad examples, not saying that they are bad examples, and point out why*

*they shouldn't have just concluded there's no difference between these two techniques because they didn't see a significant p value… If they still get accepted in their journals then, they think they're okay."*

Most participants believed that the RCSEng Surgical Trials Initiative [196] and the nationwide surgical trainee collaboratives had made an important contribution to improving the surgical research infrastructure through helping to raise the profile of surgical research and improving surgeon participation in trials. However, uncertainty about the genuine ability of these groups to deliver trials was highlighted by some funders, who felt methodological research into how and when surgical trainee collaboratives are successful, would be useful.

*011 (Funder): "Sometimes we get proposals for surgical trials where they seem to be quite light on the clinical trials unit support, heavily reliant on the [trainee] network and we don't know well is this gonna be one where the network is enthusiastic and it takes off and it really works or, is it gonna be one that falls through. Whereas, if you've got a professional CTRU clinical trials unit then you can say, 'well ok, they should be able to deliver this', but it does add another layer of uncertainty to things. And working out how and why and when it works would be very interesting."*

#### 5.6.1.2 Accessible guidance

Both surgeons and methodologists considered the available theoretical and/or conceptual literature as poorly understood and inaccessible to the wider surgical community. It was proposed that accessible guidance, which serves to operationalise the key concepts into practical recommendations for undertaking PFS and trials, would be beneficial.

*021 (Surgeon): "It's one of the areas that I'm not aware of any guidance, so some guidance would be sensible, just sensible guidance along the lines we've been talking about what to look out for, to highlight the areas of concern, how to set it up, who to talk to and when, I think it would be valuable."*

*015 (Methodologist/Funder): "My sense is that the solution would be better guidance from, well better guidance full stop, which is very clear and very practical, better guidance from the funders […] I think we could give much clearer guidance and actually it keeps coming up and we… nobody ever has the time to do it […] so there might be different versions of it so if you think of it like layers of guidance."*

Methodologists perceived that developing guidance to transcend the interface between surgeons and methodologists was, therefore, a key solution to improve PFS in surgery.

*025 (CTU Director/Funder): "I think that one of the key things is providing this kind of guidance which sits at the interface between the real methodology and the clinician and that I think is the important thing because there is stuff out there but some of it is a bit too specific or a bit too technical […] so I think there's a huge amount to do on that interface between clinicians and methodologists."*

Methodologists felt that in addition to distilling and operationalising the most important messages, new guidance may also signpost surgeons to some of the conceptual and theoretical guidance documents, such as the CONSORT reporting framework for PFS [79], in order to complement and enhance surgeon's understanding.

> *020 (CTU Director/Funder): "If that guidance only brings forward the fact, that puts the reporting into a framework for people to look at when they're designing [PFS], they're more likely to look at it in that sense. I'm not sure clinicians will automatically go to the reporting literature when thinking about designing something."*

> *014: "I don't think the reporting guidelines have got through to people in other countries yet, so until we point out; 'you might find it helpful to look at this', they may not have come across them."*

In addition, methodologists perceived it to be important, that any guidance developed, would advertise and encourage the use of existing methodological resources, for example, the NIHR Research and Design Service.

> *019 (CTU Director/Funder): "I think the best guidance I ever give the surgeons, so surgeons who are comparatively research naïve and they want to do research, is to go and talk to their research design service, if they're in England anyway. So, I think that's a great resource and I think if you need a bit of guidance at this end of the project, you shouldn't be doing it on your own."*

All participants agreed that any new guidance, should be simple, including some key recommendations, and encourage consideration of individual study uncertainties, rather than didactically listing certain issues to explore in PFS.

> *022 (Surgeon/Funder): "As long as they're not didactic because they need flexibility around different kinds of trials, but yes it would be very useful."*

> *017 (CTU Director/Funder): "I think something in which you say the principles and why we do them, what's in it for the investigators, what's in it for the funders, some things to think about have you got your questions straight, does it need to be internal or is it standalone to answer them […] something that gives us a sort of framework. I mean they're all very different so that's why I wouldn't want it say you must have this; you must have that."*

Methodologists also felt that examples of how PFS have been designed to inform main trials in surgery would assist in informing better understanding about PFS design amongst surgeons.

> *025 (CTU Director/Funder): "I think people need to start engaging with what's out there, and we need to find some good examples of how that's helped people."*

> *017 (CTU Director/Funder): "In your area from surgery, saying this is where a feasibility study was done, this is how it informed the main study, this is how a pilot was done and either it stopped or it went on, so some concrete examples so they can get their heads round what the variety is. A lot of people learn by example anyway so I think that's particularly true in areas*

*like surgery […] a high-level check list that should apply to everything but with concrete examples and making it clear that it's not a kind of a recipe, it's a thing to think about."*

*020 (CTU Director/Funder): "I'm thinking you know really focus your examples on the challenges and feasibility around surgery."*

Some methodologists felt that layers of guidance, starting with key points that always need to be considered for PFS, and developing this with layers of complexity for different end users would be helpful. This approach was analogised to a detailed appliance manual with a point of reference page at the front signposting the absolute requirements for safety before use.

*015 (Methodologist/Funder): "So there might be different versions of it, so if you think of layers of guidance so something which is widely readable, if you like, in a sense that this is meant for the widest possible audience and gets across key messages so if you see nothing else or take onboard nothing else these are the things you really need to think about. And then for those who are more interested, whose job it is perhaps to be working with trial teams to do this on a day to day basis across many trials, something a bit more detailed I think would be appropriate and there might be a third layer where it's very very detailed and this is what a trial methodologist or a CTU ought to be looking at."*

#### 5.6.1.3 Content of accessible guidance for surgeons

Methodologists and surgeons illustrated some items they felt to be fundamental for inclusion in a list of key recommendations within guidance for surgical PFS.

##### *Considering specific uncertainties of the research question*

Methodologists were very clear throughout the interviews that it was paramount to consider all uncertainties relating to the research question in PFS, and that those uncertainties will be different for different studies. A criticism of some current guidance documents was that offering lists of issues to consider as examples, led researchers to believe they are the only items to explore in PFS.

*026 (CTU Director/Funder): "That's the danger of saying; these are always the things that you should look at in feasibility, that people stop thinking. Actually, you need to think about the context, you need to think about the question that you're wanting to answer, the big question. And then what is it that you need to be able to show to convince people that that main question is answerable."*

As addressed, within the previous theme (Section 5.5 Challenges of PFS for surgical trials), considering uncertainties around the intervention was perceived as fundamental for surgical trials. Exploring the uncertainty around surgical interventions would need emphasising in any guidance, given that surgeons were less explicit about this, tending to focus on recruitment.

### *Selecting a PFS design*

Methodologists felt guidance on what type of PFS to perform and in which circumstances, would be particularly useful. It was perceived that the type of PFS would need to be considered in the context of the research question, the types and number of uncertainties and the experience of the trial team.

> *018 (CTU Director): "Some kind of uncertainty index […] a short questionnaire, ten items or fewer, that people could answer, possibly a scale on each question and I could see you'd have like three cut-points, one where you probably didn't need a pilot, one where you probably needed a pilot, one where you probably needed a feasibility, of increasing uncertainty. I'm sure if you did something like that it would be well cited because I think a lot of people, well certainly applications for pilot/feasibility studies, people would probably use it for that, to justify why they wanted a pilot or didn't want a pilot for that matter."*

In contrast, and illustrating the urgency felt by surgeons to move on quickly to a main trial, surgeons perceived that considering when you didn't need a PFS was particularly useful to include in guidance on PFS.

> *021 (Surgeon): "So maybe the most important part of your piece of advice would be what is the information you need before you start a trial, to mean that you don't need to do a pilot feasibility study, that might be the way around looking at it."*

### *Early planning for the main trial*

Methodologists and surgeons perceived that having clear intentions to undertake a main trial and, therefore, justifying the importance of the research question that the main trial will ultimately answer, was fundamental to the design and conduct of PFS.

> *019 (CTU Director/Funder): "So you must know what the main trial you want to do looks like. You might not have the detail […] but I think you need to write down, and certainly if you're applying for funding, your funder will want to see an idea of what you're working towards, and that helps you both justify and think through what the unknowns are that's stopping you doing it right now."*

> *010 (Methodologist/Funder): "I think that you do have to have the main trial in sight because, if you don't, why are you doing the pilot anyway […] it's a waste of resources."*

> *002 (Surgeon): "You've got to have a big picture first, and then decide do you need a pilot study. You can't do one without the other can you really […] you've got to know where you want to be and where you're going first."*

> *023 (Surgeon/Funder): "I don't think you should be running a pilot or feasibility unless you can see the full practice- changing trial at the end of it. I mean what are you doing, you're just playing around. […] it's irresponsible of anyone to be funding you if you can't demonstrate what the whole game is."*

### *Handling clinical outcomes measured in PFS*

The ongoing practices of publishing underpowered RCTs labelled as PFS, and of undertaking formal hypothesis testing on PFS data were highlighted in Section 5.5.5 (Challenges of reporting and interpreting PFS). Methodologists perceived it was important, therefore, for any guidance to explain and demonstrate the issues around outcome analysis for PFS.

*025 (CTU Director/Funder): "I think what you do with clinical outcomes is another area because a lot of people want to analyse those just to say we're doing a main trial."*

#### 5.6.1.4 Complementary guidance and endorsement from funding and regulatory bodies

As explained in section 5.4 (Differential understanding of PFS), funder guidance was perceived as limited in offering operationalised detail for designing and conducting PFS in practice. Whilst most funders offer some guidance on their websites, this is mainly limited to brief definitions of PFS, which often differ between funders. Methodologists and surgeons felt, therefore, that consistent guidance for designing and conducting PFS across funders and other regulating bodies, such as the Royal College of Surgeons, would be key to encourage better practice.

*015 (Methodologist/Funder): "I think it would be important for funders to effectively endorse it […] I think that would be really nice and if the funders put it up, the funders clearly have the stick where the other organisations don't. They hold the money, and so if the funders have asked me to do it I think you're much more likely to do it and if you don't do it, it makes it easier to reject it."*

*007: "I think it's important for the… you can try and preach to the individuals but ultimately, it's down to the associations and to the colleges to try to come out with guidance of what is acceptable behaviour and what isn't."*

In addition, participants felt that endorsement of guidance for the design and conduct of PFS from all the major funders and relevant journals would also be beneficial.

*022 (Surgeon/Funder): "I think the major funders are going to be paying for it so it would be good if, not just NIHR, but all the major funders will endorse it. Say if you are gonna run a feasibility or a pilot study, this is how we'd expect you to go about it and of course it would be nice if some of the journals…"*

However, some methodologists suggested that whilst formal endorsement of guidance would be advantageous, it would be unusual for this to come from funders.

*016 (Methodologist/Funder): "It would be quite unusual wouldn't it to endorse guidance for how to go about designing a pilot or feasibility study […] I think that it should, endorsement*

*would help because then it would hopefully make people take notice of it, but I'm not sure where's the best place to actually get that from."*

*027 (Funder): "I think there's always a case for something like that [endorsed guidance], but whether that's the NIHR's business, certainly it's the Research Design Services', they should have it all nailed down as to what kind of [PFS to perform]. My experience of the RDS is people don't use them sufficiently."*

Another difficulty surrounding the endorsement of guidance highlighted by some methodologists was reaching agreement across different funding bodies, given that not all funding bodies currently agree on the role and purpose of PFS.

*024 (CTU Director/Funder): "If you've got two funding bodies that are actually coming at it from very different angles, then it may not be very helpful to have broad guidance, it's about what your target funding agency is thinking about. So, if you've applied to CRUK then you need to know what the CRUK rules are for feasibility, if you're applying to NIHR you need to know their rules. So, maybe there's a space for funding bodies to discuss this, and if there is a lining up of view then it could be useful to generate something generic."*

### 5.6.2 Collaboration

Methodologists and more methodologically experienced surgeons perceived that collaboration between clinical groups and methodologists/trialists would aid more efficient design and conduct of PFS whilst simultaneously contributing to educating the surgical community (see also section 5.6.1). Improving surgeon collaboration and engagement with methodologists and CTUs was therefore seen as a key solution to optimise PFS.

*026 (CTU Director/Funder): "And so a trial unit might not be involved necessarily in running the feasibility due to expense or capacity or something, but they have to be involved in designing it, because you've got to think of the design of the main trial. You can't think about feasibility in isolation, so it's absolutely crucial that you get on well with them."*

*018 (CTU Director): "I think you'll probably need a methodologist involved […] even if it's just the pilot and you're going to throw all the data away and start again. You don't want to learn a lesson that someone who runs loads of trials knew beforehand."*

*019 (CTU Director/Funder): "You probably do need some CTU advice because if you're a clinician who's not done a lot of trial work, you probably don't know what you don't know. You know all the clinical uncertainties, but you don't know the methodological uncertainties."*

Additionally, methodologists perceived how working collaboratively with surgeons might be mutually beneficial and that improving surgeons' perceptions of the value of methodological expertise would enable better collaborative working.

*010 (Methodologist/Funder): "I mean it's always lovely if you've got an experienced CI but more importantly it's lovely if you've got a CI who's able to not think that their job is to know it*

*all and who is able to take direction and listen to and act on advice […] responsive to the fact that you're bringing your expertise in and they've got theirs and it is that blending of it."*

*012 (CTU Director): "It's just when busy clinicians try to do everything it seems really difficult for them and that's where I just think a trials unit helps because you get the database done and tested, you get someone taking care of all the documents […] some of the big units don't, but we're kind of helpful to our local investigators."*

Both methodologists and methodologically experienced surgeons understood that collaborating with a fit-for-purpose trial team with the necessary expertise to consider the research question fully, was important for improving PFS.

*008 (Surgeon/Funder): "You need the right team but […] the constituents of the right team depend on what elements of the feasibility you're addressing."*

*004 (Surgeon): "If you've got a good team they recruit well, work together […] and they enjoy it and therefore it all gets better and better […] relationships and teamwork is fundamental across trials actually."*

Methodologists also perceived that the PFS team might require broader types of expertise than the main trial team.

*020 (CTU Director/Funder): "I think the team is potentially wider in a feasibility study than it would be in the main trial because there are so many uncertainties. You need to engage probably in a more diverse group of people."*

Collaborating with patient and public involvement groups as early as the PFS stage, was also perceived by methodologists as key to improving the design of surgical PFS.

*024 (CTU Director/Funder): "I think probably patient groups are even more important, because a lot of the time the feasibility fails because the investigators haven't worked out what are the patient centred issues."*

*016 (Methodologist/Funder): "Patient input I think is essential, not just token patient input, but having patients who are engaged and can be co-applicants and can offer proper input rather than token input. Because again from a patient's point of view if you're proposing things in a study that patients are just not going to do, then you need to know that quite early on, so that you design your study in a different way."*

### 5.6.3 Efficient funding infrastructure

Surgeons were less explicit in considering education about research methodology a fundamental necessity to improve PFS. Instead, surgeons focused more generally on the current research infrastructure being too inefficient to allow thorough and timely evaluation of surgical interventions. This was considered particularly relevant to surgical research where new technologies and techniques are often introduced quickly and without the same level of regulation as, for

example, the pharmaceutical industry. Improvements in surgical research infrastructure were perceived, therefore, as a key solution to improve PFS by surgeons.

> *023 (Surgeon): "I think we probably need to look at more efficient approaches, and I think that the current trials structure is very inefficient and is probably not fit for purpose for new technologies in the future because they're coming through so quickly and so fast we have to be able to evaluate them in a much more efficient manner, and discard ones that aren't working. At the moment, our structures are completely hopeless for that."*

Methodologists also supported improvements to the research infrastructure but were much more specific in how they thought it could be improved to allow PFS studies to be undertaken more efficiently. Some funders, for example, described how there should be caveats where an external pilot can run seamlessly into a main trial if, the outcome data has not been analysed, the pilot and main trial are not too dissimilar, and the PFS shows that the main trial is feasible.

> *008 (Surgeon/Funder): "If you've not analysed that data and you've just stood it alone and set it aside until you get funding for the major study, then personally, on a funding panel, I wouldn't have a problem with that. In fact you just saved me an X amount of pounds cos you've already got 100 patients […] the reality is that we've got to run these things efficiently, so to discard 100, 200 patients within a study that seems ridiculous if you don't change the interventions."*

> *015 (Methodologist/Funder): "So perhaps the cleanest thing would be, which might address both funders' needs and researchers' needs, is that there's one proposal which includes an external pilot and there's the progression criteria and if the external pilot is fine then the remaining 3 million is released."*

However, this approach still incurs a lag time between the pilot work and the definitive study, and most funders felt that this was the clear disadvantage of a standalone (external) PFS design.

> *010 (Methodologist/Funder): "I think you have to understand the pace of research and the clinical area as well, to determine whether or not it's achievable or reasonable to do the external pilot."*

Many methodologists felt that, whilst applicants should not waste time designing a main trial if there was still considerable uncertainty, if the question was sufficiently important, then funding for a main trial should be planned as a staged approach.

> *017 (CTU Director/Funder): "With commissions somewhere we've said maybe this should be a pilot, a standalone pilot and we say; 'well they know what they're doing, if they can't recruit then that will stop'. It maybe harsh but it will stop at the pilot stage, and that always feels nasty cos you feel you've had a big grant snatched away from you. But if it is going to work that would be a much better model, cos everything will be set up and they can roll in to it, rather than as you say, taking an 18 month hiatus to reapply."*

*027 (Funder): "I think some flexibility. Sometimes you'll know in advance whether it's going to be an external pilot or, an external feasibility study and, sometimes you won't. There should be enough leeway in judgement at the point of having finished the feasibility study or the pilot to say, actually this isn't going to change and therefore the data can be used as an internal pilot."*

*011 (Funder): "If you're convinced there's a case that we need a trial to answer it, and that we're pretty clear about what the patients, the intervention, the control and the outcome should be then it should be an internal pilot, but with perhaps phased funding to allow withdrawal of funding if it's not feasible."*

Conversely, some funders perceived inevitable limitations with this approach, explaining how funding can't be entirely open-ended without the definitive study being fully costed. Furthermore, funding bodies were regarded as not currently set up to make quick decisions on large sums of money.

*019 (CTU Director/Funder): "If you're in that position so you've got an external pilot which has gone nicely and you want to roll straight through but...so currently funders are not set up to make quick decisions about large amounts of money, so if you come along asking for another million pounds, the chances that that's gonna go through on a nod are pretty low."*

In addition, one funder perceived that standalone PFS should be reported and made publicly available, from a standpoint of probity and good ethical practice, before further funding is released.

*011 (Funder): "There's an issue of probity that if you funded the standalone…you're committed then, to wanting the standalone pilot to be written up, be reviewed and made publicly available before you fund anything else, because you've committed public funding to that and we have a right to know what the pilot study shows and that's a slow process."*

However, overall, funders recognised that modifications to the current system of funding PFS, both to improve efficiency and reduce waste, were necessary and possible. For example, many funders felt a more formal and linked funding set-up between different funding streams within the same funding body, would offer a solution to improve the efficiency of funding infrastructure. This approach is more in line with the model used in the NIHR Programme Grants for Applied Research (PGfAR) funding stream. Funders felt this approach would help to prevent the funding of PFS that are unlikely to be sufficiently important to warrant major investment in a definitive study. Consequently, it was felt this would improve the efficient progression of standalone PFS to feasible main trials, and therefore prevent the waste of valuable funding resources.

*027 (Funder): "I think it's one of the major advantages of programmatic funding is that you can fund the whole thing potentially."*

*025 (CTU Director/Funder): "I think that there is an argument for doing something that's a bit more joined up. I mean the programme grants I think are really good because they do allow people to go through those stages […] I think that's a really helpful model and in some ways it would be better if everything went under that model."*

### 5.6.4 Wider dissemination of pilot and feasibility studies

As illustrated in section 5.5.5 (Challenges of reporting and interpreting PFS), the methodologists identified that dissemination of PFS results and reports was currently limited and sub-standard. This was perceived to be due, in part, to a lack of requirement by all funders to release the report into the public domain, and also due to variable understanding of journal editors regarding the purpose and value of PFS. As solutions to this problem, methodologists described how threaded publications and open access online repositories for all documents relating to a research study, not just the primary journal publication, would be beneficial.

*017 (CTU Director/Funder): "They can certainly publish a register of studies, they could certainly publish the protocols easily on line so and whether there could be some grossly light touch way, even if people were uploading their own report. So yea, something so fit for purpose and proportionality."*

*027: (Funder): "I think digital publication is, and the idea of threaded publications, so you go to an online site where there's the original protocol, there's all the referees comments, there's the revised protocol, here's the feasibility study protocol, there's the main trial protocol following the feasibility study and then all the publications are threaded, attached to that."*

Methodologists also perceived that, as part of any grant requirements, there should be greater emphasis from funders on incentivising and prioritising reporting of PFS.

*017 (CTU Director/Funder): "[for most funders] you give the funder a report for internal purposes but then you publish it separately. But in general you want at least some kind of report somewhere and again how you do that is the whole question of open publishing and making data available but in principle absolutely […] they should be in the public domain, and exactly how that should be done, I guess it's an infrastructure thing, but I think now more are being exclusively funded that…it follows on from that."*

Methodologists, therefore, felt that improved dissemination of PFS would augment the recognised value of PFS through providing learning opportunities across clinical areas, avoiding repetition of failed ideas and, consequently, reducing research waste.

*014: (Methodologist): "Well, so that people can learn from each other […] so that people don't keep re-inventing the same wheel. People can learn, but you learn across disciplines as well […] most people are very discipline specific, on focus and outlook."*

*027: (Funder): "Whether that's via the report in the journal's library or whether it's via papers and journals to some extent doesn't matter. My point is that the aim should be 100% of publicly funded research does get an airing so that people can learn."*

## 5.7 Summary

The PEPSTAR study has further identified and expanded on exactly what the key challenges are for the design and conduct of surgical PFS, why these challenges and barriers to optimal practice exist, and what solutions might be necessary to improve future practice in this area. The findings build on knowledge and understanding of the issues affecting surgical trials more widely, to understand specifically, from the perspective of all key stakeholders, why research practice around surgical PFS is perceived as difficult and therefore not currently optimised.

The findings of the PEPSTAR qualitative interview study (phase two), and the targeted review and systematic analysis of NIHR funded surgical PFS (phase one), were synthesised and interpreted, as described in chapter three, to produce detailed recommendations from this work for how to optimise future PFS in surgery (phase three). The next chapter (Chapter six) will describe and explain the results of this synthesis process, and present detailed recommendations for improving research practice more widely amongst all key stakeholder groups and a 'Top Tips' tool to operationalise these recommendations for practical use by surgeons designing and conducting PFS in surgery.

# CHAPTER 6

## *Phase III*

## *A synthesis and interpretation of the findings in this thesis to inform recommendations for the optimisation of pilot and feasibility studies in surgery*

## 6.1 Introduction

This chapter provides a complete written account describing and illustrating the results of phase three of this work. Objectives one and two of this thesis were completed through phases one and two of this work. The objectives of phase three of this work were, therefore:

> **Objective 3** To synthesise and interpret the findings of the targeted review and systematic analysis of protocols of pilot and feasibility studies funded by the NIHR and the in-depth qualitative interview findings, in light of the methodological and theoretical literature around pilot and feasibility studies (as reviewed in chapter one and two) and;
>
> **Objective 4** To develop recommendations from this work for all key stakeholders and, specifically, accessible and practical top tips for surgeons to optimise the design and conduct of pilot and feasibility studies to inform main trials in surgery.

This chapter will be divided into two parts. The first part of this chapter will address objective three and present the synthesis and interpretation of the findings

from phases one and two. The second part of this chapter will focus on objective four to present the recommendations from this work.

For clarity and ease of use, the diagram illustrating the process of data collection and analysis leading to the synthesis and interpretation of the findings in phase 3 of this work is repeated from Chapter 3 (Methods) below.

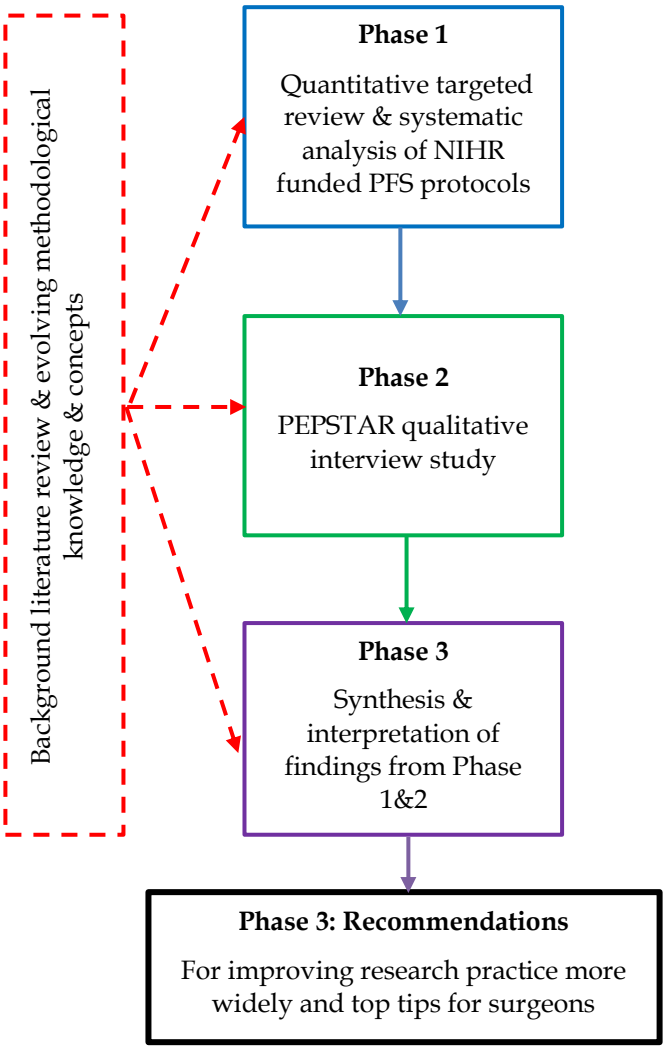



***Figure 3c*** *Process of data collection and analysis in this thesis (repeated from chapter 3).*

The key findings from phase one (Chapter 4) and phase two (Chapter 5) of this thesis are summarised and illustrated in Figure 6a. All the findings from this work were interpreted whilst also considering the existing methodological concepts surrounding PFS and the background literature and knowledge in this area, which has evolved over the course of this thesis. Figure 6a below therefore maps onto Figure 3c above. Reference to the methodological and theoretical literature

surrounding PFS generally, and the guidance documents relevant to PFS in surgery, are made in this chapter where appropriate. This is because the synthesis process occurred within the context of these documents and the evolving concepts around the design and conduct of PFS that occurred during the course of this thesis.

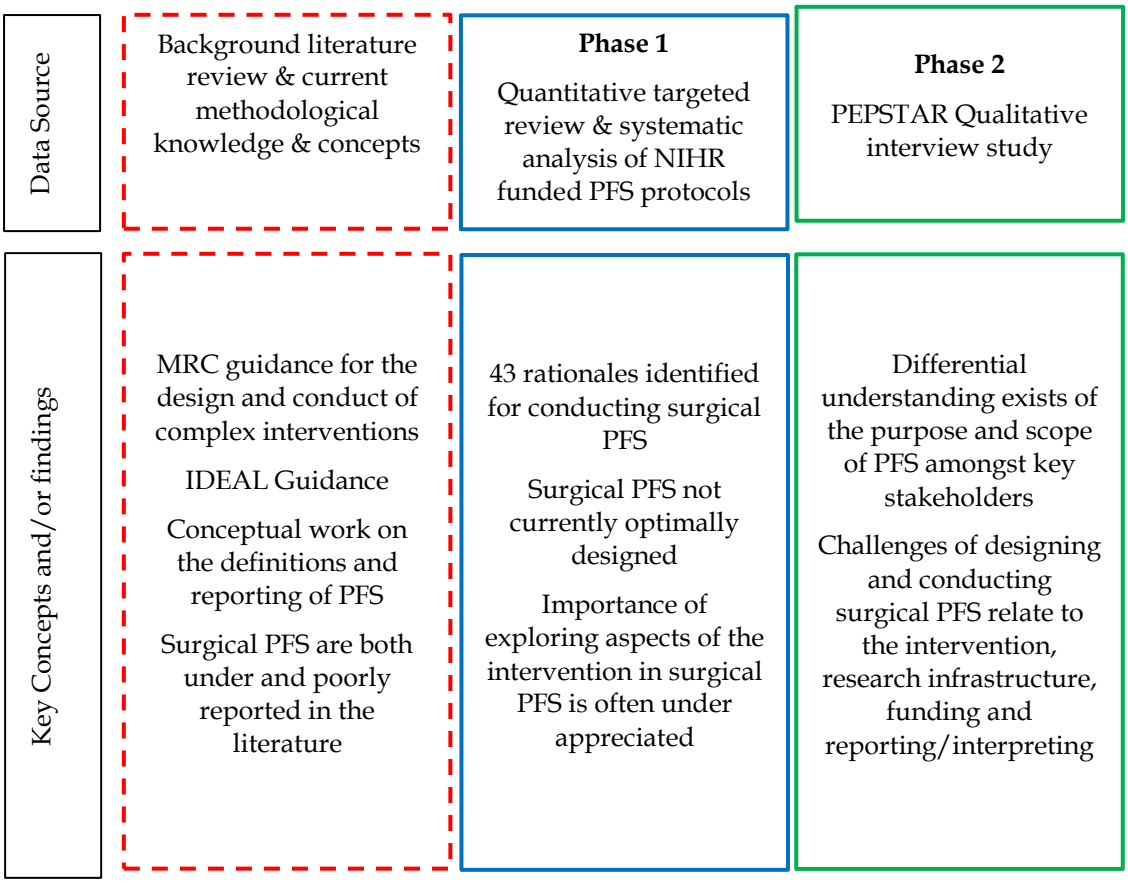



***Figure 6a*** *Summary of the key findings from phase I and II of this thesis, interpreted in light of the known methodological and theoretical literature.*

## 6.2 Part 1: Results from a synthesis and interpretation of the findings in this thesis

### 6.2.1 Defining a cyclical problem surrounding the conduct of surgical pilot and feasibility studies

During the synthesis and interpretation of the findings from phases one and two of the research, a cyclical model linking a variety of sub-optimal research practices around the design and conduct of surgical PFS emerged and is illustrated in

Figure 6b. This model comprises four root causes of why surgical PFS are not currently optimised, which are: 1) A differential appreciation and /or misunderstanding of the breadth and scope for PFS to inform main trials in surgery ('not optimally understood'); 2) Consequently, surgical PFS are potentially not optimally designed or conducted, or not perceived as valuable to complete at all ('not optimally conducted'); 3) Many surgical PFS that are done, are not reported or reported incorrectly ('not optimally reported') and; 4) This leads to PFS being undervalued by surgeons, journal editors, academic institutions and potentially funders .

In addition to the root causes identified for why surgical PFS are not currently optimised, there are also compounding factors, which are linked to both the root causes and to each other. These compounding factors are: 1) The challenges relating to current guidance and; 2) The challenges relating to the cultural issues surrounding both surgical research in general and surgical PFS more specifically. These challenges can variably impact at different and multiple points in the cycle acting as barriers to improved research practice, as illustrated in Figure 6b.

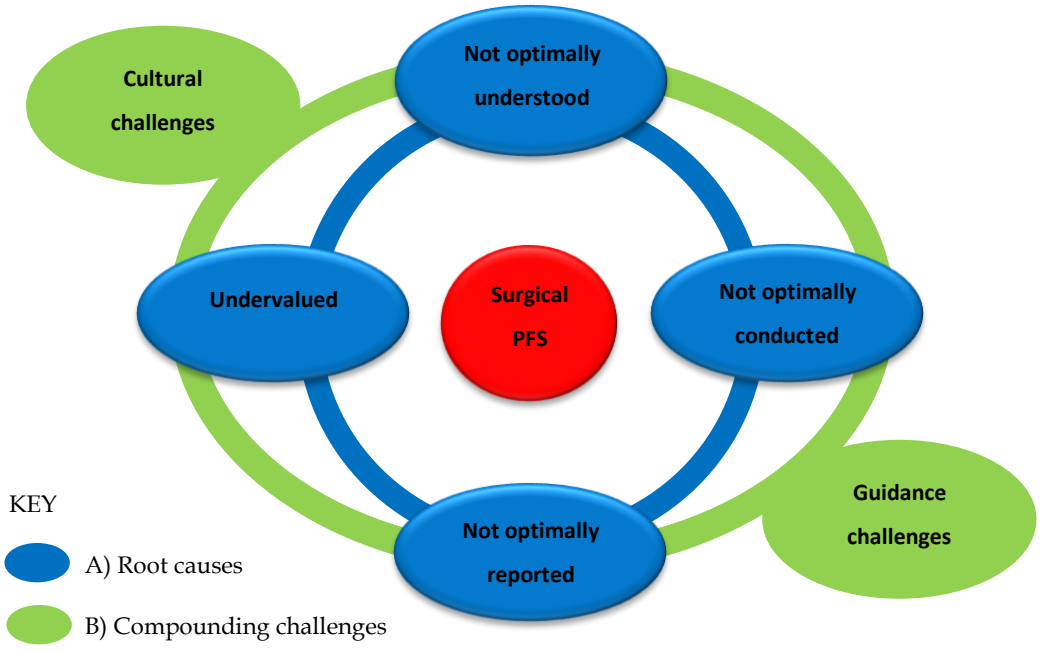



***Figure 6b*** *Cyclical model of sub-optimisation of PFS illustrating the linked and co-influential problems and challenges underpinning why surgical PFS are not currently optimised*

Recognition of the root cause problems and the compounding challenges causing sub-optimal design and conduct of surgical PFS, and how they interact and influence each other, offers insight into what solutions are needed to improve research practice in this area. Each of the components in the cyclical model will now be discussed in turn, with consideration of how the components interact and influence each other. Throughout, how the different data sources compared, converged, or complemented each other in a process of triangulation, to form the cyclical model, will be highlighted and discussed.

### 6.2.2 A) Root causes of why surgical pilot and feasibility studies are not optimised

#### Surgical PFS are not optimally understood

Differential and problematic understanding of the purpose and scope of PFS amongst the surgical research community was a finding that converged from all data sources. For example, the targeted review and systematic analysis of NIHR funded PFS protocols, demonstrated that nearly a quarter of PFS studies planned to conduct formal hypothesis testing. This finding converges with that of the PEPSTAR interviews, which found that many surgeons still perceive PFS as small underpowered RCTs, designed and reported with statistically tested outcome measures of safety and effectiveness. Hypothesis testing is, therefore, still commonly done in PFS, and held up as evidence of effectiveness of interventions in the published literature.

Even the more methodological experienced surgeons in the PEPSTAR interviews, viewed PFS in more pragmatic terms, as a tick box exercise or steppingstone to achieving main trial funding. Whilst not incorrect, this purely practical application of PFS potentially limits their optimal use. These findings converged with the findings from phase one, where exploitation of the full potential of PFS to explore all relevant uncertainties to the research question was not observed.

These differences and difficulties around understanding PFS could be described as a limited understanding amongst surgeons, but this would be too simplistic. From the PEPSTAR study, understanding seems to be intrinsically related to both

experience and professional roles within research and clinically, and in particular to roles relating to clinical trials. However, the majority of surgeons in the UK do not have significant methodological experience as a major part of their professional role. The PEPSTAR study showed that whilst methodologists had a more complete and accurate understanding of the principal purpose of PFS in optimising the main trial, in general most surgeons, lacked a basic understanding of what PFS are, and how they should be utilised.

### Surgical PFS are not optimally conducted

If there is misunderstanding of what PFS are and why they should be done, this will impact on the range of issues that surgeons will seek to explore in PFS, meaning PFS are not optimally conducted. Sub-optimal conduct of surgical PFS was demonstrated in the targeted review and systematic analysis of NIHR funded surgical PFS protocols. A tendency to focus on issues in surgical PFS that are generic to all trials, such as recruitment, was observed rather than considering the breadth of uncertainty that is of specific importance to trials of surgical interventions. This finding converged with data from the PEPSTAR interviews, with methodologists perceiving that there was a propensity for clinical and inexperienced applicants, to follow 'example' lists provided by funders, of what to consider in PFS, rather than thinking more widely and specifically as to the individual study needs.

More specifically, complementarity was demonstrated between the quantitative and qualitative findings, of an under appreciation of the importance of exploring the intervention in PFS in surgery. The findings in phase one illustrated that only a quarter of PFS sought to explore uncertainties specific to the surgical intervention. This was complemented by findings in the PEPSTAR study, where the full range of, and reasons for, the many challenges facing trials of surgical interventions, were not always recognised by surgeons. More importantly, the potential for using PFS to explore the unique challenges of surgical trials was generally not fully appreciated by surgeons.

### Surgical PFS are not optimally reported

The findings from phases one and two converged in agreement that surgical PFS are currently not optimally reported. The data from the targeted review and systematic analysis of NIHR funded surgical PFS illustrates categorically, that PFS in surgery are frequently under reported with only two thirds publishing the results to date. Complementary to this finding, the continued practice of inappropriate publication of PFS masquerading as RCTs and/or underpowered RCTs being badged as PFS *a posteriori,* was identified in the PEPSTAR study. The PEPSTAR data also demonstrated that there were several cultural barriers impacting on the publication and appropriate interpretation of surgical PFS, resulting in suboptimal reporting of such studies (see section 6.2.3).

### Surgical PFS are undervalued

The findings from the targeted review and systematic analysis of NIHR funded PFS protocols demonstrated that PFS are sub-optimally conducted and reported. The PEPSTAR study complemented these findings by illustrating the undervaluation of PFS by all key stakeholder groups, which leads to suboptimal conduct and reporting of PFS. Funders for example, perceived that many PFS had historically been conducted as standalone pieces of work with no intention of the study team, funders or both to progress to a main trial. PFS were consequently undervalued as being ineffectual and not worth investment. Similarly, academic institutions were perceived to undervalue PFS, considering them low impact studies, which do not contribute significantly to the Research Excellence Framework (REF) as three- or four-star papers, and often have no outputs at all. This is in part, perpetuated because journal editors undervalue PFS as not offering definitive practice changing results, and therefore are of limited interest to the readership and clinical/research community. This perception means PFS are often challenging to publish, meaning authors don't attempt to publish them. Finally, surgeons perceived PFS to lengthen the process of trials research, meaning answers to important questions took longer to attain. Consequently, the research question may become obsolete in fast moving clinical areas like surgery, before the research was finished. Undervaluation of PFS by all key stakeholders therefore perpetuates the cyclical model of sub-optimal design and conduct of such studies.

### 6.2.3 B) Challenges compounding the cyclical model of sub-optimisation of surgical pilot and feasibility studies

The synthesis and interpretation of findings from phases one and two of this work has produced a cyclical model of sub-optimisation of PFS which links the issues underpinning why surgical PFS are sub-optimally designed and conducted (see Figure 6b). There were two key compounding factors identified from this synthesis for the perpetual sub-optimisation of surgical PFS. The first is the variability and inconsistency of both the guidance available from funders, and the guidance available in the published literature. It is important to note, however that the guidance available relevant to surgical PFS has changed and evolved during the course of this thesis (see chapter 7 for full discussion of this). The second is the cultural challenges surrounding surgical trials more generally which also impact on surgical PFS, and the challenges of PFS more specifically. The next sections will discuss the compounding challenges identified from this synthesis, contributing to why surgical PFS are sub-optimally designed and conducted.

#### GUIDANCE CHALLENGES

#### Funder guidance

The PEPSTAR study findings highlighted that funder guidance was perceived to be limited to brief and variable definitions of PFS, which are not the same and/or complimentary to the definitions given in the theoretical and conceptual guidance. The NIHR glossary definitions [77, 78] of PFS were, for example, until recently the only guidance offered by the NIHR on conducting PFS. These NIHR definitions of PFS were perceived by the PEPSTAR participants as both linear and succinct but lacking any detailed explanation of when to choose different PFS designs. Whilst some methodologists and trialists quoted the definitions given by the NIHR as 'guidance' during the PEPSTAR interviews, there was acceptance that ongoing confusion exists in the wider surgical research community. Using the terms 'pilot' and 'feasibility' interchangeably may seemingly cause little harm. However, underlying the interchangeable use of terms by surgeons applying for PFS funding is a fundamental misunderstanding of the full purpose, value and

potential of pre-trial work and how it applies to the individual research area as demonstrated in phase one of this work. This results in using the nuances of different definitions to design studies which may not be fit for purpose, so PFS will consequently continue to be sub-optimally designed and conducted.

### Theoretical/conceptual guidance

The PEPSTAR data illustrated that whilst most methodologists recognise the extensive work already done to conceptualise the types, purpose and reporting of PFS [71, 79-81, 90, 132], many felt this work to be inaccessible and poorly disseminated to surgeons. There was limited awareness of its existence amongst surgeons, and the few surgeons who did mention the conceptual work, felt that it was mostly theoretical and generic, making it largely unhelpful. This finding indicates that the conceptual work is poorly understood and not widely acknowledged beyond the methodological community. Indeed, the newest NIHR guidance published in 2019 [197], does not quote any of the theoretical/conceptual work as further information on best practice in this area. Even the IDEAL guidance [9, 100], which is the conceptual work most aligned with surgeons and surgical trials, was not perceived in the PEPSTAR study, to be widely accepted or utilised amongst surgeons.

If available guidance is inaccessible, not publicised and/or not endorsed by funders and research design services, it will not be pragmatically incorporated in research practice. The PEPSTAR interviews perceived that amongst those without, or not incorporated within groups with, methodological expertise, the available theoretical/conceptual guidance on the design and conduct of PFS was both inaccessible and/or often unheard of. Furthermore, there was a perception that the available guidance (including the IDEAL guidance) is too generic for surgeons to operationalise.

In summary, if none of the available guidance is operationalised specifically so that it is of practical use by surgeons, its inaccessibility is compounded, further adding to confusion so that it is consequently misunderstood or ignored. The cycle

of misunderstanding and undervaluation of surgical PFS is, therefore, perpetuated, by a lack of consistent, accessible and operationalised guidance.

## CULTURAL CHALLENGES

### Surgical trials

The PEPSTAR findings showed that many of the existing challenges for surgical trials, also impact on PFS, and in doing so, make PFS even more important in surgery. In addition, the findings in the PEPSTAR study emphasised the blossoming culture of surgical research in recent years, through the formation of nationwide surgical trainee research collaboratives and the RCSEng Surgical Trials Initiative, for example. These newer developments have undoubtedly contributed significantly to raising the profile of surgical research and, more specifically collaboratively conducting trials in surgery. However, the methodological robustness of some research performed by surgical trainee collaboratives was highlighted as potentially problematic by some PEPSTAR participants. Furthermore, findings in both phases one and two, suggest there is some way to go before methodologically sound surgical trials are the domain of the majority rather than the interested few.

### Reporting of surgical PFS

Inappropriate reporting of PFS was identified as a root cause of suboptimal research practice in this area, as illustrated in Figure 6b. However, there are also cultural influences on the reporting of PFS. The PEPSTAR findings demonstrated ongoing misunderstanding even amongst surgical journal editors, some of whom desire only to publish definitive results and therefore request that these are included for the publication of a PFS. This editorial practice will consequently perpetuate the cycle of misunderstanding; if definitive results are requested, authors may feel compelled to produce them in order to achieve publication. In addition, the findings from phases one and two converged in confirming that variable funder policy on public reporting of PFS, impacts on the dissemination and perceived value of PFS. Whilst there is now a journal dedicated to the reporting and publication of PFS (*The Journal of Pilot and Feasibility Studies* [134]),

without investment of journal editors, academic institutions and funders to drive the importance of publishing pre-trial work, the cycle of sub-optimisation of PFS will continue.

### Funding surgical PFS

The targeted review and systematic analysis of NIHR funded PFS studies demonstrated that almost four years beyond the end of the funding period (2015) for included studies, only three PFS (3/35, 8.6%), have progressed with certainty to a main trial, with at least a third (11/35, 31.4%) falling into relative obscurity. This finding was complemented by the PEPSTAR study findings, where it was perceived that standalone PFS had historically been funded, completed (or not) and reported (or not) but importantly, without any intention of funders, study teams or both, of progressing to a main trial. Whilst this practice, was perceived in the PEPSTAR interviews to now be less common in recognition of this issue, the structure of many funding streams remains unchanged. Most funders still offer uncoupled funding where a PFS may be funded initially, without firm promise of main trial funding.

The inefficiencies in the current funding structure were demonstrated by the convergence of findings from phases one and two of this work. Phase one showed that PFS are sub-optimally reported and rarely progress to a main trial despite demonstrating feasibility. The PEPSTAR study illustrated that the additional time and resources needed to perform standalone pre-trial work, represents a major barrier to completing them. In addition, there was scepticism demonstrated in the PEPSTAR study, over whether PFS (especially standalone/external studies) make enough difference to main trial success and/or preventing research waste to justify the time and additional resources necessary to complete them properly and effectively. The benefits of pre-trial work in avoiding failed main trials, may therefore be far less valuable, if the pre-trial work adds so much time to the research pathway, that the question and interventions being studied become outdated and irrelevant before definitive answers are acquired.

### 6.2.4 Summary of the synthesis to produce a cyclical model of sub-optimisation of pilot and feasibility studies

Overall, the synthesis and interpretation of phases one and two of this work, produced a cyclical model explaining the sub-optimisation of surgical PFS. The root causes of sub-optimisation of surgical PFS were identified, along with the compounding challenges. The root causes and compounding challenges were shown to interact and impact on each other in a cyclical and thus co-influential way. Each of the root causes and compounding challenges have been discussed above in terms of how the concepts were derived in light of the findings of this work. The cyclical model was then used, in the context of the methodological or theoretical literature and the available guidance in this area, to produce recommendations for all key stakeholders for optimising the design and conduct of PFS to inform main trials in surgery.

## 6.3 Part 2: Recommendations from this work for optimising the design and conduct of surgical pilot and feasibility studies

The recommendations from this work to improve research practice around the design and conduct of surgical PFS are described in the second part of this chapter. Firstly, recommendations are made for changes to wider research practice. These broad recommendations are followed by the presentation of a 'top tips' guidance tool for surgeons to specifically operationalise these recommendations and disseminate them to the wider surgical community.

The key recommendations derived from this work for how to optimise future surgical PFS, and which phase of the work contributed to each recommendation, are outlined in Table 6.1. These recommendations identify four key areas for improvement to research practice necessary to optimise future PFS in surgery, termed education, collaboration, funding and dissemination. Each of these areas will now be discussed in turn, including how they relate to and address the

cyclical model of sub-optimisation of surgical PFS produced from the synthesis process.

***Table 6.1*** *Recommendations derived from this work to optimise surgical pilot and feasibility studies*

| | Recommendation | Further detail | Issues to consider | Informed by | |
|---|---|---|---|---|---|
| | | | | Phase 1 | Phase 2 |
| EDUCATION | **Improved guidance on designing and conducting PFS** | Consensus based guidance endorsed by regulatory bodies, funders and journals. | Theoretical/conceptual guidance operationalised specifically for application by surgeons in practice. | X | X |
| | **Grass roots training for surgeons from earlier in career** | Training in trials methodology through courses, conferences, publication and guidelines. | Might also be achieved through collaboration. | | X |
| COLLABORATION | **Collaboration of surgeons with methodologists & CTUs** | Working closely with methodologists and CTUs from earlier in the research process to ensure the future main trial is in sight. | Practicalities of funding collaborations. | | X |
| | **Accessibility of CTUs & methodology support** | Highlight where to go/who to ask for assistance in the new guidance. | Consensus process should consider what level of methods support is enough for PFS. | | X |
| FUNDING | **Improved efficiency of funding structure** | PFS should only be funded if main trial would be funded (i.e. important enough research question) so PFS are not wasteful. More joined up funding so no lag time between successful PFS and main trial. | More programmes offering staged funding like NIHR PGfAR to improve efficiency and reduce waste. | X | X |
| | **Raising the profile of the importance of funding surgical studies** | To achieve proportional funding.<br>More surgeons on funding panels.<br>Regulatory requirements for industry to contribute to surgical research. | How to promote surgical involvement on funding panels. | X | X |
| DISSEMINATION | **Funder requirement to publish PFS** | Publication in journals and/or through publicly available funder reports. | How to fund process of publication. | X | X |
| | **Journal editors stop publishing underpowered RCTs as PFS or PFS as underpowered RCTs** | Both wrong! Educate through guidance. | Involve editors of surgical journals specifically in the consensus process for guidance. | X | X |
| | **Academic institutions to value PFS as potentially essential for main trial development** | May not be 3 or 4* alone but need to have improved value for the often essential role they play in the success of the main trial. If academic institutions don't value PFS, researchers will not value disseminating their findings. | How to engage academic institutions in considering the value of PFS | | X |

**KEY:** CTUs Clinical Trials Units, PFS Pilot and Feasibility Studies, PGfAR Programme grants for Applied Research, RCTs Randomised Controlled Trials

### 6.3.1 Education

The first recommendation identified from this work to improve the design and conduct of PFS for surgical trials is enhancing education which can be done in several ways.

#### Improved guidance on designing and conducting PFS

The cyclical model of sub-optimisation established evidence that the available guidance around the design and conduct of PFS in surgery, is considered by surgeons and other key stakeholders, to be largely generic, theoretical and as such inaccessible. New guidance is needed which crosses the interface between practice (surgeons) and theory (methodologists). Such guidance would incorporate the extensive methodological work already performed but importantly, build on this to operationalise it for surgeons and those with less methodological expertise. New guidance should ideally be endorsed by funders, journal editors and other regulatory bodies such as the Royal College of Surgeons in order that it is incorporated fully into research practice.

#### Grass roots training for surgeons from earlier in career

A need for more 'grass roots' education of surgeons from earlier in their careers, is also identified from this work. This would mean educating the surgical community about methodological research practice through courses, conferences, relevant publications and finally, again, accessible guidance. There are steps being taken to achieve this educational strategy nationally, as already discussed in chapter one of this thesis, through the RCSEng Surgical Trials Initiative and the surgical trainee research collaboratives. In addition, courses such as the Bristol Oxford Surgical Trials Course (BOSTiC): Making trials stick [198] and GRANULE (GeneRAtiNg StUdent Recruiters for Surgical TriaLs) [199] are reaching out to trainees and medical students, through education in an area that is currently not taught elsewhere in the surgical training pathway. Whilst such courses are not yet mandatory across all surgical training programmes, there is certainly an argument for this, so that all surgeons are educated in trials methodology as part of their training programme [200]. In addition, the research requirements for surgeons in

training would be far more usefully focused on collaborative trial involvement and training in recruitment, rather than the number of first author publications. This 'publish or perish' phenomenon, potentially only perpetuates poor surgical research, for trainees with neither the time, resources or access to methodological expertise to perform high quality, meaningful studies. Surgeons will also become more educated in research methodology through collaboration with methodologists, and this is the second key recommendation identified for optimising future surgical PFS and is discussed below.

### 6.3.2 Collaboration

Improved collaboration between methodologists and surgeons is the second area of recommendation made from this work, both for improving the design and efficient conduct of PFS, but also to enhance the education of surgeons through learning from working with expert trial teams. Those surgeons interviewed in the PEPSTAR study who had worked closely with methodological experts on surgical trials, recognised the essentiality of this approach. However, many surgeons 'on the outside' of Clinical Trials Units, found accessing such support difficult and sometimes obstructive. This may be due to the relatively limited funding packages available for PFS and the time and input that is still required by methodologists with inexperienced clinical teams. Some methodologist participants however, recognised that the responsibility might lie with the methodological community to translate and transcend this interface, to at the very least provide accessible guidance for PFS. Future consensus-based guidelines should consider the importance and proportionality of methodological expertise for PFS design and conduct. This should particularly consider funder requirements for methodological input, and how this might be achieved within the limited funding parameters of PFS grants.

### 6.3.3 Funding

A third key area of recommendation identified with the potential to optimise PFS in surgery, was improvements to current funding practices and infrastructure.

**Improved efficiency of funding structure**

These improvements included a more cohesive approach between funders and funding streams, at the point of a PFS design and funding application. Here, consideration of the importance of the research question overall, should have more impact on whether a PFS is funded as the initial step. Funders need to consider, whether a main trial would be funded in the future, and whether the study applicants are appropriately supported by the necessary expertise enabling continuation to a main trial. Certainly, this advice has now appeared in the most recent NIHR RfPB guidance on the NIHR website, where an outline of the main trial is requested as part of the PFS application and, if the main trial is proved feasible it is expected the application for the main trial will be submitted in the timeframe of the PFS [197]. In addition, more opportunities for staged funding, like the NIHR PGfAR funding stream may be beneficial to improve the value and efficiency of PFS for surgical trials. A key recommendation from this thesis is, therefore, that PFS are, essential to optimising surgical trials, but more efficient ways to fund them need to be found. An improved funding structure for PFS, would allow progress to a main trial, if the PFS shows one is possible, to be seamless. This would reduce the additional time added by a PFS in getting to a definitive answer, thus avoiding waste through trials becoming obsolete before finishing, and improving the perceived value of PFS overall.

### Raising the profile of the importance of funding surgical studies

A related issue to improving the funding structure for PFS, is the fact that funding for surgical studies is known to be proportionately underrepresented. The findings in phase one, demonstrated that of all clinical research funded by the NIHR HTA and RfPB programmes from 2005-2015, only 10.4% (140/1341) were studies of surgery as the main intervention, and of these just a quarter were PFS (35/140, 25%). This converged with findings in the PEPSTAR study, where surgeon applicants felt disadvantaged in the funding process. However, whilst the qualitative work did consider the views of a wide variety of funders from different UK funding bodies, this thesis did not quantitatively examine the proportion of surgical research funded by other funding bodies. Nevertheless, this work suggests that some funding bodies do not currently prioritise surgical research,

though some of these findings may also be explained by fewer surgical studies seeking funding, than other clinical areas. Another explanation could be that applications for surgical studies are of poorer quality and are therefore not passing the peer review process.

Certainly, it is apparent that there is under representation by surgeons on funding panels. This is likely due to the evolving culture of research amongst surgeons and should change as interest and involvement in surgical trials by surgeons continues to grow. Funding bodies could advertise specifically for members in underrepresented areas, like surgery, and employers should recognise this as a meaningful use of surgeons' time.

In addition, surgical trials have historically, and probably often still are, being dismissed as too difficult or risky to perform, because of the many challenges of surgical trials identified and explored in the context of PFS in this thesis. Unlike pharmaceutical trials, there is no industry infrastructure around trials of surgical interventions, and no regulatory requirement for the device and implant industry to perform trials, meaning sources of funding are limited. These are all cultural challenges representing barriers to completing surgical trials research in general, therefore equally affecting PFS. Addressing these challenges is not straightforward, and the identified problems represent the need for a cultural shift amongst all key stakeholders in surgical research including funders, regulatory bodies such as the RCSEng, industry, academic institutions and surgeons.

### 6.3.4 Dissemination

Finally, the fourth recommendation identified from this work is the wider dissemination of PFS results. The findings from phase one demonstrated that PFS are sub-optimally reported, with an ongoing propensity for hypothesis testing and that funders have variable requirements on the publication of PFS. Improved dissemination would happen through improvements to funder policies and practices on the mandatory publication of PFS, and journal editorial practice being open to the publication of 'true' PFS. These changes were perceived as essential in the PEPSTAR study for improving understanding of the purpose and value of

PFS, allowing cross disciplinary learning, and consequently breaking the identified cyclical process of sub-optimisation of PFS. Wider and required dissemination of PFS, should subsequently increase their perceived value and importance, thereby cultivating improvements in the quality of design and conduct of such studies.

In addition, academic institutions need to shift their appreciation of the value of PFS. Without this, even researchers who understand the purpose and importance of PFS, will continue to undervalue the dissemination of PFS. It will be important in any consensus process to include academic institution leaders, so that the value of publishing PFS to the Research Excellence Framework (REF) can be reconsidered. For example, instead of measuring the worth of PFS in terms of 3* or 4* outputs and publications, the value of PFS in terms of its impact on reducing waste could be considered, for instance where a main trial is shown to not be possible. The impact of PFS could also be measured in terms of successful optimisation of a main trial where the PFS substantially contributed to main trial design, and ultimately trial conduct and successful completion.

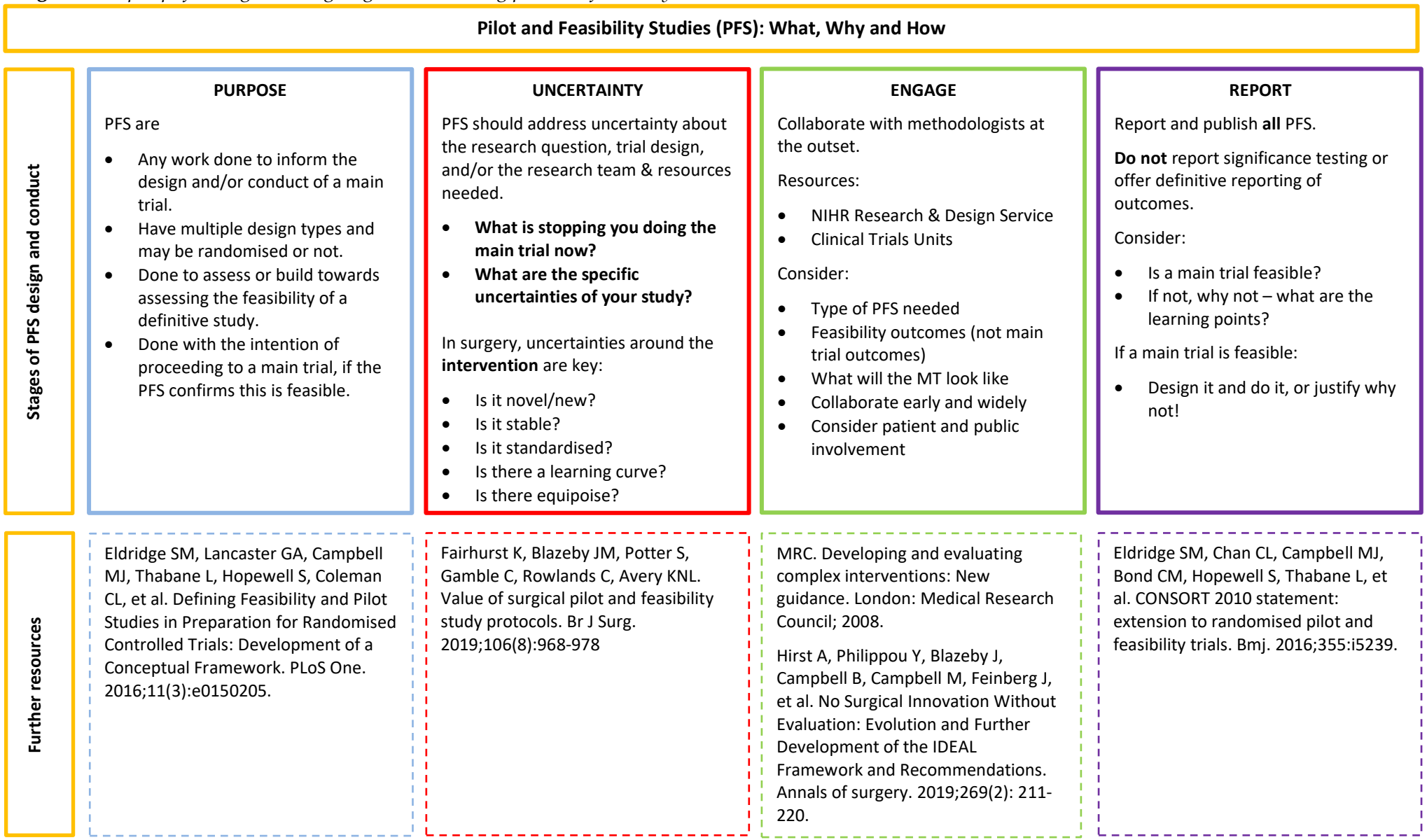
Finally, journal editors need to publish the results of PFS correctly, by recognising their true value and importance and understanding the damaging consequences of publishing underpowered RCTs as PFS or PFS as underpowered RCTs. This shift in practice would be achieved via education from accessible and endorsed guidance, as described earlier (see section 6.3.1).

## 6.4 Top Tips for surgeons designing and conducting surgical pilot and feasibility studies

Practical and accessible guidance that serves to operationalise and bridge the gap between recommendations produced in this thesis and surgeons involved in designing and conducting surgical PFS, is something that this work has identified as acutely necessary in parallel to the wider recommendations for research practice presented in the previous section. The following 'Top Tips' tool illustrated in Figure 6c, is a brief and simplified summary of the key findings and practical

recommendations from this work. The purpose of this diagram is to condense the findings of this work into an accessible tool for those surgeons who are considering or involved with participating in surgical PFS but who may have minimal methodological expertise. This tool was created as a direct result of the work in this thesis but should undergo assessment and validation as part of a future consensus process to produce formal guidelines for PFS in surgery (see Chapter 7 for discussion of future work).

***Figure 6c*** *Top Tips for surgeons designing and conducting pilot and feasibility studies*



## 6.5 Dissemination of the findings and recommendations from this doctoral thesis

### Outputs relating to Phase I of the work in this thesis

Work from Phase 1 of this thesis, relating to the background literature review and the targeted review and systematic analysis of NIHR funded PFS, has already been presented and disseminated in several formats.

**Publication in a peer-reviewed journal:**

Fairhurst K, Blazeby JM, Potter S, Gamble C, Rowlands C, Avery KNL. Value of surgical pilot and feasibility study protocols. *Br J Surg*. 2019;106(8):968-978

This work was published in the *British Journal of Surgery* 201, with the specific aim of disseminating best research practice in this area to a surgical audience.

**Conference contributions:**

1) Katherine Fairhurst, Jane Blazeby, Ceri Rowlands, Shelley Potter, Carrol Gamble, Kerry Avery. **A systematic analysis of UK nationally funded surgical pilot and feasibility study protocols from the last 10 years to inform and optimise future surgical trials.** *Clinical Trials* 2018; 15(S2): 118 (A82)

   **Oral presentation at the Society of Clinical Trials** (SCT), Portland Oregon (international). May 2018. (NB Given on my behalf by Kerry Avery as on maternity leave)

2) Katherine Fairhurst, Jane Blazeby, Shelley Potter, Amanda Blatch-Jones, Ceri Rowlands, Carrol Gamble, Kerry Avery. **Key design features of pilot and feasibility studies to inform successful surgical trials: A systematic analysis of funded studies.** *Trials* 2017; 18(Suppl 1):P16

   **Poster presented at 4th International Clinical Trials Methodology Conference** (ICTMC) & 38thAnnual meeting of the Society of Clinical Trials (SCT), Liverpool (international). May 2017.

3) Katherine Fairhurst, Kerry Avery, Elaine O'Connell Francischetto, Chris Metcalfe, Jane Blazeby. **How can pilot work optimally inform surgical RCTs? A review of current evidence**. *Trials* 2015; 16(Suppl 2):P17

   **Poster presented at 3rd International Clinical Trials Methodology Conference** (ICTMC), Glasgow, (international). October 2015

## Outputs relating to phases II and III of the work in this thesis

**Conference contributions:**

1) Katherine Fairhurst, Shelley Potter, Jane Blazeby, Carrol Gamble, Kerry Avery. **Unique challenges and proposed solutions for designing and conducting pilot and feasibility work to optimise surgical trials.** *Trials* 2019; 20(Suppl 1):PS9D-O5

   **Poster presented at 5th International Clinical Trials Methodology Conference** (ICTMC), Brighton, (international). October 2019.

2) Katherine Fairhurst, Kerry Avery, Alicia O'Cathain, Pat Hoddinott, Jane. When to do an external or internal pilot study: Findings from an interview study with research. *Trials* 2019; 20(Suppl 1):PS5A-O1

   **Poster presented at 5th International Clinical Trials Methodology Conference** (ICTMC), Brighton, (international). October 2019.

The PEPSTAR qualitative interview study was presented at the 5th International Clinical Trials Methodology Conference (ICTMC) in October 2019. There are plans for further dissemination of the work in this thesis, and specifically the key messages for surgeons, at surgical conferences later this academic year (for example at Association of Surgeons of Great Britain and Ireland, ASGBI; Association of Surgeons in Training, ASiT; Association of Breast Surgery, ABS). With this in mind, the 'top tips' tool (see section 6.4) will present the key messages regarding designing and conducting PFS in an accessible and operationalised format for surgeons to use. The aim is to also publish the results of the PEPSTAR study (Phase II), and the synthesis and interpretation of the

findings (Phase III) including the 'top tips' tool for surgeons, in a surgical journal, to disseminate this work through publication.

## 6.6 Summary

The next and final chapter of this thesis will discuss and critique the key findings of this work both in the context of other published literature and guidance documents in this area, and its strengths and limitations. Consideration of how this work adds to the field of research will be made, and proposals for future work will also be presented.

# CHAPTER 7

# *Discussion*

## 7.1 Introduction

This chapter will discuss the key findings from this work and how they enhance understanding of the potential for PFS to optimise future surgical trials. In addition, how these findings contribute and relate to the research field in the context of other work, and in particular other relevant published guidance in this area, will be critiqued and discussed. Strengths and limitations for each phase of the work will be considered, and an appraisal made of the applicability and generalisability of the findings to different audiences and current research practice. Finally, proposals for future work building on the findings in this thesis will be presented.

## 7.2 Key findings from this work

This work used an interpretative process to synthesise the findings from a targeted review and systematic analysis of NIHR funded PFS protocols, and a qualitative interview study with key stakeholders to produce a cyclical model illustrating why the design and conduct of surgical PFS is currently sub-optimal. This model delineates the cyclical relationship between misunderstanding of the purpose, scope and different PFS design types in the context of surgical trials and the sub-optimal design, conduct and reporting of PFS, and how this continues to perpetuate undervaluation and dismissal or disregard of PFS by surgeons, journal editors, academic institutions, and sometimes funders. The root causes outlined in the model are compounded by additional challenges relating to inconsistent and dense, inaccessible guidance and the cultural challenges inherent to the surgical community. The cyclical model developed in this thesis

was used to inform the development of key recommendations for surgeons and study teams for how to improve research practice.

The recommendations from the work in this thesis, intended to comprehensively address the cyclical model illustrating why surgical PFS are sub-optimised. This is achieved by ensuring that the recommendations incorporate necessary and achievable changes to current research practice. The premise at the outset of this thesis was: 'If we could just change the way surgeons think and behave regarding surgical research, everything would be much improved!' However, during the course of the research, it became apparent that not all the challenges identified in this work, are entirely rectifiable by changes to surgeons' behaviour and practice, as initially hypothesised. Much of what is recommended by this work, in fact, requires a wider cultural shift in research practice amongst funders, academic institutions, regulatory bodies and journal editors, as well as amongst surgeons.

The justification for why PFS might be key for optimising trials more generally, has been well described and considered by others, as explored in chapter two of this thesis. These works include the published methodological work on the definitions, study types and reporting of PFS in general, [71, 79-81, 90, 132], the MRC guidance document on developing and evaluating complex interventions in general [13] and the IDEAL guidance on the development and evaluation of surgical interventions specifically [101, 102]. This guidance discusses PFS more generally, but to date no work has specifically considered the value and scope of PFS in surgery. As explained and illustrated in the introductory chapters, when compared to other clinical areas, surgical trials face unique challenges meaning pre-trial work to explore challenges and uncertainties may be even more necessary.

Perhaps in recognition of the limitations of the guidance already available on the design and conduct of PFS, further guidance has evolved during the course of this thesis. These new guidance documents will be critiqued in terms of how they relate to the work in this thesis.

## IDEAL guidance

To reiterate, the IDEAL framework [144] is the currently available guidance most specific to surgery, and describes a pathway for new surgical interventions from the initial Idea (Stage 1) to Development (Stage 2a)/Exploration (Stage 2b), Assessment (Stage 3) and Long-term monitoring (Stage 4). Within this framework, PFS are considered as stage 2a/2b studies, aiming to explore uncertainties before a stage 3 assessment in a definitive RCT. The recently updated IDEAL recommendations were published in print in 2019 [102]. This is four years after the final studies in the targeted review and systematic analysis of NIHR funded protocols were funded, and 14 months after the PEPSTAR qualitative interviews were completed. The updated guidance does provide some clarification and examples of what to consider when designing PFS. In particular, several feasibility issues are suggested for consideration in stage 2a/2b studies including: Estimating effect size (though importantly, what is meant by this this is not clarified in this guidance); defining intervention quality and standards; evaluating learning curves; exploring subgroup differences; eliciting key stakeholder values and preferences and analysis of adverse events. However, as demonstrated in this thesis, this is far from an exhaustive list. In addition, such example lists were perceived in the PEPSTAR study as limiting PFS study applicants into thinking the listed items are the only uncertainties to consider, rather than thinking more laterally about uncertainty specific to the study and research question, and the potential for PFS to explore these uncertainties. In addition, offering 'estimating effect size' as a use of PFS without any clarification, is potentially incorrect and misleading (see next section on hypothesis testing in PFS guidance).

The perception that the IDEAL guidance is not widely accepted or utilised was explicitly highlighted in this work, and further corroborated recently, by a publication in *The Lancet* from the IDEAL collaboration in 2019. This study analysed how surgical research has changed since the first IDEAL recommendations were published in 2009. Random samples of n=25 studies

published in the general literature in the periods 2000-2004 and 2010-2014 were compared [103]. The study found that explanation of modifications of IDEAL 2a studies had not changed over this 10-year period (4/25, 16%), though the use of prospective cohort studies had slightly improved (5/25, 20% to 9/25, 36%). However, the mention of PFS to prepare for an RCT had, at best, not improved at all over this time (2/25, 4% to 1/25, 1%) [103]. The authors also note, that statistical testing was still being employed inappropriately in pre-trial studies [103]. This work from the IDEAL group suggests that the original IDEAL guidance, has had little impact on the design and conduct of surgical PFS over the time period (2005-2015) of the PFS analysed in this thesis.

It would be reasonable to surmise, therefore, that the IDEAL guidance is currently not adequate in terms of operationalised detail for surgeons to improve the design and conduct of surgical PFS specifically.

**Guidance on hypothesis testing in PFS**

There has been ongoing and evolving contention during the course of this thesis regarding if and how hypothesis testing should be performed in PFS. This issue permeates through multiple components of the cyclical model of sub-optimisation, and consideration of this issue is, therefore, discussed here in terms of published guidance and debate of this issue.

Phase II studies are usually defined as early phase pharmaceutical trials, and have been methodologically established for longer, with very different feasibility issues from studies of complex interventions. In the context of pharmaceutical trials, a Phase II study could be regarded as synonymous with doing a PFS and means specifically testing for signs of 'promise', typically an early signal of safety (e.g. toxicity) and effectiveness (e.g. reduction in a viral count). It is possible, therefore, that because of the more established methodology for Phase II trials, with the defined purpose of demonstrating preliminary evidence of efficacy, this has influenced the perceived purpose of PFS for complex clinical interventions. It is evident, from the work in this thesis, that understanding around when, how and if to use hypothesis testing in PFS has not yet become widespread. In

addition, whilst some might argue that the same rationale of seeking early signs of promise can be applied to PFS of surgical interventions, there are challenges with taking this approach.

Firstly, safety is always reported as part of an adverse event profile in any surgical trial, including a PFS. Unacceptable death or major complication rates would be monitored by the Data Monitoring Committee (DMC) and cause a PFS or trial to be reviewed and/or stopped by the Trial Steering Committee (TSC). However, the measurement of clinically important outcomes in surgical trials may have extrapolated endpoints, for example, recurrence of cancer or quality of life assessments, meaning effectiveness may be difficult to determine in the timescale of a PFS. In addition, surrogate outcomes for surgical trials may be difficult to establish [202]. Short-term outcomes such as in-hospital or 30-day mortality, may, therefore, have limited value in determining a signal of promise for the true safety or effectiveness of surgical interventions.

Secondly, safety and effectiveness data from a PFS reported without any warning or caveat, may be accepted for publication as the definitive result, when the PFS is not powered to assess such outcomes. Because of this, it is advised that outcomes from PFS should either not be analysed at all, or if a suggestion of promise is sought, this should only be analysed and reported with a very clear caveat that the study was underpowered to test for significance. Not analysing PFS data may be especially important if there are plans for, or the potential to proceed to a main trial, as any analysis of outcomes may bias the results. Unfortunately, as is evident from the work in this thesis, there are ongoing issues with the definitive reporting of outcomes of surgical PFS, without appropriate caution during interpretation of results.

Whist there has been general acceptance in the literature for some time that any suggestion of promise or significance should be reported with caution, given the underpowered sample size of most PFS [72, 75, 92, 124, 125], it is important to note that historically, it has been justified as appropriate to use PFS efficacy data for assisting the calculation of a sample size for the main trial [82, 203]. In addition,

estimates of effect derived from PFS data have been used as a criterion to assess progression to a main trial [204]. More recently it has been explicitly proposed that estimating efficacy outcomes in PFS is inappropriate, as the sample size is rarely large enough to support or adequately refute a hypothesis [205]. Some propose that using estimates of effect from pilot and feasibility study data might, therefore, do more to *"mislead than to enlighten"* [205] , and should not be used in such important decisions as whether or not to proceed to a main trial [206], unless perhaps combined with more statistically robust evidence such as from a Bayesian decision model [207] .

Reporting results from a PFS as significant or definitive is, therefore, potentially inaccurate and misleading, both statistically and from inappropriate interpretation, because of misunderstanding about PFS more generally. It could be argued, therefore, that publicly reporting any suggestion of promise from a PFS is inappropriate, and new guidance must seek to fully clarify the appropriate handling of hypothesis testing in surgical PFS.

### NIHR Guidance

Perhaps in recognition of limitations of the existing advice from the NIHR on PFS, a guidance document was published [197] online on 18th June 2019 (after the data analysis for phases one and two of this work were complete). This relatively brief guidance offers applicants more information on designing PFS. The original NIHR definitions are included and concise guidance on which NIHR funding stream to apply to, both in terms of PFS grant size and how likely a definitive trial is in the immediate future. The guidance also stipulates that PFS should be distinguished from Phase II trials, as described above. However, whilst this more recent NIHR guidance [197], states that Phase II trials are different from PFS, it does not clarify why, and if and /or how any suggestion of promise of safety or effectiveness should be appropriately handled.

The most recent NIHR guidance [197], also offers advice on specifically applying to the RfPB funding stream for a PFS. Examples of key design parameters that a PFS may consider are listed, and it is stipulated that an outline of the proposed main

trial should be included. As part of the full trial proposal within the PFS application, applicants are requested to outline progression criteria. If the main trial is deemed feasible the guidance states applicants will be expected to write the main trial proposal within the timeframe of the PFS. Only three references on best practice for designing and conducting PFS are given [72, 74, 208] and these are all earlier references in the evolution of PFS methodology (published 9-14 years ago). Notably, this new NIHR guidance does not include reference to any of the more recent conceptual work [71, 132] or the complex interventions guidance from the MRC [13].

In addition, this guidance is only relevant for NIHR funding applications. The work in this thesis highlighted that if different funders adopt different definitions and guidance, this results in mixed messages and further confusion for study applicants. This variable and inconsistent message, consequently, serves to restrict applicants to fitting their study into a specific funder definition. As demonstrated, this further perpetuates the cycle of sub-optimisation, meaning PFS are not designed and conducted to optimally explore all the issues relevant to a specific study question.

In terms of the existing funding structure for surgical PFS, the NIHR have also published data on how this may be a barrier to completing PFS. Work conducted by the NIHR assessing the value of PFS was published in 2018 [209]. This work sought to identify the outcome of 89 completed RfPB funded feasibility studies (in all specialties) by sending questionnaires to the Principal Investigator of each study. The authors identified a mean time trajectory of 8 years from the start of the PFS to the completion of a main trial and definitive results, meaning that a PFS adds on average three years to achieving a definitive answer to a research question through a main trial. In fast moving clinical areas such as surgery, this represents another barrier to completing PFS, as an intervention may become obsolete before trials are complete, meaning PFS are avoided by surgeon trialists. The authors concluded [209] that whilst PFS may be avoiding waste through fewer trials failing secondary to poor research questions and trial design as previously

described [105], the failure of feasibility studies to progress to definitive studies in a time efficient manner, therefore, represents an alternative source of waste [209].

The study also considered the financial value of PFS in terms of saving monetary resources. Of the 89 feasibility studies, funded by the NIHR RfPB funding stream that had completed by May 2016, 57 studies judged the main trial feasible, 20 were deemed unfeasible and 12 had uncertain feasibility. The authors conclude that these 32 studies had saved in the region of £20-30 million of funding (the 12/89 studies that went on to achieve main trial funding had an average awarded grant of just over £1m per trial: mean £1,163,966, range £321,403 to £2,099,813). However, of the 57 studies that were judged feasible, 15 had secured further funding, 17 had been unsuccessful in securing further funding and 17 had not yet applied for further funding. So, whilst just over half of the successful feasibility studies had applied for further funding, less than half of these had secured it. Another key finding in this paper, therefore, is that an alternative source of research waste exists, when PFS show that a main trial is feasible, but fail to progress to a main trial. This could, therefore, be considered a fault of the current funding structure, and corroborates the findings in this thesis. However, one contradictory finding in the PEPSTAR study was that funders perceive PFS as a way of researching high-risk questions or clinical areas, of which surgery was perceived to be one in the PEPSTAR study. It may be challenging therefore, to alter the funding structure for research questions perceived to be high-risk, such as in surgery.

### A summary of the key recommendations from this work

Underpinning all the recommendations made from this work, is the urgent need for accessible and operationalised guidance for surgeons on the design and conduct of PFS. The creation of clear guidance, endorsed across funding and regulatory bodies and journals, would drive up the quality of PFS in surgery. Consequently, improved PFS, would lead a mechanism for change to a more efficient and thriving research structure and culture around surgical trials. The development of targeted and accessible guidance will bring key stakeholders

together to consider, and come to consensus, over what guidance is needed and how this should be delivered and endorsed to greatest effect.

The strengths and limitations of this work will now be discussed, followed by plans for future work to build on the findings in this thesis.

## 7.3 Strengths of this work

### 7.3.1 A novel contribution to the field of pilot and feasibility study methodology in the context of surgical trials

This is the first work to specifically consider the current research practice for PFS in surgery and the explicit challenges and barriers preventing optimal conduct of PFS in surgery. Producing surgery-specific recommendations for the optimal design and conduct of surgical PFS is also entirely novel.

A key strength of this project is the incorporation of a variety of different research methods to understand the potential for PFS to optimise future surgical trials. The rationale for this mixed methods approach was thoroughly described and justified in Chapter 3 (sections 3.10-3.12). Each component of this project involved researching a novel area of research practice, and sometimes used techniques that had not been employed previously. The targeted review and systematic analysis of NIHR funded PFS protocols, for example, was an innovative method for exploring current research practice and seeking out detailed understanding of how and why PFS are currently designed. Reviewing and analysing surgical PFS protocols, offered insights into research practice that could not have been achieved with a more traditional systematic or narrative review of the literature.

The PEPSTAR study involved the purposive sampling of participants, to consider the perceptions, experiences, views and opinions of key stakeholders involved in the design, conduct and funding of PFS in surgery. The in-depth qualitative interviews included participant surgeons, funders, trial methodologists and journal editors with the aim of maximum variation sampling

to consider the breadth of participant perspectives, experience, and research practice amongst different key stakeholder groups. The perspectives of these stakeholders have not been previously explored in the context of the design, conduct and funding of surgical PFS.

The synthesis and interpretation of the findings from the different phases of research, enhanced the breadth and strength of understanding regarding the difficulties facing PFS in surgery. The mixed methods approach therefore increased the validity and applicability of the results of this work. In addition, whilst surgical PFS are a relatively niche area in the broader realms of trials research, the results are likely to be more widely generalisable, further enhancing their impact, which is discussed next.

### 7.3.2 Applicability of the recommendations of this work to current practice and different audiences

Surgical trials have long been recognised as a challenging area of research and many of the unique and complex challenges of surgical research have been further highlighted and explored within the context of PFS for this thesis. The work to explore and understand how surgical PFS might be optimised in the future has been performed within the context of surgery, and the recommendations from this work are, therefore, specific to surgery. Whilst the recommendations from this work are focused on PFS in surgery, many may be relevant to the wider context of complex interventions as a whole. As part of a future consensus process (see section 7.5), this question could be addressed, especially in light of other similar work going on in other areas [210] (see also section 7.4.3 for details of this work). Most methodologists and funders have been and are involved in a wide range of clinical areas and are well placed to assess the potential for cross-pollination of useful strategies.

Examining the literature, it seems PFS may be less commonly done in countries outside the UK. A systematic review by the methodology group who produced the conceptual framework of the definitions of PFS, looked at the quality of reporting of 18 pilot and feasibility cluster randomised trials conducted

published between 2011 and 2014 [211]. This study found that half (56%) were set in the UK, with all other studies represented only once, apart from Canada (three studies) and the USA (two studies). In addition, it was noted in the PEPSTAR study, that the UK produced methodological guidance [71, 79] was perceived to have not been incorporated into practice by authors from overseas yet.

Whilst the focus of this work was entirely on research and funding practice in the UK, it is perhaps reasonable to tentatively conclude, that the UK is ahead of the curve in terms of methodological developments for the design and conduct of PFS, and that with further exploration and collaboration, the findings of this research could well be relevant to researchers in other countries.

## 7.4 Limitations of this work

### 7.4.1 Phase 1: A targeted review and systematic analysis of NIHR funded surgical pilot and feasibility studies

This targeted review and systematic analysis demonstrated some key limitations of PFS in surgery but focused solely on studies funded by the NIHR HTA and RfPB funding streams. While these represent two of the major funding bodies for surgical PFS in the UK, they do not include all possible funders of such work. Analysing PFS funded by NIHR funding streams, other than HTA and RfPB were considered initially. However, these other funders were excluded early on, as they fund surgical research far less commonly. To illustrate, the 2012 NIHR call for surgical research resulted in a number of 'surgical' studies being funded through a variety of NIHR funding streams [212]. A total of n=25 studies were funded as a result of this call:

**EME** (Efficacy and Mechanism Evaluation), which predominantly funds translational research: (n=1/25);

**HS&DR** (Health Services and Delivery Research), which funds research to produce evidence to impact on the quality, accessibility and organisation of health and social care services: (n=4/25);

**HTA** (Health Technology Assessment): (n=11/25)

**RfPB** (Research for Patient Benefit): (n=10/25);

**i4i** (Invention for Innovation), which is a translational funding scheme aimed for an advanced or clinically validated prototype medical device, technology or intervention: (n=3/25)

**PGfAR** (Programme Grants for Applied Research): Open to applications for surgical research but has predominantly funded applications relating to primary care historically: (n=0/25).

A review of these studies against the inclusion criteria for phase one of this work, showed that n=20/25 were not PFS, n=1 was not primarily surgical (assessing surgical outcomes for osteoarthritis of the knee following a short-term psychological intervention), leaving n=4/25 which are studies of surgical interventions and PFS, all of which were funded by RfPB, and all were included in our study.

Whilst the NIHR HTA and RfPB research programmes are the major funders of studies of surgical interventions in the UK, it is accepted there are other funders, for example the British Heart Foundation (BHF) and Arthritis UK (ARUK), which may also consider and fund such work. However, including work from other funders would have been logistically challenging, given that very few funders make study protocols publicly available. Another potential limitation of this work is that NIHR funded protocols, are likely to be of relatively high quality so may provide an overly positive perception of the quality of PFS performed in surgery more generally.

This work also focused only on so-called external PFS. It would also be informative to review internal PFS, but as stated in the methods (Chapter 3), trial experts are increasingly agreeing that internal pilots should not be termed 'pilot' studies at all, being so intrinsically part of the main trial, in terms of design, funding and outcome analysis. In addition, external PFS were considered to yield more relevant information for this thesis as this design is typically chosen when

there is greater uncertainty about main trial design and conduct. Despite these recognised limitations, this work provided valuable insights through identifying several challenges and areas of difficulty with surgical PFS design and conduct, which warranted further exploration.

Ideally, it would be useful to compare the findings of this work regarding PFS exploring surgical interventions, with those of non-surgical complex interventions. This comparison would give further insight into the differences in approach between these research areas and allow for cross specialty learning. Unfortunately, this was not feasible within the time limitations of this project but could be considered for future work.

### 7.4.2 Phase 2: The PEPSTAR Study

Seniority of participants in the PEPSTAR qualitative interviews was deemed important to ensure extensive experience in the area of PFS design and conduct, to allow the extraction of information about the challenges and barriers to conducting pre-trial research from those with the most extensive experience. However, a limitation of this study is in not also sampling less senior surgeons, for example, those leading the trainee surgical collaboratives. This group might have given different views of the current difficulties facing PFS in surgery.

Another limitation of the sampling strategy might have been in not sampling research-naïve surgeons for a comparison of knowledge, understanding and perceptions of the challenges in this area. However, throughout this thesis the long-standing issues with inappropriate reporting of both underpowered RCTs as PFS, and vice versa has been emphasised, indicating a widespread misunderstanding of the value and purpose of PFS in surgery. It was therefore, felt most important within the time and resource limitations of this PhD thesis, to concentrate on extracting data from the most experienced and data-rich sources. This stance was taken for both the targeted review and systematic analysis of NIHR funded PFS (as opposed to performing a traditional systematic or narrative literature review) and the PEPSTAR interviews (where the most senior and experienced participants were sampled).

Another limitation of this work might have been in considering only UK based participants for the interview study. It would certainly have added a further dimension to the qualitative work, to consider stakeholder participants from outside the UK. However, including participants in the PEPSTAR study from outside of the UK, might have added additional complexity to an already poorly understood area. Further complexity would potentially be created by the relatively different funding systems outside the UK, as well as a potential language barrier with participants at interview. Certainly, considering research practices around pre-trial work outside of the UK, might be an area for future work and collaboration.

### 7.4.3 Phase 3: A synthesis and interpretation of the findings in this thesis to inform recommendations for the optimisation of pilot and feasibility studies in surgery

There was not time within the scope and timeframes of this thesis to derive consensus-based recommendations, though this is considered for future work (see next section 7.5). This was partly because considerably more preparatory work was necessary to arrive at the recommendations presented in Chapter six than originally thought. As the work progressed, therefore, it became apparent that a consensus-based process to produce recommendations, would be out of the scope of this project, both in terms of time and resources.

This is a complex and contentious area of research methodology. The methodological research already performed to derive guidelines on the reporting PFS [79] and a conceptual framework of the definitions of PFS [71] took five years and a whole team of methodological experts. Similarly, in public health the GUEST (GUidance for Exploratory STudies of complex public health interventions) study funded by the MRC/NIHR Methodology Research Programme, aims to develop guidance on the design and conduct of PFS (termed exploratory studies) in complex public health interventions (Grant: MR/N015843/1, £246,955, awarded 2015/2016) [213]. Specifically, the research team planned a systematic

review of current guidance, an audit of current practice, a web based DELPHI exercise to identify expert consensus on the design and conduct of PFS studies and a horizon scan to identify approaches to intervention optimisation and PFS design from other contexts within and outside health care research. The systematic review from this work has now been published, concluding that existing guidance was inconsistent with little evidence available to inform when to proceed to a main trial [214]. The Delphi study has also completed [210], but guidance is not yet published. The GUEST study is another example of the work and resources needed to create guidance in a similar area of complex healthcare to surgery, and why creating consensus-based guidance for PFS in surgery was, therefore, outside the realms of possibility within this PhD.

However, the work in this thesis has concluded with a clear list of recommendations for how to take the findings forward. The planned dissemination pathway of this work and the recommendations produced from it, were discussed at the end of chapter six (section 6.5). The next section will discuss further work proposed to build on the recommendations from the work in this thesis.

## 7.5 Proposals for further work to implement the key recommendations from this thesis

### 7.5.1 New guidance

Ultimately, consensus-based recommendations are needed to allow this work to be incorporated into clear, operationalised guidelines, which would ideally be endorsed by major funders, journals and regulatory bodies. A consensus process would need to include representation from all key stakeholder groups including funders, methodologists/clinical trials units, journal editors, the IDEAL collaboration and surgeons.

Improved, accessible and operationalised guidance for surgeons and those applying for surgical PFS funding is recognised as key to improving surgical PFS. The new guidance needs to consider the methodological work done by

others, but most importantly this needs distilling and shaping into operationalised content for study applicants. Some key considerations around the content and processes for producing guidance in this area have been highlighted in this work.

### Key Stakeholders

To be implemented fully into practice, guidance needs to be derived from a consensus-based process including all key stakeholders. A list of key stakeholders to consider for inclusion at a consensus process are:

**Funders:** The NIHR RfPB and HTA funding programme are the major funders of surgical research in the UK. Representatives from these funding streams are essential. It would also be relevant to invite representatives from the PGfAR funding stream, as this is a funding stream that may become more relevant for a streamlined approach to conducting PFS. Other charity funders should be considered, though they fund PFS less commonly;

**Methodologists:** To include CTU directors, senior statisticians and those with an interest in surgical research and/or PFS;

**Pilot and Feasibility Studies Collaboration**: The group who produced the CONSORT extension for reporting PFS [79];

**NIHR Research Design Service (RDS)**: The RDS will be pivotal in dissemination and uptake of new guidance, so senior representation from this group is necessary at a consensus process to determine the level to pitch guidance at and what might be included;

**Journal editors**: Representation from journals who publish PFS and/or surgical studies for example, *Trials*, *Pilot and Feasibility Studies*, and *British Journal of Surgery*;

**Surgeons**: Particularly methodologically experienced surgeons (both junior and senior) for example those leading the RCS Surgical Trial Initiative, RCS Surgical Specialty Leads, Trainee Collaborative Leads, NIHR Surgical Clinician Scientists;

**IDEAL collaboration**: Senior representation from this group to make sure complementary messages are filtered through into other guidance documents.

### Endorsement

Endorsement by funders and relevant journals and regulatory bodies is vital to optimise future research practice in this area. Funders are key in producing and endorsing this guidance, as they drive the format and selection of research projects for funding. The current RfPB guidance [197] was written by the RfPB programme director, as a 'top tips' type of guide, with the NIHR definitions of PFS agreed across certain NIHR funding programmes (HTA, RfPB, EME, PHR). However, there was no input or consensus from other key stakeholders, and a need is identified to bring these groups together to produce accessible, accurate and operationalised guidance in order to improve pre-trial surgical research.

Some PEPSTAR participants highlighted potential problems with achieving funder-endorsed guidance, given that funding bodies do not currently agree on the role and purpose of PFS. A consensus process amongst all key stakeholders, including representatives from all major funders, journals and regulatory bodies would be paramount to overcoming this obstacle. This would probably also mean that the guidance would need to be broad and accessible to research practice across all specialties, but with specifics as necessary for areas such as surgery that have unique challenges.

### Key questions to consider in a consensus process to produce guidance

This work identified several important questions that should be considered as part of a consensus process to produce operationalised guidance for performing PFS in surgery. These are:

1) *Methodological input to PFS design and conduct*: This is clearly needed as part of PFS study design as a minimum. Consideration is also necessary for what this may consist of within the realms of PFS funding, and indeed for creating a funding application. The NIHR Research and Design Service (RDS) are well placed to assist at the initial stages of PFS design, but

appropriate input from methodological teams/Clinical Trials Units needs to be determined.

2) *Funding structure*: A need was identified for a more efficient and streamlined funding structure and the reality of creating this will be an important discussion point for the key stakeholders.
3) *Publication and dissemination of PFS*: The consensus-based process to produce guidance also needs to consider how to encourage improved rates of appropriate publication of PFS. The CONSORT extension to PFS offer comprehensive guidance on the appropriate reporting of PFS, but there is also a need for funders to require publication, perhaps as an incentive to receiving grant instalments. This may be as a publicly available report, or as a journal publication. With journal editors present as part of the stakeholder group, this will allow channels of communication for how PFS can be appropriately and consistently published and disseminated.

### 7.5.2 Education & training

Whilst improved consensus based guidance, incorporating the input of all key stakeholders is shown to be necessary from the work in this thesis, there is also an identified need to engage and educate the surgical community on the use of PFS in surgical research. The PEPSTAR study demonstrated that there had been relatively little engagement with the currently available guidance, such as the IDEAL framework. Mandatory training as part of surgical training pathways would assist with early collaborative engagement with surgical trials and allow those with a specific interest to develop their skills appropriately. Existing courses designed to engage both medical students and surgical trainees were discussed in section 6.3.1. As such courses are already in existence, further work would involve advocating for the mandatory training status of these courses, through engagement and collaboration with the Royal College of Surgeons and the Joint Committee on Surgical Training (JCST). In particular, involvement with surgical trials in terms of recruitment for example, rather than first author

publication in quantity rather than quality, should become the primary focus of research training for surgeons.

### 7.5.3 Wider collaboration

Improved collaborative practice between surgeons and methodologists/CTUs was given as one of four key recommendations for improving research practice around PFS in surgery (see section 6.3.2). Publishing these recommendations and highlighting the need for improved collaboration alone as is planned (see section 6.5), will not suffice to improve practice. The practicalities of improving collaboration dictate that further work is needed to showcase the available methodological support and opportunities to surgeons interested in trials, and better collaborative models of working need to be developed. These might include, for example, the Royal College of Surgeons Surgical Specialty Lead roles, being mutually beneficial and collaborative positions between surgeons and methodologists.

This work was presented and well received at ICTMC (see section 6.5), where a collaboration lead by the RfPB to produce guidance for the design and conduct of PFS more generally was formed and instigated. The challenge for this future work is to advocate for the nuances and specific needs of surgical research and ensure that the surgical 'voice' is not lost.

## 7.6 Conclusion

It is increasingly acknowledged that PFS are a methodological solution that can be utilised to explore, and address uncertainties surrounding the design and conduct of a future main trial. The unique and complex challenges that have been proven to obstruct the conduct and completion of surgical trials may, therefore, be usefully explored in PFS, to increase the chances of main trial success. This work has addressed a gap in the clinical trials research field, by providing a detailed understanding of the potential for PFS to optimise future surgical trials and by presenting clear recommendations for surgeons and research teams.

The mixed methods approach taken to achieve the objectives of this work, allowed comprehensive exploration of the challenges and barriers obstructing the optimal design and conduct of PFS in surgery. A synthesis and interpretation of quantitative and qualitative findings from the first two phases of work revealed that misunderstanding, sub-optimal design, conduct, reporting and consequent devaluation of PFS are perpetuated in a cyclical model of sub-optimal practice amongst the surgical research community. The members of this surgical research community include methodologists, funders, academic institutions, journal editors and surgeons. Confounding issues further preventing optimal surgical PFS design and conduct, include the cultural challenges inherently surrounding surgical trials and a lack of targeted and accessible guidance.

To address this model of sub-optimisation of PFS, a set of recommendations relevant to all key stakeholders was produced delineating the key approaches to optimise the design and conduct of future PFS. In addition, to operationalise these recommendations, a 'Top Tips' tool was developed for use by surgeons designing and conducting surgical PFS. At the foundation of all the recommendations made from this work, is the need for guidance that is specifically operationalised and accessible to surgery as an area of research with known complexity and challenges.

This work has demonstrated that well-performed PFS, as part of a more streamlined and efficient research process, will enhance the success of surgical main trials and optimise the use of limited funding resources, therefore providing a solid foundation for the high-quality evidence base necessary to inform future surgical practice. Future work is directed towards developing consensus-based operationalised guidance for surgeons designing and conducting PFS, recognising that this guidance could also be more widely applicable to other complex interventions.

# *References*


1. Evidence-Based Medicine Working G. Evidence-based medicine. A new approach to teaching the practice of medicine. *Jama* 1992;**268**(17): 2420-2425.

2. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. *Evid Based Med* 2016;**21**(4): 125-127.

3. Greenhalgh T, Howick J, Maskrey N, Group EBMR. Evidence based medicine: a movement in crisis? *Bmj* 2014;**348**: g3725.

4. Horton R. Surgical research or comic opera: questions, but few answers. *Lancet* 1996;**347**(9007): 984-985.

5. Lane R. Jane Blazeby: passionate leader in evidence-based surgery. *Lancet* 2016;**388**(10052): 1367.

6. Bell P.R.F CD, Ruckley C.V, Russell R.C.G, Grant A, Kennedy, A, McCulloch P, John, L.C.H, Black N. Letters to the Editor. *The Lancet* 1996;**34**(9013): 1479-1482.

7. Barkun JS, Aronson JK, Feldman LS, Maddern GJ, Strasberg SM, Altman DG, Barkun JS, Blazeby JM, Boutron IC, Campbell WB, Clavien PA, Cook JA, Ergina PL, Flum DR, Glasziou P, Marshall JC, McCulloch P, Nicholl J, Reeves BC, Seiler CM, Meakins JL, Ashby D, Black N, Bunker J, Burton M, Campbell M, Chalkidou K, Chalmers I, de Leval M, Deeks J, Grant A, Gray M, Greenhalgh R, Jenicek M, Kehoe S, Lilford R, Littlejohns P, Loke Y, Madhock R, McPherson K, Rothwell P, Summerskill B, Taggart D, Tekkis P, Thompson M, Treasure T, Trohler U, Vandenbroucke J. Evaluation and stages of surgical innovations. *Lancet* 2009;**374**(9695): 1089-1096.

8. Ergina PL, Cook JA, Blazeby JM, Boutron I, Clavien PA, Reeves BC, Seiler CM, Altman DG, Aronson JK, Barkun JS, Campbell WB, Cook JA, Feldman LS, Flum DR, Glasziou P, Maddern GJ, Marshall JC, McCulloch P, Nicholl J, Strasberg SM, Meakins JL, Ashby D, Black N, Bunker J, Burton M, Campbell M, Chalkidou K, Chalmers I, de Leval M, Deeks J, Grant A, Gray M, Greenhalgh R, Jenicek M, Kehoe S, Lilford R, Littlejohns P, Loke Y, Madhock R, McPherson K, Rothwell P, Summerskill B, Taggart D, Tekkis P, Thompson M, Treasure T, Trohler U, Vandenbroucke J. Challenges in evaluating surgical innovation. *Lancet* 2009;**374**(9695): 1097-1104.

9. McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC, Nicholl J, Aronson JK, Barkun JS, Blazeby JM, Boutron IC, Campbell WB, Clavien PA, Cook JA, Ergina PL, Feldman LS, Flum DR, Maddern GJ, Nicholl J,

Reeves BC, Seiler CM, Strasberg SM, Meakins JL, Ashby D, Black N, Bunker J, Burton M, Campbell M, Chalkidou K, Chalmers I, de Leval M, Deeks J, Ergina PL, Grant A, Gray M, Greenhalgh R, Jenicek M, Kehoe S, Lilford R, Littlejohns P, Loke Y, Madhock R, McPherson K, Meakins J, Rothwell P, Summerskill B, Taggart D, Tekkis P, Thompson M, Treasure T, Trohler U, Vandenbroucke J. No surgical innovation without evaluation: the IDEAL recommendations. *Lancet* 2009;**374**(9695): 1105-1112.

10. McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *Bmj* 2002;**324**(7351): 1448-1451.

11. National Cancer Research Institute. Challenges and opportunities in surgical cancer research in the UK. London: National Cancer Research Institute; 2012.

12. Cook JA. The challenges faced in the design, conduct and analysis of surgical randomised controlled trials. *Trials* 2009;**10**: 9.

13. Medical Research Council. Developing and evaluating complex interventions: New guidance. London: Medical Research Council; 2008.

14. Blencowe NS, Brown JM, Cook JA, Metcalfe C, Morton DG, Nicholl J, Sharples LD, Treweek S, Blazeby JM, Members of the MRCHfTMRNW. Interventions in randomised controlled trials in surgery: issues to consider during trial design. *Trials* 2015;**16**: 392.

15. Blencowe NS, Boddy AP, Harris A, Hanna T, Whiting P, Cook JA, Blazeby JM. Systematic review of intervention design and delivery in pragmatic and explanatory surgical randomized clinical trials. *The British journal of surgery* 2015;**102**(9): 1037-1047.

16. Weijer C, Enkin MW, Shapiro SH, Glass KC. Clinical equipoise and not the uncertainty principle is the moral underpinning of the randomised controlled trial. *Bmj* 2000;**321**(7263): 756-758.

17. Donovan JL, de Salis I, Toerien M, Paramasivan S, Hamdy FC, Blazeby JM. The intellectual challenges and emotional consequences of equipoise contributed to the fragility of recruitment in six randomized controlled trials. *Journal of clinical epidemiology* 2014;**67**(8): 912-920.

18. Donovan JL, Paramasivan S, de Salis I, Toerien M. Clear obstacles and hidden challenges: understanding recruiter perspectives in six pragmatic randomised controlled trials. *Trials* 2014;**15**: 5.

19. Rooshenas L, Elliott D, Wade J, Jepson M, Paramasivan S, Strong S, Wilson C, Beard D, Blazeby JM, Birtle A, Halliday A, Rogers CA, Stein R, Donovan JL, group A-s, group B-B-Ss, group Cs, group Cs, group Ops, group Ps. Conveying Equipoise during Recruitment for Clinical Trials: Qualitative Synthesis of

Clinicians' Practices across Six Randomised Controlled Trials. *PLoS medicine* 2016;**13**(10): e1002147.

20. Potter S, Mills N, Cawthorn SJ, Donovan J, Blazeby JM. Time to be BRAVE: is educating surgeons the key to unlocking the potential of randomised clinical trials in surgery? A qualitative study. *Trials* 2014;**15**: 80.

21. Harrop E, Kelly J, Griffiths G, Casbard A, Nelson A, Published on behalf of the BTMG. Why do patients decline surgical trials? Findings from a qualitative interview study embedded in the Cancer Research UK BOLERO trial (Bladder cancer: Open versus Lapararoscopic or RObotic cystectomy). *Trials* 2016;**17**: 35.

22. Winters ZE, Griffin C, Horne R, Bidad N, McCulloch P. Barriers to accrue to clinical trials and possible solutions. *British journal of cancer* 2014;**111**(4): 637-639.

23. Reed MW, Wyld L, Ellis P, Bliss J, Leonard R, Action, Groups ESTM. Breast cancer in older women: trials and tribulations. *Clin Oncol (R Coll Radiol)* 2009;**21**(2): 99-102.

24. Elliott D, Husbands S, Hamdy FC, Holmberg L, Donovan JL. Understanding and Improving Recruitment to Randomised Controlled Trials: Qualitative Research Approaches. *European urology* 2017;**72**(5): 789-798.

25. Donovan JL, Rooshenas L, Jepson M, Elliott D, Wade J, Avery K, Mills N, Wilson C, Paramasivan S, Blazeby JM. Optimising recruitment and informed consent in randomised controlled trials: the development and implementation of the Quintet Recruitment Intervention (QRI). *Trials* 2016;**17**(1): 283.

26. Donovan JL, Opmeer B, Young GJ, Mills N, Martin RM, Lane JA, Metcalfe C, Peters TJ, Davis M, Turner EL, Walsh E, Neal DE, Hamdy FC, Group PS. Factors associated with trial recruitment, preferences, and treatments received were elucidated in a comprehensive cohort study. *J Clin Epidemiol* 2019;**113**: 200-213.

27. Sedrakyan A, Campbell B, Graves S, Cronenwett JL. Surgical registries for advancing quality and device surveillance. *Lancet* 2016;**388**(10052): 1358-1360.

28. Sedrakyan A, Campbell B, Merino JG, Kuntz R, Hirst A, McCulloch P. IDEAL-D: a rational framework for evaluating and regulating the use of medical devices. *Bmj* 2016;**353**: i2372.

29. MJ B. Problems in the economic appraisal of new health technology: the evaluation of heart transplants in the UK. In: *Economic Appraisal of Health Technology in the European Community* MF D (ed). Oxford Medical Publications: Oxford, 1987; 103-118.

30. Harvey KL, Clark SE. A guide to breast implants for the non-breast specialist. *Womens Health* 2016;**12**(6): 533-537.

31. Smith AJ, Dieppe P, Vernon K, Porter M, Blom AW, Wales NJRE. Failure rates of stemmed metal-on-metal hip replacements: analysis of data from the National Joint Registry of England and Wales. *Lancet* 2012;**379**(9822): 1199-1204.

32. Heneghan CJ, Goldacre B, Onakpoya I, Aronson JK, Jefferson T, Pluddemann A, Mahtani KR. Trials of transvaginal mesh devices for pelvic organ prolapse: a systematic database review of the US FDA approval process. *BMJ open* 2017;**7**(12): e017125.

33. Angelos P. The ethical challenges of surgical innovation for patient care. *Lancet* 2010;**376**(9746): 1046-1047.

34. The Mid Staffordshire NHS Foundation Trsut Public Inquiry. http://webarchive.nationalarchives.gov.uk/20150407084003/http://www.midstaffspublicinquiry.com/ [27th September 2017].

35. National Cancer Research Institute. Trials are only as credible as their endpoints: Defining the future outcomes of surgical research; July 2016.

36. Blencowe NS, Mills N, Cook JA, Donovan JL, Rogers CA, Whiting P, Blazeby JM. Standardizing and monitoring the delivery of surgical interventions in randomized clinical trials. *The British journal of surgery* 2016;**103**(10): 1377-1384.

37. Cook JA, Ramsay CR, Fayers P. Statistical evaluation of learning curve effects in surgical trials. *Clinical trials* 2004;**1**(5): 421-427.

38. Hawasli A, Lloyd LR. Laparoscopic cholecystectomy. The learning curve: report of 50 patients. *The American surgeon* 1991;**57**(8): 542-544; discussion 545.

39. Lee DY, Oh DJ, Kang KR, Kim MS, Oh KH, Baek SK, Kwon SY, Woo JS, Jung KY. Comparison of Learning Curves for Retroauricular and Transaxillary Endoscopic Hemithyroidectomy. *Annals of surgical oncology* 2016;**23(**12): 4023-4028..

40. Cook JA, Ramsay CR, Carr AJ, Rees JL, group Ut. A questionnaire elicitation of surgeons' belief about learning within a surgical trial. *PloS one* 2012;**7**(11): e49178.

41. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, Altman DG, Barbour V, Macdonald H, Johnston M, Lamb SE, Dixon-Woods M, McCulloch P, Wyatt JC, Chan AW, Michie S. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Bmj* 2014;**348**: g1687.

42. http://www.consort-statement.org/ [16th August 2016].

43. http://www.spirit-statement.org/ [27th August 2019].

44. Lewis S, Warlow C. How to spot bias and other potential problems in randomised controlled trials. *Journal of Neurology Neurosurgery and Psychiatry* 2004;**75**(2): 181-187.

45. Karanicolas PJ, Farrokhyar F, Bhandari M. Practical tips for surgical research: blinding: who, what, when, why, how? *Canadian journal of surgery Journal canadien de chirurgie* 2010;**53**(5): 345-348.

46. Boutron I, Guittet L, Estellat C, Moher D, Hrobjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS medicine* 2007;**4**(2): e61.

47. Probst P, Grummich K, Heger P, Zaschke S, Knebel P, Ulrich A, Buchler MW, Diener MK. Blinding in randomized controlled trials in general and abdominal surgery: protocol for a systematic review and empirical study. *Systematic reviews* 2016;**5**: 48.

48. Bluebelle Study G, Reeves BC, Andronis L, Blazeby JM, Blencowe NS, Calvert M, Coast J, Draycott T, Donovan JL, Gooberman-Hill R, Longman RJ, Magill L, Mathers JM, Pinkney TD, Rogers CA, Rooshenas L, Torrance A, Welton NJ, Woodward M, Ashton K, Bera KD, Clayton GL, Culliford LA, Dumville JC, Elliott D, Ellis L, Gould-Brown H, Macefield RC, McMullan C, Pope C, Siassakos D, Strong S, Talbot H. A mixed-methods feasibility and external pilot study to inform a large pragmatic randomised controlled trial of the effects of surgical wound dressing strategies on surgical site infections (Bluebelle Phase B): study protocol for a randomised controlled trial. *Trials* 2017;**18**(1): 401.

49. Biere SS, van Berge Henegouwen MI, Maas KW, Bonavina L, Rosman C, Garcia JR, Gisbertz SS, Klinkenbijl JH, Hollmann MW, de Lange ES, Bonjer HJ, van der Peet DL, Cuesta MA. Minimally invasive versus open oesophagectomy for patients with oesophageal cancer: a multicentre, open-label, randomised controlled trial. *Lancet* 2012;**379**(9829): 1887-1892.

50. Mariette C, Robb WB. Minimally invasive versus open oesophagectomy for oesophageal cancer. *Lancet* 2012;**380**(9845): 883; author reply 885-886.

51. Parotto M, Valenza F, Ori C, Spieth PM. Minimally invasive versus open oesophagectomy for oesophageal cancer. *Lancet* 2012;**380**(9845): 884; author reply 885-886.

52. Spengler CM, Verges S, Walder B. Minimally invasive versus open oesophagectomy for oesophageal cancer. *Lancet* 2012;**380**(9845): 885; author reply 885-886.

53. Swisher S, Ajani J, Correa A, Komaki R, Hofstetter W. Minimally invasive versus open oesophagectomy for oesophageal cancer. *Lancet* 2012;**380**(9845): 883; author reply 885-886.

54. Bruce J, Russell EM, Mollison J, Krukowski ZH. The measurement and monitoring of surgical adverse events. *Health technology assessment* 2001;**5**(22): 1-194.

55. Cornish JA, Tilney HS, Tan E, Thompson MR, Smith JJ, Tekkis PP. The National Bowel Cancer Audit Project: what do trusts think of the National Bowel Cancer Audit and how can it be improved? *Tech Coloproctol* 2011;**15**(1): 53-59.

56. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Annals of surgery* 2004;**240**(2): 205-213.

57. https://www.invo.org.uk/ [27th November 2019].

58. COMET Initiative. http://www.comet-initiative.org/ [5th September 2016].

59. Macefield R, Nicholson A, Milne T, Pinkney T, Calvert M, Avery K, Reeves B, Blazeby J. Development of a single, practical measure of surgical site infection (SSI) for patient self-report and observer completion: a novel method of questionnaire design. *Trials* 2015;**16**(Suppl 2): O6.

60. https://www.methodologyhubs.mrc.ac.uk/about/hubs/ [22nd June 2017].

61. http://www.bristol.ac.uk/social-community-medicine/centres/conduct2/ [22nd June 2017].

62. https://www.rcseng.ac.uk/standards-and-research/research/surgical-trials-initiative/ [19th June 2017].

63. Royal College of Surgeons of England. Surgical Research Report 2017-2018; 2017.

64. Jamjoom AA, Phan PN, Hutchinson PJ, Kolias AG. Surgical trainee research collaboratives in the UK: an observational study of research activity and publication productivity. *BMJ open* 2016;**6**(2): e010374.

65. Pinkney TD, Calvert M, Bartlett DC, Gheorghe A, Redman V, Dowswell G, Hawkins W, Mak T, Youssef H, Richardson C, Hornby S, Magill L, Haslop R, Wilson S, Morton D, Collaborative WMR, Investigators RT. Impact of wound edge protection devices on surgical site infection after laparotomy: multicentre randomised controlled trial (ROSSINI Trial). *Bmj* 2013;**347**: f4305.

66. Blencowe NS, Cook JA, Pinkney T, Rogers C, Reeves BC, Blazeby JM. Delivering successful randomized controlled trials in surgery: Methods to optimize collaboration and study design. *Clin Trials* 2017;**14**(2): 211-218.

67. Strong S, Paramasivan S, Mills N, Wilson C, Donovan JL, Blazeby JM. 'The trial is owned by the team, not by an individual': a qualitative study exploring

the role of teamwork in recruitment to randomised controlled trials in surgical oncology. *Trials* 2016;**17**(1): 212.

68. Royal College of Surgeons of England. From theory to theatre: Overcoming barriers to innovation in surgery; 2011.

69. Oxford English Dictionary: Oxford University Press.

70. O'Cathain A, Thomas KJ, Drabble SJ, Rudolph A, Goode J, Hewison J. Maximising the value of combining qualitative research and randomised controlled trials in health research: the QUAlitative Research in Trials (QUART) study--a mixed methods study. *Health technology assessment* 2014;**18**(38): 1-197, v-vi.

71. Eldridge SM, Lancaster GA, Campbell MJ, Thabane L, Hopewell S, Coleman CL, Bond CM. Defining Feasibility and Pilot Studies in Preparation for Randomised Controlled Trials: Development of a Conceptual Framework. *PLoS One* 2016;**11**(3): e0150205.

72. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. *Journal of evaluation in clinical practice* 2004;**10**(2): 307-312.

73. Arnold DM, Webert KE, Adhikari NK, Cook DJ. Pilot trials in transfusion medicine. *Transfusion* 2009;**49**(7): 1293-1295.

74. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC medical research methodology* 2010;**10**: 67.

75. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, Robson R, Thabane M, Giangregorio L, Goldsmith CH. A tutorial on pilot studies: the what, why and how. *BMC medical research methodology* 2010;**10**: 1.

76. Jeray KJ, Tanner SL. Pilot randomized trials: is there a need? *The Journal of bone and joint surgery American volume* 2012;**94 Suppl 1**: 15-18.

77. http://www.invo.org.uk/posttypejargon/pilot-studies/ [20th June 2016].

78. http://www.invo.org.uk/posttypejargon/feasibility-studies/ [20th June 2016].

79. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, Lancaster GA, group Pc. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Bmj* 2016;**355**: i5239.

80. Thabane L, Hopewell S, Lancaster GA, Bond CM, Coleman CL, Campbell MJ, Eldridge SM. Methods and processes for development of a CONSORT

extension for reporting pilot randomized controlled trials. *Pilot and Feasibility Studies* 2016;**2**(1): 1-13.

81. Eldridge S BC, Campbell M, Hopewell S, Thabane L, Lancaster G, Coleman C. Defining feasibility and pilot studies in preparation for randomised controlled trials: using consensus methods and validation to develop a conceptual framework. *Trials* 2015;**16**(Suppl 2): 087.

82. Eldridge SM, Costelloe CE, Kahan BC, Lancaster GA, Kerry SM. How big should the pilot study for my cluster randomised trial be? *Statistical methods in medical research* 2016;**25**(3): 1039-1056.

83. Avery K WP, Gamble C, O'Connell Francischeto E, Metcalfe C, DAvidson P, Blazxeby J. Optimising the design of internal pilot work to inform efficient randomised controlled trials: issues to consider when developing progression criteria. *Trials* 2015;**16**(Suppl 2): P10.

84. Avery KN, Williamson PR, Gamble C, O'Connell Francischetto E, Metcalfe C, Davidson P, Williams H, Blazeby JM, members of the Internal Pilot Trials Workshop supported by the Hubs for Trials Methodology R. Informing efficient randomised controlled trials: exploration of challenges in developing progression criteria for internal pilot studies. *BMJ open* 2017;**7**(2): e013537.

85. Cook JA, Beard DJ, Cook JR, MacLennan GS. The curious case of an internal pilot in a multicentre randomised trial-time for a rethink? *Pilot Feasibility Studies* 2016;**2**: 73.

86. Leuck AM, Johnson JR, Hunt MA, Dhody K, Kazempour K, Ferrieri P, Kline S. Safety and efficacy of a novel silver-impregnated urinary catheter system for preventing catheter-associated bacteriuria: a pilot randomized clinical trial. *American journal of infection control* 2015;**43**(3): 260-265.

87. Romero-Cerecero O, Zamilpa A, Tortoriello J. Pilot study that evaluated the clinical effectiveness and safety of a phytopharmaceutical elaborated with an extract of Ageratina pichinchensis in patients with minor recurrent aphthous stomatitis. *Journal of ethnopharmacology* 2015;**173**: 225-230.

88. Woodhall SC, Nichols T, Alexander S, da Silva FC, Mercer CH, Ison C, Gill ON, Soldan K. Can we use postal surveys with anonymous testing to monitor chlamydia prevalence in young women in England? Pilot study incorporating randomised controlled trial of recruitment methods. *Sexually transmitted infections* 2015;**91**(6): 412-414.

89. Shanyinde M, Pickering RM, Weatherall M. Questions asked and answered in pilot and feasibility randomized controlled trials. *BMC medical research methodology* 2011;**11**: 117.

90. Eldridge S BC, Campbell M, Lancaster G, Thabane L, Hopewell S. Definition and reporting of pilot and feasibility studies. *Trials* 2013;**14** (Suppl 1): O18.

91. Van Teijlingen ER, Rennie AM, Hundley V, Graham W. The importance of conducting and reporting pilot studies: the example of the Scottish Births Survey. *Journal of advanced nursing* 2001;**34**(3): 289-295.

92. Lancaster GA, Campbell MJ, Eldridge S, Farrin A, Marchant M, Muller S, Perera R, Peters TJ, Prevost AT, Rait G. Trials in primary care: statistical issues in the design, conduct and evaluation of complex interventions. *Statistical methods in medical research* 2010;**19**(4): 349-377.

93. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. *Journal of psychiatric research* 2011;**45**(5): 626-629.

94. Bugge C, Williams B, Hagen S, Logan J, Glazener C, Pringle S, Sinclair L. A process for Decision-making after Pilot and feasibility Trials (ADePT): development following a feasibility study of a complex intervention for pelvic organ prolapse. *Trials* 2013;**14**: 353.

95. Kistin C, Silverstein M. Pilot Studies: A Critical but Potentially Misused Component of Interventional Research. *Jama* 2015;**314**(15): 1561-1562.

96. Tickle-Degnen L. Nuts and bolts of conducting feasibility studies. The American journal of occupational therapy : official publication of the American Occupational Therapy Association 2013;**67**(2): 171-176.

97. Morin KH. Value of a pilot study. *The Journal of nursing education* 2013;**52**(10): 547-548.

98. Prospero Registry. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=47843. [30th April 2018]

99. Medical Research Council. http://gtr.rcuk.ac.uk/projects?ref=MR%2FN015843%2F1 [30th April 2018]

100. IDEAL Collaboration. http://www.ideal-collaboration.net/ [22nd June 2016].

101. McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC, Nicholl J, Balliol C, Aronson JK, Barkun JS, Blazeby JM, Boutron IC, Campbell WB, Clavien PA, Cook JA, Ergina PL, Feldman LS, Flum DR, Maddern GJ, Nicholl J, Reeves BC, Seiler CM, Strasberg SM, Meakins JL, Ashby D, Black N, Bunker J, Burton M, Campbell M, Chalkidou K, Chalmers I, de Leval M, Deeks J, Ergina PL, Grant A, Gray M, Greenhalgh R, Jenicek M, Kehoe S, Lilford R, Littlejohns P, Loke Y, Madhock R, McPherson K, Meakins J, Rothwell P,

Summerskill B, Taggart D, Tekkis P, Thompson M, Treasure T, Trohler U, Vandenbroucke J. No surgical innovation without evaluation: the IDEAL recommendations. *Lancet* 2009;**374**(9695): 1105-1112.

102. Hirst A, Philippou Y, Blazeby J, Campbell B, Campbell M, Feinberg J, Rovers M, Blencowe N, Pennell C, Quinn T, Rogers W, Cook J, Kolias AG, Agha R, Dahm P, Sedrakyan A, McCulloch P. No Surgical Innovation Without Evaluation: Evolution and Further Development of the IDEAL Framework and Recommendations. *Annals of surgery* 2019;269(2): 211-220.

103. McCulloch P, Feinberg J, Philippou Y, Kolias A, Kehoe S, Lancaster G, Donovan J, Petrinic T, Agha R, Pennell C. Progress in clinical research in surgery and IDEAL. *Lancet* 2018;**392**(10141): 88-94.

104. Al-Shahi Salman R, Beller E, Kagan J, Hemminki E, Phillips RS, Savulescu J, Macleod M, Wisely J, Chalmers I. Increasing value and reducing waste in biomedical research regulation and management. *Lancet* 2014;**383**(9912): 176-185.

105. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gulmezoglu AM, Howells DW, Ioannidis JP, Oliver S. How to increase value and reduce waste when research priorities are set. *Lancet* 2014;**383**(9912): 156-165.

106. Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gotzsche PC, Krumholz HM, Ghersi D, van der Worp HB. Increasing value and reducing waste: addressing inaccessible research. *Lancet* 2014;**383**(9913): 257-266.

107. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, Michie S, Moher D, Wager E. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;**383**(9913): 267-276.

108. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, Schulz KF, Tibshirani R. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;**383**(9912): 166-175.

109. Chen R, Desai NR, Ross JS, Zhang W, Chau KH, Wayda B, Murugiah K, Lu DY, Mittal A, Krumholz HM. Publication and reporting of clinical trial results: cross sectional analysis across academic medical centers. *Bmj* 2016;**352**: i637.

110. Jones CW, Handler L, Crowell KE, Keil LG, Weaver MA, Platts-Mills TF. Non-publication of large randomized clinical trials: cross sectional analysis. *Bmj* 2013;**347**: f6104.

111. Fairhurst K AK, O'Connell Francischetto E, Metcalfe C, Blazeby J. How can pilot work optimally inform surgical RCTs? A review of current evidence. *Trials* 2015;**16(Suppl 2)**: 17.

112. O'Connell Francischetto E AK, Metcalfe C, Williamson P, Gamble C, Blazeby J. Optimising the design and evaluation of pilot work to inform the main

trial: a review of current evidence and consideration of future practices. *Trials* 2013;**14**(Suppl 1): O17.

113. Cook D, Meade M, Guyatt G, Walter SD, Heels-Ansdell D, Geerts W, Warkentin TE, Cooper DJ, Zytaruk N, Vallance S, Berwanger O, Rocha M, Qushmaq I, Crowther M. PROphylaxis for ThromboEmbolism in Critical Care Trial protocol and analysis plan. *Journal of critical care* 2011;**26**(2): 223 e221-229.

114. Cook DJ, Rocker G, Meade M, Guyatt G, Geerts W, Anderson D, Skrobik Y, Hebert P, Albert M, Cooper J, Bates S, Caco C, Finfer S, Fowler R, Freitag A, Granton J, Jones G, Langevin S, Mehta S, Pagliarello G, Poirier G, Rabbat C, Schiff D, Griffith L, Crowther M, Investigators P, Canadian Critical Care Trials G. Prophylaxis of Thromboembolism in Critical Care (PROTECT) Trial: a pilot study. *Journal of critical care* 2005;**20**(4): 364-372.

115. Audrey S, Cooper AR, Hollingworth W, Metcalfe C, Procter S, Davis A, Campbell R, Gillison F, Rodgers SE. Study protocol: the effectiveness and cost effectiveness of an employer-led intervention to increase walking during the daily commute: the Travel to Work randomised controlled trial. *BMC public health* 2015;**15**: 154.

116. Audrey S, Procter S, Cooper A, Mutrie N, Hollingworth W, Davis A, Kipping R, Insall P, Garfield K, Campbell R. Employer schemes to encourage walking to work: feasibility study incorporating an exploratory randomised controlled trial. *Public Health Research* 2015;**3**(4).

117. Reinforcement of Closure of Stoma Site C, the West Midlands Research C. Feasibility study from a randomized controlled trial of standard closure of a stoma site vs biological mesh reinforcement. *Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland* 2016;**18**(9): 889-896.

118. Reinforcement of Closure of Stoma Site C, the West Midlands Research C. Randomized controlled trial of standard closure of a stoma site vs biological mesh reinforcement: study protocol of the ROCSS trial. *Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland* 2018;**20**(2): O46-O54.

119. Berents TL, Ronnevig J, Soyland E, Gaustad P, Nylander G, Loland BF. Topical treatment with fresh human milk versus emollient on atopic eczema spots in young children: a small, randomized, split body, controlled, blinded pilot study. *BMC dermatology* 2015;**15**: 7.

120. Juhasz M, Nagy VL, Szekely H, Kocsis D, Tulassay Z, Laszlo JF. Influence of inhomogeneous static magnetic field-exposure on patients with erosive gastritis: a randomized, self- and placebo-controlled, double-blind, single centre, pilot study. *Journal of the Royal Society, Interface / the Royal Society* 2014;**11**(98): 20140601.

121. Perera AP, Howell AM, Sodergren MH, Farne H, Darzi A, Purkayastha S, Paraskeva P. A pilot randomised controlled trial evaluating postoperative packing of the perianal abscess. *Langenbeck's archives of surgery / Deutsche Gesellschaft fur Chirurgie* 2015;**400**(2): 267-271.

122. Wijayasinghe N, Andersen KG, Kehlet H. Analgesic and Sensory Effects of the Pecs Local Anesthetic Block in Patients with Persistent Pain after Breast Cancer Surgery: A Pilot Study. *Pain Pract* 2016;**17**(2): 185-191

123. Perera T, Lewis PM, Davidson AJ, Junor P, Bottrell S. A pilot study to determine whether visually evoked hemodynamic responses are preserved in children during inhalational anesthesia. *Paediatric anaesthesia* 2015;**25**(3): 317-326.

124. Kraemer HC, Mintz J, Noda A, Tinklenberg J, Yesavage JA. Caution regarding the use of pilot studies to guide power calculations for study proposals. *Archives of general psychiatry* 2006;**63**(5): 484-489.

125. Arnold DM, Burns KE, Adhikari NK, Kho ME, Meade MO, Cook DJ. The design and interpretation of pilot trials in clinical research in critical care. *Critical care medicine* 2009;**37**(Suppl 1): S69-74.

126. Koizumi M, Koyama M, Tada K, Nishimura S, Miyagi Y, Makita M, Yoshimoto M, Iwase T, Horii R, Akiyama F, Saga T. The feasibility of sentinel node biopsy in the previously treated breast. *European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology* 2008;**34**(4): 365-368.

127. Jung HS, Kim HR, Choi SH, Kim YH, Kim DK, Park SI. Clinical feasibility and efficacy of video-assisted thoracic surgery (VATS) anatomical resection in patients with central lung cancer: a comparison with thoracotomy. *Journal of thoracic disease* 2015;**7**(10): 1774-1779.

128. Loscalzo J. Pilot trials in clinical research: of what value are they? *Circulation* 2009;**119**(13): 1694-1696.

129. HRA and Research Transparency: Clinical trial registration. In. Version 4.2 August 2015 ed; 2015.

130. Friedman L. Commentary: Why we should report results from clinical trial pilot studies. *Trials* 2013;**14**: 14.

131. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Bmj* 2010;**340**: c332.

132. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, Lancaster GA, group Pc. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Pilot Feasibility Stud* 2016;**2**: 64.

133. http://pilotfeasibilitystudies.biomedcentral.com/ [16th August 2016].

134. Lancaster GA. Pilot and feasibility studies come of age! *Pilot and feasibility studies* 2015;**1**(1): 1-4.

135. Hodgkinson GP, Ford JK. Narrative, meta-analytic, and systematic reviews: What are the differences and why do they matter? *J Organ Behav* 2014;**35**: S1-S5.

136. Oxford English Dictionary. http://www.oed.com/ [5th November 2018]

137. Brown SR, Tiernan JP, Watson AJM, Biggs K, Shephard N, Wailoo AJ, Bradburn M, Alshreef A, Hind D, team HS. Haemorrhoidal artery ligation versus rubber band ligation for the management of symptomatic second-degree and third-degree haemorrhoids (HubBLe): a multicentre, open-label, randomised controlled trial. *Lancet* 2016;**388**(10042): 356-364.

138. Seymour M, Morton D, Investigators obotIFT. FOxTROT: an international randomised controlled trial in 1052 patients (pts) evaluating neoadjuvant chemotherapy (NAC) for colon cancer. *Journal of Clinical Oncology* 2019;**37**: 3504.

139. Keeler BD, Dickson EA, Simpson JA, Ng O, Padmanabhan H, Brookes MJ, Acheson AG, Group IT. The impact of pre-operative intravenous iron on quality of life after colorectal cancer surgery: outcomes from the intravenous iron in colorectal cancer-associated anaemia (IVICA) trial. *Anaesthesia* 2019;**74**(6): 714-725.

140. Daniels L, Ünlü Ç, de Korte N, van Dieren S, Stockmann HB, Vrouenraets BC, Consten EC, van der Hoeven JA, Eijsbouts QA, Faneyte IF, Bemelman WA, Dijkgraaf MG, Boermeester MA, Group DDDDCS. Randomized clinical trial of observational versus antibiotic treatment for a first episode of CT-proven uncomplicated acute diverticulitis. *Br J Surg* 2017;**104**(1): 52-61.

141. Cousins S, Blencowe NS, Blazeby JM. What is an invasive procedure? A definition to inform study design, evidence synthesis and research tracking. *BMJ Open* 2019;**9**(7): e028576.

142. National Institute for Health Research. https://www.nihr.ac.uk/research-and-impact/research/nihr-studies/research-for-patient-benefit.htm. [28th November 2017].

143. National Institute for Health Research. https://www.nihr.ac.uk/funding-and-support/funding-for-research-studies/funding-programmes/health-technology-assessment/ [28th November 2017].

144. IDEAL Collaboration http://www.ideal-collaboration.net/ [28th November 2017].

145. Ritchie J, Lewis J, McNaughton Nicholls C, Ormston R. *Qualitative research practice : a guide for social science students and researchers* (Second edition edn). SAGE Publications Ltd: London, 2014; xxiv, 430 pages.

146. Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P. Qualitative research methods in health technology assessment: a review of the literature. *Health technology assessment* 1998;**2**(16): iii-ix, 1-274.

147. Al-Busaidi ZQ. Qualitative research and its uses in health care. *Sultan Qaboos Univ Med J* 2008;**8**(1): 11-19.

148. Kuper A, Reeves S, Levinson W. An introduction to reading and appraising qualitative research. *Bmj* 2008;**337**: a288.

149. O'Cathain A, Murphy E, Nicholl J. Why, and how, mixed methods research is undertaken in health services research in England: a mixed methods study. *BMC health services research* 2007;**7**: 85.

150. Mays N, Pope C. Qualitative research in health care. Assessing quality in qualitative research. *Bmj* 2000;**320**(7226): 50-52.

151. O'Cathain A. Mixed Methods Research in the Health Sciences A Quiet Revolution. *J Mix Method Res* 2009;**3**(1): 3-6.

152. Mills N, Donovan JL, Smith M, Jacoby A, Neal DE, Hamdy FC. Perceptions of equipoise are crucial to trial participation: a qualitative study of men in the ProtecT study. *Control Clin Trials* 2003;**24**(3): 272-282.

153. Phelps EE, Tutton E, Griffin X, Baird J, co-applicants Ts. Facilitating trial recruitment: A qualitative study of patient and staff experiences of an orthopaedic trauma trial. *Trials* 2019;**20**(1): 492.

154. Horwood J, Johnson E, Gooberman-Hill R. Understanding involvement in surgical orthopaedic randomized controlled trials: A qualitative study of patient and health professional views and experiences. *Int J Orthop Trauma Nurs* 2016;**20**: 3-12.

155. Ziebland S, Featherstone K, Snowdon C, Barker K, Frost H, Fairbank J. Does it matter if clinicians recruiting for a trial don't understand what the trial is really about? Qualitative study of surgeons' experiences of participation in a pragmatic multi-centre RCT. *Trials* 2007;**8**: 4.

156. Cook JA, Campbell MK, Gillies K, Skea Z. Surgeons' and methodologists' perceptions of utilising an expertise-based randomised controlled trial design: a qualitative study. *Trials* 2018;**19**(1): 478.

157. Donovan JL, Rooshenas L, Jepson M, Elliott D, Wade J, Avery K, Mills N, Wilson C, Paramasivan S, Blazeby JM. Optimising recruitment and informed

consent in randomised controlled trials: the development and implementation of the Quintet Recruitment Intervention (QRI). *Trials* 2016;**17**(1): 283.

158. Mills N, Gaunt D, Blazeby JM, Elliott D, Husbands S, Holding P, Rooshenas L, Jepson M, Young B, Bower P, Tudur Smith C, Gamble C, Donovan JL. Training health professionals to recruit into challenging randomized controlled trials improved confidence: the development of the QuinteT randomized controlled trial recruitment training intervention. *J Clin Epidemiol* 2018;**95**: 34-44.

159. Rooshenas L, Scott LJ, Blazeby JM, Rogers CA, Tilling KM, Husbands S, Conefrey C, Mills N, Stein RC, Metcalfe C, Carr AJ, Beard DJ, Davis T, Paramasivan S, Jepson M, Avery K, Elliott D, Wilson C, Donovan JL, group B-B-Ss, group Cs, group H-s, group Ops, group Rfs. The QuinteT Recruitment Intervention supported five randomized trials to recruit to target: a mixed-methods evaluation. *J Clin Epidemiol* 2019;**106**: 108-120.

160. Clement C, Edwards SL, Rapport F, Russell IT, Hutchings HA. Exploring qualitative methods reported in registered trials and their yields (EQUITY): systematic review. *Trials* 2018;**19**(1): 589.

161. Qualitative Research Guidelines Project. http://www.qualres.org/index.html [15th July 2019].

162. Silverman D. Interpreting Qualtitative Data: Methods for analysing Talk, Text and Interaction. SAGE publications, 1993.

163. Reichardt CSR, S.F. (Eds.). *The qualitative-quantitative debate: New perspectives*. Jossey-Bass: San Francisco, 1994.

164. King NH, C. *Interviews in Qualitative Research*. SAGE Publications Ltd, 2010.

165. Paley J, Lilford R. Qualitative methods: an alternative view. *Bmj* 2011;**342**: d424.

166. Reeves S, Kuper A, Hodges BD. Qualitative research methodologies: ethnography. *Bmj* 2008;**337**: a1020.

167. Willig C. Beyond appeearances: a critical realist approach to social constructionism. In: *Social constructionist psychology: a critical analysis of theory and practice*, Nightingal DJC, J. (ed). Open University Press, 1999; 37-51.

168. Hanson WE, Creswell JW, Clark VLP, Petska KS, Creswell JD. Mixed methods research designs in counseling psychology. *J Couns Psychol* 2005;**52**(2): 224-235.

169. Mays N, Pope C. Rigour and qualitative research. *Bmj* 1995;**311**(6997): 109-112.

170. Golafshani N. Understanding reliability and validity in qualitative research. *The Qualitative Report* 2003;**8**: 597-607.

171. https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1d-qualitative-methods/principles-qualitative-methods-exercise1answers. [4th August 2019].

172. Glaser BG, Strauss AL. The discovery of grounded theory : strategies for qualitative research. Aldine Pub: Chicago, 1967; x, 271 pages.

173. Kuper A, Lingard L, Levinson W. Critically appraising qualitative research. *Bmj* 2008;**337**: a1035.

174. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, Burroughs H, Jinks C. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant* 2018;**52**(4): 1893-1907.

175. Strauss AL, Corbin JM. *Basics of qualitative research : techniques and procedures for developing grounded theory* (2nd edn). Sage Publications: Thousand Oaks, 1998; xiii, 312.

176. Mason M. Sample size and saturation in PhD studies using qualitative interviews. *Forum Qualitative Social Research* 2010;**11**(3).

177. Braun V, Clark V. Successful qualitative research: a practical guide for beginners. SAGE Publication Ltd, 2013.

178. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. *Bmj* 2000;**320**(7227): 114-116.

179. Lingard L, Albert M, Levinson W. Grounded theory, mixed methods, and action research. *Bmj* 2008;**337**: a567.

180. Braun VC, V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006;**3**: 77-101.

181. http://www.qsrinternational.com/nvivo. [12th August 2019].

182. http://atlasti.com/ [12th August 2019].

183. Barbour RS. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? *Bmj* 2001;**322**(7294): 1115-1117.

184. Bloor M. Techniques of validation in qualitative research: a critcial commentary. In: *Context and method in qualitative research*, Miller GD, R. (ed). SAGE: London, 1997; 37-50.

185. Braun VC, V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006;**3**(2): 77-101.

186. Greene J CV, Graham W. Towards a conceptual framework for mixed-methods evaluation designs. . *Educational Evaluation and Policy* 1989;**11**(3): 255-274.

187. Tashakkori AT, C. *Handbook of mixed methods in social and behavioural research*. SAGE: Thousand Oaks, California, 2003.

188. Cresswell JW. Research design: Quantitative, qualitative, and mixed methods approaches (2nd edn). SAGE: Thousand Oaks, California, 2003.

189. A. B. Integrating quantitative and qualitative research: how is it done? . *Qualitative Research* 2006;**6**(1): 97-113.

190. O'Cathain A ME, Nicholl J. Integration and Publications as Indicators of "Yield" From Mixed Methods Studies. . *J Mix Method Res* 2007;**1**(2): 147-163.

191. Punch KF. Introduction to social research: Quantitative and qualitative approaches. SAGE: Thousand Oaks, California, 1998.

192. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. *Bmj* 2010;**341**: c4587.

193. Adamson J, Ben-Shlomo Y, Chaturvedi N, Donovan J. Exploring the impact of patient views on 'appropriate' use of services and help seeking: a mixed method study. *The British journal of general practice : the journal of the Royal College of General Practitioners* 2009;**59**(564): e226-233.

194. O'Cathain A, Murphy E, Nicholl J. Multidisciplinary, interdisciplinary, or dysfunctional? Team working in mixed-methods research. *Qualitative health research* 2008;**18**(11): 1574-1585.

195. Farmer T, Robinson K, Elliott SJ, Eyles J. Developing and implementing a triangulation protocol for qualitative health research. *Qualitative health research* 2006;**16**(3): 377-394.

196. https://www.rcseng.ac.uk/standards-and-research/research/surgical-trials-initiative/ [28th June 2017].

197. RfPB programme guidance on applying for feasibility studies. https://www.nihr.ac.uk/documents/nihr-research-for-patient-benefit-rfpb-programme-guidance-on-applying-for-feasibility-studies/20474?pr= 2019]. [2nd September 2019]

198. Bristol Oxford Surgical Trials Course (BOSTiC). https://www.bristol.ac.uk/population-health-sciences/centres/surgical-research/education/courses/ 2019]. [2nd September 2019]

199. http://www.starsurg.org/granule-courses. [11th December 2019].

200. Joint Committee on Surgical Training. https://www.jcst.org/quality-assurance/certification-guidelines-and-checklists/ [11th December 2019].

201. Fairhurst K, Blazeby JM, Potter S, Gamble C, Rowlands C, Avery KNL. Value of surgical pilot and feasibility study protocols. *Br J Surg* 2019;**106**(8): 968-978.

202. Macefield RC, Boulind CE, Blazeby JM. Selecting and measuring optimal outcomes for randomised controlled trials in surgery. *Langenbeck's archives of surgery / Deutsche Gesellschaft fur Chirurgie* 2014;**399**(3): 263-272.

203. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. *Statistical methods in medical research* 2016;**25**(3): 1057-1073.

204. Ruzicka M, McCormick B, Magner P, Ramsay T, Edwards C, Bugeja A, Hiremath S. Thiazide diuretic-caused hyponatremia in the elderly hypertensive: will a bottle of Nepro a day keep hyponatremia and the doctor away? Study protocol for a proof-of-concept feasibility trial. *Pilot Feasibility Stud* 2018;**4**: 71.

205. Sim J. Should treatment effects be estimated in pilot and feasibility studies? *Pilot Feasibility Stud* 2019;**5**: 107.

206. E W, EA S. The nonuse, misuse, and proper use of pilot studies in experimental evaluation research. *Am J Eval* 2016;**38**(2): 246-261.

207. EC L, AL W, RM J, SA J. The statistical interpretation of pilot trials: should significance thresholds be reconsidered? *BMC Med Res Methodol* 2014;**14**.

208. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. *Pharm Stat* 2005;**4**(4): 287-291.

209. Morgan B, Hejdenberg J, Hinrichs-Krapels S, Armstrong D. Do feasibility studies contribute to, or avoid, waste in research? *PloS one* 2018;**13**(4): e0195951.

210. Peter Craig AM, Susan Browne, Sharon Anne Simpson, Danny Wight, Mike Robling, Graham Moore, Britt Hallingberg, Jeremy Segrott, Ruth Turley, Simon Murphy, Laurence Moore. Development of guidance for feasibility studies to decide whether and how to proceed to full-scale evaluation of complex public health interventions: a systematic review. *The Lancet* 2018;**392**: S7.

211. Chan CL, Leyrat C, Eldridge SM. Quality of reporting of pilot and feasibility cluster randomised trials: a systematic review. *BMJ open* 2017;**7**(11): e016970.

212. National Institute for Health Research: Applied Health Research in Surgery (2012). https://www.nihr.ac.uk/funding-and-support/documents/themed-calls/Surgery.pdf [14th January 2019].

213. Moore L, Hallingberg B, Wight D, Turley R, Segrott J, Craig P, Robling M, Murphy S, Simpson SA, Moore G. Exploratory studies to inform full-scale evaluations of complex public health interventions: the need for guidance. *Journal of epidemiology and community health* 2018;**72**(10): 865-866.

214. Hallingberg B, Turley R, Segrott J, Wight D, Craig P, Moore L, Murphy S, Robling M, Simpson SA, Moore G. Exploratory studies to decide whether and how to proceed with full-scale evaluations of public health interventions: a systematic review of guidance. *Pilot Feasibility Stud* 2018;**4**: 104.

# *Glossary of Abbreviations*

| | |
|---|---|
| ABS | Association of Breast Surgery |
| ARUK | Arthritis Research UK |
| ASGBI | Association of Surgeons of Great Britain and Ireland |
| ASiT | Association of Surgeons in Training |
| BHF | British Heart Foundation |
| BOSTiC | Bristol Oxford Surgical Trials Course: Making trials stick |
| BSc | Batchelor of Science |
| CRUK | Cancer Research UK |
| COMET | Core Outcome Measures in Effectiveness Trials |
| ConDuCT-II | Collaboration and innovation in Difficult and Complex randomised controlled Trials In Invasive procedures |
| CONSORT | CONsolidated Statement Of Reporting of Trials |
| COS | Core Outcome Set |
| CSO | Chief Scientist Office |
| CT | Computerised Tomography |
| CTU | Clinical Trials Unit |
| DMC | Data Monitoring Committee |
| ECV | External Cephalic Version |
| EME | Efficacy and Mechanism |
| ESTEeM trial | Endocrine +/- Surgical Therapy for Elderly women with Mammary cancer |

| | |
|---|---|
| FREC | University of Bristol Faculty of Health Sciences Research Ethics Committee |
| GUEST | GUidance for Exploratory STudies of complex public health interventions |
| HPB | Hepatopancreaticobiliary |
| HS&DR | Health Services and Delivery Research |
| HTA | Health Technology Assessment |
| HTA CB | Health Technology Assessment Commissioning Board |
| HTA CEAT | Health Technology Assessment Clinical Evaluation and Trials Board |
| HTA GB | Health Technology Assessment General Board |
| ICTMC | International Clinical Trials Methodology Conference |
| IDEAL | Idea, Development, Exploration, Assessment Long-term Study |
| i4i | Invention for Innovation |
| ISRCTN | International Standard Randomised Controlled Trials Number |
| JCST | Joint Committee on Surgical Training |
| KF | Katherine Fairhurst |
| MBChB | Batchelor of Medicine, Batchelor of Surgery |
| MRC | Medical Research Council |
| NETSCC | NIHR Evaluation, Trials and Studies Co-ordinating Centre |
| NHS | National Health Service |
| NIHR | National Institute for Health Research |
| PEPSTAR | Exploring Perceptions and Experiences of Pilot Work for Surgical Trials: A Qualitative Research Study |
| PFS | Pilot/Feasibility Studies |

| | |
|---|---|
| PhD | Doctor of Philosophy |
| PHR | Public Health Research |
| PGfAR | Programme Grants for Applied Research |
| PPI | Patient and Public Involvment |
| PRISMA | Transparent Reporting of Systematic Reviews and Meta-Analyses |
| PROTECT trial | Prophylaxis for ThromboEmbolism in Critical Care Trial |
| RCSEng | Royal College of Surgeons of England |
| RCT | Randomised Controlled Trial |
| RDS | Research Design Service |
| REF | Research Excellence Framework |
| RfPB | Research for Patient Benefit |
| ROCSS trial | Reinforcement of Closure of Stoma Site |
| ROSSINI trial | Reduction Of Surgical Site Infection using several Novel Interventions |
| QUEST trial | Quality of life after mastectomy and breast reconstruction |
| SCT | Society of Clinical Trials |
| SPIRIT | Standard Protocol Items: Recommendations for Interventional Trials |
| TIDieR | Template for Intervention Description and Replication |
| TIME trial | Traditional Invasive versus Minimally invasive Esophagectomy |
| TSC | Trial Steering Committee |
| UK | United Kingdom |
| USA | United States of America |

# *Appendices*

**Appendix I** Data extraction form for a targeted review and systematic analysis of NIHR funded pilot and feasibility studies

NB Detail is related to all proposed data points to be collected, including to enable follow up of studies as they complete and progress/not to a main trial

| Allocated study code: | NIHR Project number: |
|---|---|

**Pilot and feasibility studies of surgical interventions:**
**A systematic review of NIHR funded studies.**

**DATA EXTRACTION FORM**
**(office use only)**

Please use all of the different data sources given to you to fill out this form. The form will in places ask for information from a specific source or guide you to where the information might be. Not all data sources are available for each study, but each study will have the minimum of a pilot/feasibility study protocol available. If data from a specific data source is needed (e.g. the pilot/feasibility results paper or study report) this will be stated.

**Background data**

1 Date of data extraction (DD/MM/YYYY)

2 Person extracting data (1=KF, 2=KA, 3=SP, 4=JB, 5=CR, 6=other) If other please state

3 Title of **pilot/feasibility** project

4 **Pilot/feasibility** project team personnel

| **Role in project** 1= CI, 2= Co-CI, 3= PI, 4 = Co-Investigator, 5= PPI representative, 6= Collaborator, 7= TSC, 8= TMC, 9=Sponsor representative, 10= Trial co-ordinator, 10= other (state) | **Specialty** 1= surgeon, 2=doctor, 3= Nurse/OPAM, 4=trialist/methodologist, 5= researcher, 6= qualitative researcher, 7= ethicist, 8= statistician, 9= Health economist, 10= R&D manager, 11 = other (state) |
|---|---|
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |

5 Please give any other information about the project team here

6 Total number of people named in the protocol

1

**Section A: General characteristics of the study**

**A1a** What type of study do the authors label this as? Please tick all that apply.
(See guidance notes)

| | | | |
|---|---|---|---|
| ☐ | 1= Randomised internal pilot | ☐ | 4= Other feasibility work |
| ☐ | 2= Randomised external pilot | ☐ | 5= Uncertain/not stated |
| ☐ | 3= Non-randomised pilot | ☐ | 6= Other |

**A1b** If Other please state

**A1c** Verbatim description

**A2** Status of pilot/feasibility study? ☐ 1= Completed, 2= In progress, 3= Not completed/stopped, 4= Uncertain

**A3** Single centre ☐ Multicentre ☐ (X in appropriate box) If multicentre, state number of centres

**A4** Minimum number of proposed participants

**A5** Source of data
(X in all available, see guidance notes)

| | |
|---|---|
| 1 | 1= Pilot/feasibility study protocol |
| 2 | 2= Pilot/feasibility grant application |
| 3 | 3= Published pilot/feasibility protocol paper |
| 4 | 4= Published pilot/feasibility study results paper |
| 5 | 5= NIHR pilot/feasibility study report |
| 6 | 6= Main trial protocol |
| 7 | 7= Published main trial protocol paper |
| 8 | 8= Published main trial results paper |

**A5** Date of publication
(Write date in box DD/MM/YYYY)

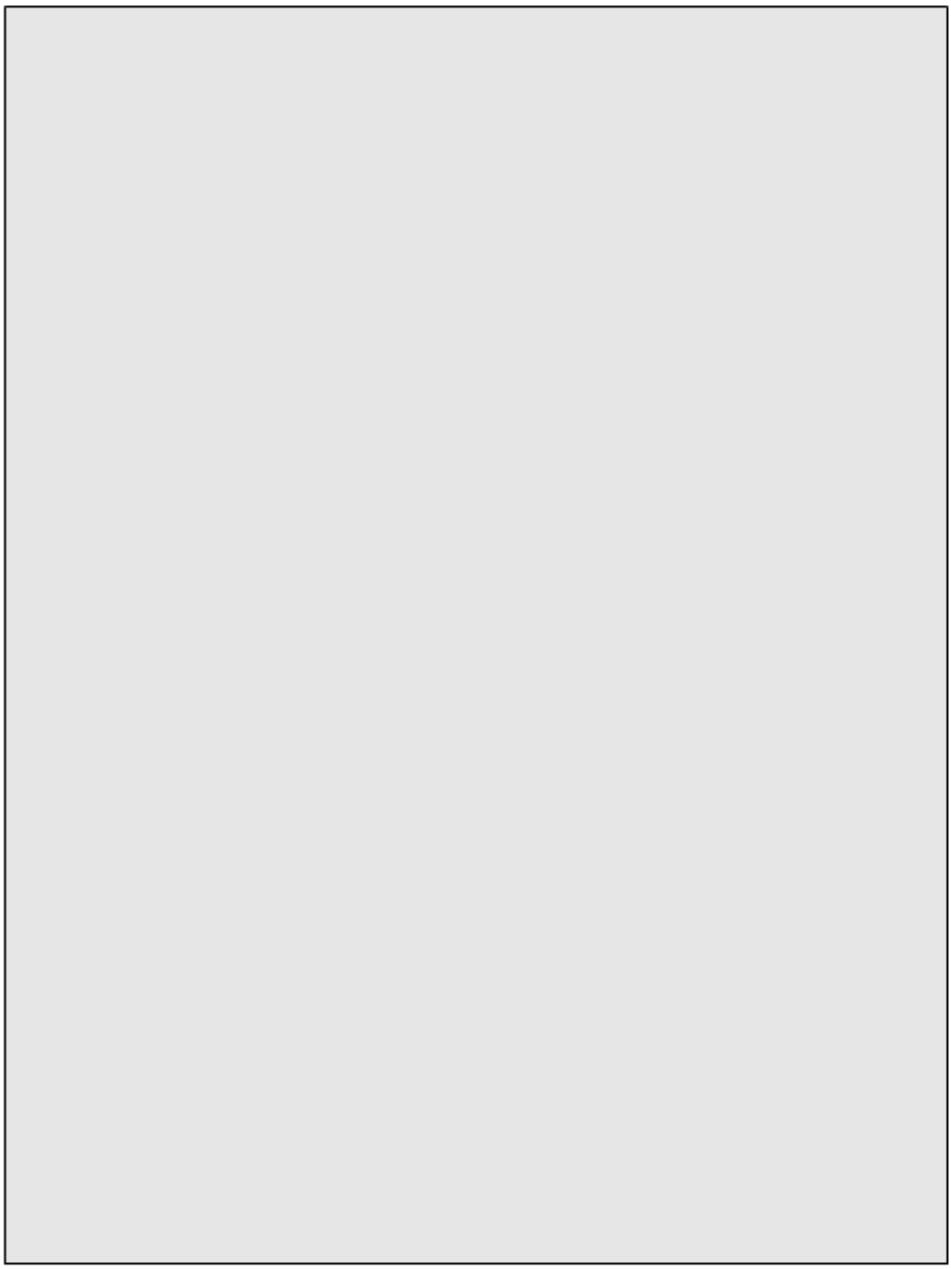
| | |
|---|---|
| 1 | |
| 2 | |
| 3 | |
| 4 | |
| 5 | |
| 6 | |
| 7 | |
| 8 | |

**A6** How long did the pilot/feasibility study run for? (MM/YY)
**OR** uncertain/not stated ☐
(please tick)

**A7** Surgical specialty of study ☐ 1=GI, 2=Urology, 3=cardiothoracic, 4=orthopaedic, 5=O&G, 6=MaxFax/ENT, 7=Breast, 8=Plastics, 9=Paeds)

**Section B: Rationale for the pilot/feasibility study**

**B1** Copy & paste *ad verbatim* any description of the rationale for the pilot/feasibility here (See guidance notes)



**B2** Categorisation of the rationale for the pilot/feasibility study (See guidance notes

| Area | | | Intervention | B2a What were the reasons stated for performing the study in the protocol? (X each that apply – see guidance notes) | B2b What areas did the study report on? (X each that apply – see guidance notes) |
|---|---|---|---|---|---|
| Main Trial design | Main trial possible / necessary | 1 | To examine and test whether a main trial is possible | | |
| | | 2 | To assess whether main trial is needed and/or produce a protocol | | |
| | | 3 | To test whether the protocol can be adhered to and modify it as necessary | | |
| | Sample size | 4 | To estimate the variability in outcomes to help determine a sample size for the main trial | | |
| | | 5 | To determine a sample size for the main trial | | |
| | Costs / funding | 6 | To assess/gather information on costs of performing the trial (direct and indirect) | | |
| | | 7 | To perform/prepare for a cost effectiveness analysis of the intervention(s) | | |
| | | 8 | To provide information/evidence to funders | | |
| | Hypothesis testing | 9 | To test the safety of an intervention | | |
| | | 10 | To test the effectiveness of an intervention | | |
| Logistics | | 11 | To test the logistics of multicentre studies | | |
| | | 12 | To develop a research network as a resource for a future main trial | | |
| | | 13 | To develop/test patient information content/forms/methods of delivery | | |
| | | 14 | To develop/test data collection forms/methods | | |
| | | 15 | To develop/test questionnaires/surveys | | |
| | | 16 | To test response rates to questionnaires/surveys | | |
| | | 17 | To prepare/plan/assess monitoring procedures | | |
| | | 18 | To determine what resources are needed for a main trial (funding/staff) | | |
| | | 19 | To assess the logistics of delivering an intervention as part of a trial in the NHS | | |
| | | 20 | To test (novel) methods of blinding | | |
| | | 21 | To assess proposed data analysis techniques | | |
| | | 22 | To lean about the day-to-day running of a trial | | |
| Recruitment | | 23 | To test/modify inclusion/exclusion/eligibility criteria | | |
| | | 24 | To estimate the expected prevalence or rate of incidence cases in the population | | |
| | | 25 | To estimate the number to be screened and proportions of eligible patients | | |
| | | 26 | To assess numbers/rates of recruitment and consent | | |
| | | 27 | To test the randomisation procedure | | |
| | | 28 | To test the acceptability of randomisation/trial design | | |
| | | 29 | To determine the acceptability of the intervention to clinicians and patients | | |
| | | 30 | To assess rates of retention in the study | | |
| Intervention | | 31 | To assess and monitor the development of an intervention and/or it's stability | | |
| | | 32 | To develop and test the implementation and delivery of the intervention | | |
| | | 33 | To train staff in delivery and assessment procedures | | |
| | | 34 | To monitor the surgical learning curve | | |
| | | 35 | To test rates of crossover | | |
| | | 36 | To examine reasons for non-adherence/cross-over for the main trial | | |
| | | 37 | To develop pathways and protocols for co-interventions | | |
| Outcome | | 38 | To select the most appropriate primary outcome measure | | |
| | | 39 | To develop and test a new outcome measure | | |
| | | 40 | To determine appropriate/important/suitability of outcome measures for patients/clinicians | | |
| Other | | | Other (state) | | |
| | | | Other (state) | | |

**B3** Data related to the rationale for pilot/feasibility study: For each item identified in B2, please fill in B3a-c
(See guidance notes)

| B3a. Reason/Area of uncertainty identified as: | | B3b. Defined *a priori* in protocol? | B3c. Data collected/Measured by? | | |
|---|---|---|---|---|---|
| Number | Actual wording of reason identified | | Data collected? | Methodology | Notes |
| | | Yes ☐<br>No ☐ | Yes ☐<br>No ☐ | Quantitative ☐<br>Qualitative ☐<br>Mixed ☐<br>Uncertain ☐<br>Other (please state) ☐<br>Not measured or reported ☐<br>Info not available ☐ | |
| | | Yes ☐<br>No ☐ | Yes ☐<br>No ☐ | Quantitative ☐<br>Qualitative ☐<br>Mixed ☐<br>Uncertain ☐<br>Other (please state) ☐<br>Not measured or Reported ☐<br>Info not available ☐ | |

**B4** What was the **main reason** for performing the study defined as (*ad verbatim* – generally overall aim)?

| |
|---|
| |

5

**B5** **Overall,** did the reasons for conducting the **pilot/feasibility study** remain the same as per the pilot/feasibility study protocol when the study was done? *(Will need to refer to pilot/feasibility protocol AND published pilot study paper/NIHR report if available)*

☐ 1= Adhered to
2= Not adhered to
3= Adhered to and additional reasons added
4= Uncertain (information not available)

**B6** **Please add any additional information here**

**Section C: Pilot/feasibility study design**

**C1a** Copy & paste *ad verbatim* description of study design given in pilot/feasibility study protocol here

**C1b** Type of pilot/feasibility study
(See guidance notes) ☐
1= Randomised internal pilot
2= Randomised external pilot
3= Non-randomised pilot
4= Other feasibility work

**C1c** If other specify
(See guidance notes)

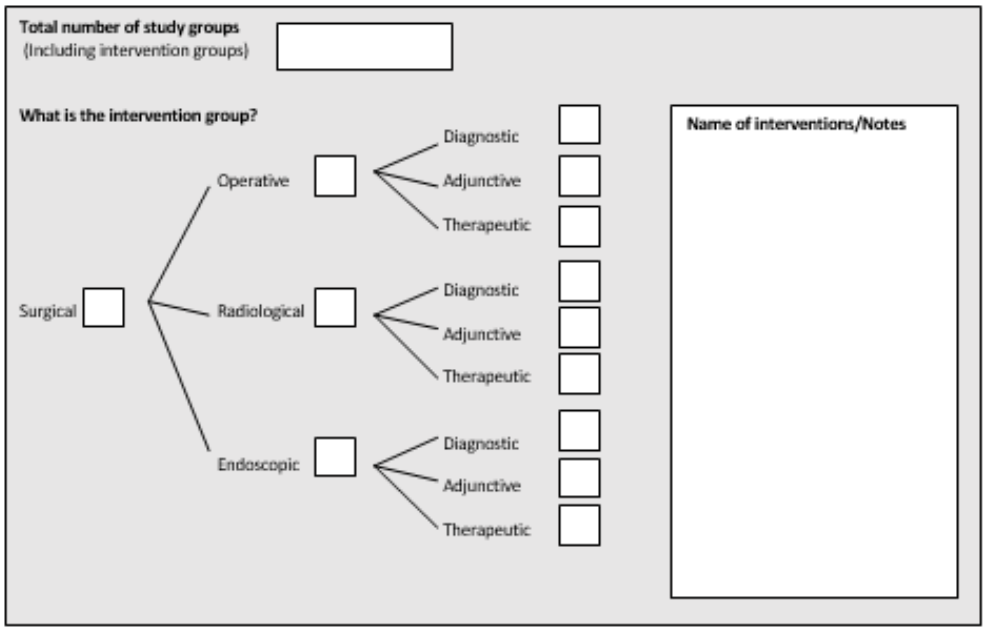
**C1d** If reported, is the pilot/feasibility actually a small RCT? ☐
(See guidance notes)
1= Yes
2= No
3= Uncertain/information not available

**C2** Patient population of the **pilot/feasibility study** (X all that apply)

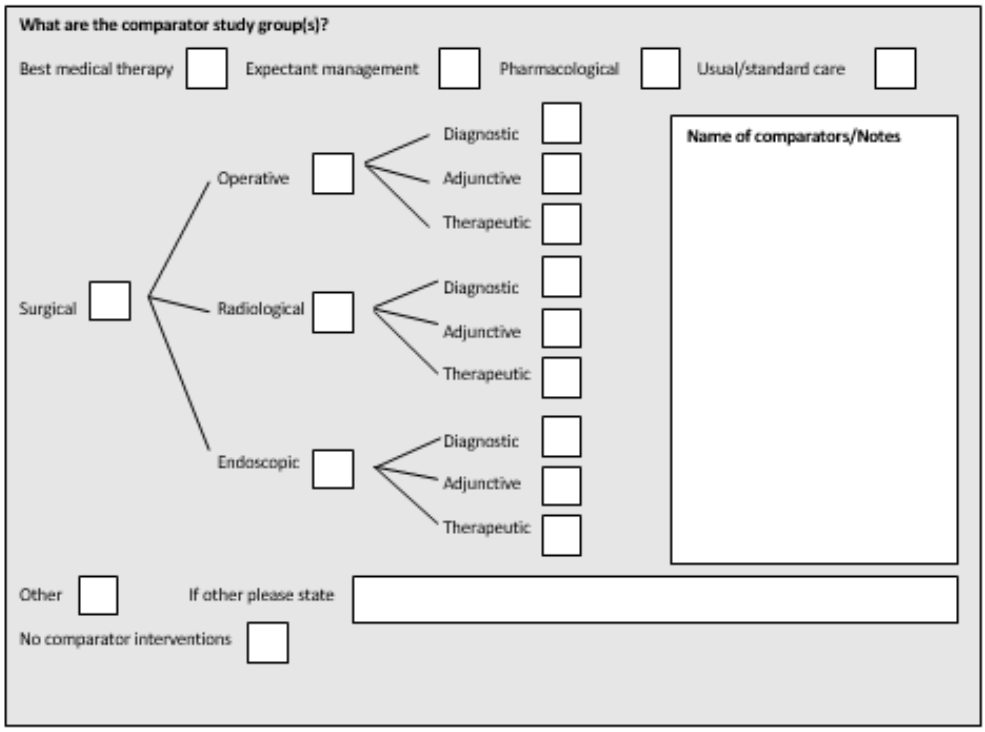
| | | | | | |
|---|---|---|---|---|---|
| Adults (≥ 18) | ☐ | Male | ☐ | UK | ☐ |
| Children (<18 yrs) | ☐ | Female | ☐ | Europe | ☐ |
| | | Both | ☐ | Worldwide | ☐ |

Countries (List all)

**C3** Intervention of the **pilot/feasibility study** (If applicable, X all that apply, See guidance notes)




**C4** Comparators of the **pilot/feasibility study**
(If applicable, enter **number** of each type of comparator group in the appropriate boxes, see guidance notes)



**C5** Outcomes of the **pilot/feasibility study** - If the protocol states primary/secondary/feasibility outcomes please copy and paste text *ad verbatim* here (See guidance notes)

| **Primary:** |
|---|

| **Secondary:** |
|---|

| **Feasibility:** |
|---|

**C6** Copy & paste *ad verbatim* description of data analysis proposed/used here

**C7a** Was a sample size calculation performed for the **pilot/feasibility study?** ☐

1= Yes
2= No
3= Uncertain/info not available

**C7b** If performed, what reasons for doing so are given?

**C7c** If performed, what assumptions were made in order to make the calculation?

**C8a** Was an economic analysis done for the **pilot/feasibility study?** ☐

1= Yes
2= No
3= Uncertain/information not available

**C8b** If yes, what were the stated end points for the pilot/feasibility economic analysis?

**C8c** If yes, what end points did the pilot/feasibility study actually measure?

**Section D: Outcome of pilot/feasibility study & progress of main trial**
*(Use all available data sources)*

**D1a** Does the report state that a definitive **main trial** is planned? ☐ 1= Yes 2= No 3= Uncertain/Information not available

**D1b** If yes, please categorise ☐ 1= Planned, 2= Not planned, 3= Uncertain/Information not available

**D2a** Does the report state that a definitive **main trial** has been funded? ☐ 1= Yes 2= No 3= Uncertain/Information not available

**D2b** If yes, please categorise ☐ 1= Funded, 2= Not funded, 3= Uncertain

**D3** If a **main trial** is planned/funded what is the status of the main trial?

☐ 1= Completed, 2= In progress, 3= Not completed/stopped, 4= Uncertain, 5= Not applicable

**D4** If a **main trial** is planed/funded/in progress, who are the project team personnel?

| **Role in project**<br>1= CI, 2= Co-CI, 3= PI, 4 = Co-Investigator, 5= PPI representative, 6= Collaborator, 7= TSC, 8= TMC, 9=Sponsor representative, 10= Trial co-ordinator, 10= other (state) | **Specialty**<br>1= surgeon, 2=doctor, 3= Nurse/OPAM, 4=trialist/methodologist, 5= researcher, 6= qualitative researcher, 7= ethicist, 8= statistician, 9= Health economist, 10= R&D manager, 11 = other (state) |
|---|---|
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |

**D4b** Please give any other information about the project team here

| |
|---|
| |

**D5a** Are there any stated decision/progression criteria regarding progress to a **main trial**?
(See guidance notes) ☐ 1= Yes 2= No 3=Uncertain/information not available

**D5b** If yes, copy and paste text *ad verbatim* here

**D5c** Were these decision/progression criteria met? ☐ 1= Yes, 2= No, 3= Uncertain/info not available, 4= Not applicable

**D5d** If no, why were these criteria not met? Copy and paste text *ad verbatim* here

**D6a** Were remaining uncertainties about the **main trial** reported? ☐ 1= Yes, 2= No, 3= Uncertain/info not available

**D6b** If yes, please copy and paste *ad verbatim* any text regarding remaining uncertainties about the feasibility/viability of the main trial here

**D7** If main trial **completed/in progress**, how was it improved?

**D8** If the main trial is **not completed/stopped**, why did this happen?

## Appendix II: Pilot and feasibility studies of surgical interventions funded by NIHR HTA and RfPB programmes between 2005 and 2015 details and outputs.

| | NIHR project ID & CI | Date funded to projected study end date | Lead centre | Title | Related data sources available | | | | Type of pilot/feasibility work | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Protocol | Protocol paper | Results paper | NIHR report | Randomised external pilot | Details of other feasibility work |
| **RfPB Funded** | | | | | | | | | | |
| 62-407 | PB-PG-0807-14131 Jane Blazeby | 01/01/2010 30/06/2013 | University Hospitals Bristol NHS Foundation Trust | Oesophageal squamous cell cancer: chemoradiotherapy versus chemotherapy and surgery - a feasibility study | ✓ | | ✓ | N/A | ✓ | Qualitative |
| 63-413 | PB-PG-0808-15115 Adrian Marchbank | 16/08/2010 15/02/2012 | Plymouth Hospitals NHS Trust | A pilot randomised controlled single-blind trial of a collagen implant for the prevention of sternal wound infection in cardiac surgery. | ✓ | | | N/A | ✓ | |
| 64-442 | PB-PG-0808-17257 Paul Roderick | 01/09/2010 31/08/2012 | University Hospital Southampton NHS Foundation Trust | Optimum care for non-morbidly obese patients with type 2 diabetes: the perspectives of patients and clinicians on the role of bariatric surgery. | ✓ | | ✓ | N/A | | Surveys (quantitative & qualitative) |
| 66-499 | PB-PG-0909-20150 Charles Knowles | 21/01/2011 28/08/2013 | Barts Health NHS Trust | A randomised mixed methods pilot (phase II exploratory trial) of sacral and percutaneous tibial nerve stimulation for faecal incontinence | ✓ | | ✓ | N/A | ✓ | Qualitative |
| 65-492 | PB-PG-0909-20079 Khaled Ismail | 01/02/2011 30/06/2014 | University Hospitals of North Midlands NHS Trust | Perineal re-suturing versus expectant management following vaginal delivery complicated by a dehisced wound (PREVIEW): A pilot randomised controlled study. | ✓ | ✓ | ✓ | N/A | ✓ | Qualitative interviews |
| 72-633 | PB-PG-1208-17025 Anjan Dhar | 15/02/2011 14/02/2014 | County Durham and Darlington NHS Foundation Trust | Biodegradable stent in benign oesophageal stricture compared to standard balloon dilatation treatment | ✓ | | ✓ | N/A | ✓ | |
| 48-472 | PB-PG-0110-21019 Alastair Sutcliffe | 01/07/2011 30/12/2013 | North Bristol NHS Trust | A feasibility study for a randomised controlled trial to measure the impact of frenotomy in breastfed infants with tongue tie. | ✓ | | ✓ | N/A | ✓ | |
| 73-639 | PB-PG-1208-18031 Sam Eldabe | 01/09/2011 31/01/2015 | South Tees Hospitals NHS Foundation Trust | A multicentre randomised controlled trial of spinal cord stimulation plus usual care vs. usual care alone in the management of refractory angina: A feasibility & pilot study. | ✓ | ✓ | ✓ | N/A | ✓ | Survey |
| 67-503 | PB-PG-0909-20214 Malcolm Crundwell | 15/11/2011 14/05/2013 | Royal Devon and Exeter NHS Foundation Trust | Pilot investigation on the effect of the Memokath® 028 prostatic stent on Quality of Life in patients with urethral obstruction – a comparison with long term catheter | ✓ | | | N/A | ✓ | |
| 49-53 | PB-PG-0110-21121 Damian Griffin | 20/01/2012 19/01/2015 | University Hospitals Coventry and Warwickshire NHS Trust | SHOULDER ARTHROPLASTY TRIAL A Randomised Controlled Trial of Total Resurfacing Versus Hemi Resurfacing in the Treatment of Primary arthritis of the Shoulder | ✓ | | | N/A | ✓ | |
| 57-352 | PB-PG-0711-25066 John de Caestecker | 04/02/2013 03/11/2015 | University Hospitals of Leicester NHS Trust | BRIDE (Barrett's Randomised Intervention for Dysplasia by Endoscopy) - a feasibility study | ✓ | | ✓ | N/A | ✓ | Qualitative interviews |
| 50-95 | PB-PG-0212-27050 Martin Tickle | 01/07/2013 01/10/2015 | Central Manchester University Hospitals NHS Foundation Trust | A pilot randomised controlled trial to compare the costs and effects of Conventional Implants with Mini Implants used to retain full lower DENTures (CIMI-DENT trial) | ✓ | | ✓ | N/A | ✓ | |
| 58-354 | PB-PG-0711-25080 Rajendra Prasad | 13/01/2014 11/08/2016 | Leeds Teaching Hospitals NHS Trust | Randomised Controlled Trial of Pringle Manoeuvre versus Portal Vein Clamping in Patients undergoing Liver Resection for Colorectal Liver Metastasis – A Pilot Study | ✓ | | | N/A | ✓ | |
| 71-595 | PB-PG-1112-29107 Thomas Pinkney | 30/06/2014 29/12/2016 | University Hospitals Birmingham NHS Foundation Trust | The feasibility of undertaking Appendicectomy to impact upon the Clinical Course of UlceRativE Colitis- The ACCURE Trial Feasibility study | ✓ | ✓ | | N/A | ✓ | Qualitative |
| 59-383 | PB-PG-0712-28094 Anju Goyal | 01/07/2014 31/08/2018 | Central Manchester University Hospitals NHS Foundation Trust | Onabotulinum toxin-A versus extended release tolterodine in the management of idiopathic overactive bladder in children: A pilot randomised controlled trial | ✓ | | | N/A | ✓ | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 60-387 | PB-PG-0712-28112<br>Emma Hall | 01/10/2014<br>30/09/2018 | The Royal Marsden NHS Foundation Trust | A phase II randomised feasibility study of Chemoresection and surgical management in Low rIsk non muscle invasive Bladder cancER (CALIBER) | ✓ | | | **N/A** | ✓ | |
| 56-307 | PB-PG-0613-31114<br>Kevin Franks | 01/11/2014<br>31/10/2017 | Leeds Teaching Hospitals NHS Trust | A study to determine the feasibility and acceptability of conducting a phase III randomised controlled trial comparing Stereotactic Ablative Radiotherapy (SABR) with surgery in patients with peripheral stage I non-small cell lung cancer (NSCLC) considered higher risk of complications from surgical resection. | ✓ | ✓ | | **N/A** | ✓ | |
| 51-128 | PB-PG-0213-30106<br>Angharad Care | 01/12/2014<br>30/04/2017 | Liverpool Women's NHS Foundation Trust | Three-arm randomised trial of Arabin pessary, cervical cerclage and progesterone to prevent spontaneous preterm birth in an asymptomatic high risk cohort: a feasibility study | ✓ | | ✓ | **N/A** | ✓ | Qualitative |
| 69-541 | PB-PG-1013-32045<br>Pradeep Bhandari | 01/02/2015<br>30/04/2018 | Portsmouth Hospitals NHS Trust | A feasibility study with a crossover design to assess the diagnostic accuracy of acetic acid targeted biopsies versus non targeted biopsies (current practice) for detection of dysplasia during Barrett's surveillance: The ABBA study | ✓ | ✓ | | **N/A** | ✓ | Qualitative interviews |
| 70-545 | PB-PG-1013-32058<br>Martin Birchall | 01/05/2015<br>30/04/2017 | University College London Hospitals NHS Foundation Trust | Does Laryngeal Reinnervation or Type I Thyroplasty give better voice results for patients with Unilateral Vocal Fold Paralysis (VOCALIST): a feasibility study | ✓ | ✓ | | **N/A** | ✓ | Qualitative (QRI) |
| 52-143 | PB-PG-0214-33065<br>Shelley Potter | 01/06/2015<br>31/05/2018 | University Hospitals Bristol NHS Foundation Trust | The iBRA (implant Breast Reconstruction evAluation) study – A prospective multicentre cohort study to inform the feasibility and conduct a pragmatic randomised clinical trial comparing new techniques of implant-based breast beconstruction. | ✓ | ✓ | ✓ | **N/A** | | National audit<br>Survey<br>Qualitative work |
| 55-298 | PB-PG-0613-31083<br>Tim Davis | 01/07/2015<br>30/07/2017 | Nottingham University Hospitals NHS Trust | Needle fasciotomy versus limited fasciectomy for the treatment of Dupuytren's contractures of the fingers: a study which investigates the feasibility, acceptability and design of a multicentre randomised trial. | ✓ | ✓ | | **N/A** | ✓ | |
| 68-537 | PB-PG-1013-32019<br>Peter Hull | 01/07/2015<br>30/06/2018 | Cambridge University Hospitals NHS Foundation Trust | KFORT – Knee Fix or Replacement Trial. A feasibility study comparing fixation vs replacement in elderly patients sustaining a distal femoral fracture | ✓ | | | **N/A** | ✓ | |
| 53-145 | PB-PG-0214-33068<br>Sumita Verma | 01/09/2015<br>06/09/2018 | Brighton and Sussex University Hospitals NHS Trust | Palliative long-term abdominal drains versus repeated drainage in individuals with untreatable ascites due to advanced cirrhosis: a feasibility randomised controlled trial | ✓ | ✓ | | **N/A** | ✓ | |
| 61-388 | PB-PG-0712-28114<br>Steven Pereira | 30/11/2015<br>30/04/2020 | University College London Hospitals NHS Foundation Trust | A phase II trial of endoscopic ultrasound-guided radiofrequency ablation of cystic tumours of the pancreas (RADIOCYST) | ✓ | | | **N/A** | | Non-randomised (cohort) study |
| **HTA Funded** | | | | | | | | | | |
| 1-22 | 03/48/01<br>Marion Campbell | 01/07/2005<br>31/10/2008 | University Court of the University of Aberdeen | Effectiveness and cost-effectiveness of arthroscopic lavage in the treatment of osteoarthritis of the knee: a mixed methods study of the feasibility of conducting a surgical placebo controlled trial (the KORAL study) | ✓ | | ✓ | ✓ | ✓ | |
| 4-275 | 09/01/20<br>Michael Clarke | 01/01/2011<br>30/06/2013 | University of Newcastle | An External Pilot Study to test the Feasibility of a Randomised Controlled Trial comparing Eye Muscle Surgery against Active Monitoring for Childhood Intermittent Distance Exotropia | ✓ | ✓ | | ✓ | ✓ | Qualitative interviews |
| 14-376 | 09/167/02<br>Kevin O'Brien | 01/12/2011<br>30/11/2013 | Central Manchester University Hospitals NHS Foundation Trust | The management of Otitis Media with Effusion in children with cleft palate (mOMEnt) | ✓ | ✓ | ✓ | ✓ | | Qualitative interviews |
| 13-375 | 09/166/01<br>Heather Fortnum | 01/01/2012<br>31/01/2013 | The University of Nottingham | Assessment of the feasibility and clinical value of further research to evaluate the management options for children with Down syndrome and otitis media | ✓ | | | ✓ | | Questionnaires<br>Qualitative interviews, focus groups, Delphi review |
| 15-387 | 10/41/02<br>Damien Griffin | 01/03/2012<br>30/11/2013 | University Hospitals Coventry & Warwickshire NHS Trust | Feasibility study for a randomised controlled trial comparing hip arthroscopy with best conservative care for patients with femoro-acetabular impingement (UK FASHIoN) | ✓ | | ✓ | ✓ | ✓ | Pre-pilot consensus work<br>Survey<br>Qualitative interviews |
| 23-472 | 11/29/01<br>Jane Daniels | 01/08/2012<br>31/07/2014 | University of Birmingham | The relationship between refluxing pelvic veins and chronic pelvic pain in women | ✓ | | ✓ | ✓ | | Systematic reviews<br>Surveys |
| 17-398 | 10/50/65<br>Chris Metcalfe | 01/01/2013<br>31/12/2015 | University of Bristol | Randomised Oesophagectomy: Minimally Invasive or Open, a feasibility study. The ROMIO trial. | ✓ | ✓ | | ✓ | ✓ | Qualitative interviews<br>Audio recordings<br>Process evaluation |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 34-569 | 12/35/32<br>Vinidh Paleri | 01/01/2014<br>30/06/2016 | The Newcastle upon Tyne Hospitals NHS Trust | A feasibility randomised controlled trial of pre-treatment gastrostomy tube versus oral feeding plus as-needed nasogastric tube feeding in patients undergoing chemoradiation for head and neck cancer (TUBE trial) | ✓ | ✓ | | ✓ | ✓ | Qualitative<br>Economic modelling |
| 27-515 | 11/107/01<br>Naeem Soomro | 01/04/2014<br>31/07/2016 | The Newcastle upon Tyne Hospitals NHS Trust | Willingness of clinicians and patients to randomise into a study comparing radio frequency ablation (RFA) with active surveillance, in the management of incidentally diagnosed small renal tumours: a feasibility study | ✓ | | | ✓ | ✓ | Qualitative interviews |
| 37-572 | 12/35/54<br>Freddie Hamdy | 01/01/2015<br>30/06/2016 | University of Oxford | A randomised controlled trial of Partial prostate Ablation versus Radical prosTatectomy (PART) in intermediate risk unilateral clinically localised prostate cancer – a feasibility study | ✓ | | | ✓ | ✓ | Qualitative |

**Appendix III** PEPSTAR FREC Favourable opinion letter

University of BRISTOL

**Faculty of Health Sciences**
**Research Ethics Committee (FREC)**

University of Bristol Faculty of Health Sciences,
First Floor South, Senate House,
Tyndall Avenue, Bristol
BS8 1TH
Tel: 0117 331 8197

Research Governance and Ethics Officer:
Liam McKervey
E-mail: Liam.McKervey@bristol.ac.uk
Tel: 0117 928 9089

Miss Katherine Fairhurst
University of Bristol

13th October, 2016

Dear Miss Fairhurst,

**Re: Application 41001**

**Title: Exploring Perceptions and Experiences of Pilot work for Surgical Trials: A Qualitative Research Study (PEPSTAR)**

Thank you for responding to the issues raised by the Faculty of Health Sciences Research Ethics Committee (FREC) as stated in our letter dated 27.09.16. Your response to the issues raised by the FREC was reviewed by the chair of the committee who agreed to grant a favourable ethical opinion for the above named study.

The committee recognises that you have been diligent in anticipating and responding to ethical issues in your preparation for the research. Please note that the FREC expects to be notified of any changes or deviations in the study.

Good luck with your study.

Yours faithfully,
Liam McKervey
pp
[redacted]
Dr Allison Fulford
*Chair, Faculty of Health Sciences Research Ethics Committee*

**Appendix IV** PEPSTAR Participant Introductory letter

**Centre for Surgical Research**
University of Bristol
School of Social and Community Medicine
Canynge Hall 1.13, 39 Whatley Road,
Clifton, Bristol. BS8 2PS
**Phone:** 0117 92 87386
**Email:** katherine.fairhurst@bristol.ac.uk

29th August 2017

## Introductory email/letter to potential participants

Dear Participant,

We write to invite you to take part in our research study and enclose details of this in the participant information sheet. The study forms part of Katherine Fairhurst's PhD thesis titled **'Optimising the design and evaluation of pilot work to inform more efficient RCTs in surgery'** and supervised by Dr Kerry Avery and Professor Jane Blazeby.

The study aims to explore the experience and opinions of individuals working in the field of designing, conducting or funding pilot and feasibility work for surgical trials and to ultimately make recommendations for the optimal use of such pilot work. Given your experience and expertise, we would hugely value your participation in this study and hope that you will agree to take part.

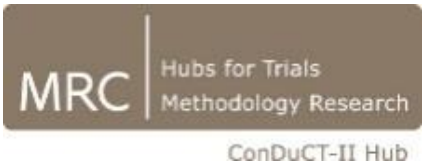
With kind regards,

**Miss Katherine Fairhurst**
MRC Clinical Research Training Fellow
MRC ConDucT - II Hub PhD Student
Specialist Training Registrar General Surgery

**Professor Jane Blazeby**
Professor of Surgery
Director of the MRC ConDucT II Hub
Director of the Bristol RCS Trials Centre

**Dr Kerry Avery**
Research Fellow

## **Appendix V** PEPSTAR Participant information sheet





**Appendix 3: Participant Information Sheet A**

**Exploring Perceptions and Experiences of Pilot work for Surgical Trials: A Qualitative Research Study (PEPSTAR)**

We would like to invite you to take part in our research study as titled above. Before you decide we would like you to understand why the research is being done and what it would involve for you. Please feel free to talk to others about the study if you wish and/or ask us if there is anything that is not clear.

**What is the purpose of the study?**
We wish to explore opinions, perceptions, experiences and attitudes of individuals involved the field of designing, conducting or funding pilot and feasibility work for surgical trials. The ultimate aim is to form recommendations for the optimal use of such pilot and feasibility work.

**Why have I been invited?**
You are being asked to participate because you have experience in the field of designing, conducting or funding pilot and feasibility work for surgical trials and we would value your expert opinion and contribution.

**Do I have to take part?**
Taking part in this study is entirely voluntary. This information sheet describes the study and what it involves. If you are interested in participating in this study please get in touch via the contact details at the end of this information sheet. We may contact you again if we don't hear from you to follow up on whether you wish to participate. If you agree to take part, we will ask you to sign a consent form. You are free to withdraw at any time, without giving a reason.

**What will happen to me if I take part and what will I have to do?**
The initial part of the study involves an interview with the researcher (lasting 30-60 minutes). The interview will be audio recorded and you will be asked questions about your experience with surgical pilot and feasibility work and your opinions about its conduct. The researcher will travel to meet you at convenient location to conduct this interview. We may then invite you to take part in a consensus process and/or meeting. We will give you plenty of notice about timings for this. If you take part in the consensus process, this may include, for example, questionnaires, surveys and group discussions to contribute to forming consensus recommendations for the design of pilot and feasibility work for surgical trials.

**Expenses and payments**
The researcher will make arrangements to travel to meet/speak with you at a convenient time and location to conduct the interview. Travel expenses incurred getting to a consensus meeting will be payable following attendance and refreshments will be provided.

**What are the possible disadvantages and risks of taking part?**
There are no anticipated risks associated with taking part in this study.

**What are the possible benefits of taking part?**
We hope that the results of this study will guide the future design and conduct of pilot and feasibility work for surgical trials. This will benefit the surgical trials community as a whole, and will ultimately also benefit patients.

**What will happen if I don't want to carry on with the study?**
You are free to withdraw from participating in the study without giving a reason. If the consensus meeting has taken place, we may be unable to remove you voice from the tape, but we will remove any data deduced from the recordings from the study.

**Will my taking part in this study be kept confidential?**
Data will be collected through audio-recordings and notes taken during the interviews. All data will be stored under the provisions of the National Data Protection Regulations and the University of Bristol requirements. Participants' contact details will be stored securely in a restricted access folder on password protected computers at the School of Social and Community Medicine, University of Bristol. The audio-recordings and interview data will be created using an encrypted digital recorder and the raw data will only be available to members of the study team. Interview data will be stored using coded study numbers to ensure that individuals cannot be identified from the file name alone. The hard copies of the consent forms and audio-recordings will be kept securely by the University of Bristol in a locked filing cabinet. They will be kept up to 10 years after termination of the study and then disposed of securely. Access to all data will be restricted to the project lead (Katherine Fairhurst) and the supervisors (Dr Kerry Avery, Professor Jane Blazeby, Miss Shelley Potter, Professor Carrol Gamble). Data from the study, excluding any personal data, may be seen and used by other researchers, for ethically approved research projects, on the understanding that confidentiality will be maintained.

**What will happen to the results of the research study?**
The results of the study will be disseminated through the resulting PhD thesis and via normal academic channels including for example, journal publication, presentation at scientific/academic meetings/conferences and talks to academic/research audiences. Participants will be anonymised and every effort made to ensure they will not be identifiable in any written or verbal reports from the research. In situations where participants may potentially be identifiable, their agreement will be gained for their data to be used in publication/presentation. When disseminating findings, no names of individuals or places will appear in any form of publication.

**Who is organising and funding the research?**
Katherine Fairhurst is the lead researcher who is an MRC Clinical Research Fellow. This research forms part of a PhD thesis, supervised by Dr Kerry Avery, Professor Jane Blazeby, Miss Shelley Potter and Professor Carrol Gamble at the University of Bristol. The research is funded by the MRC Network of Hubs for Trials Methodology Research (MRC HTMR).

**Who has reviewed the study?**
The Faculty of Health Sciences Research Ethics Committee at the University of Bristol has reviewed and approved this research.

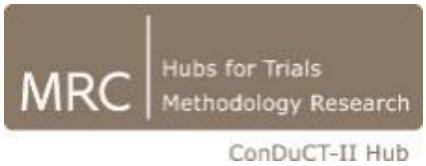
**If you would like to take part in this study, have any further questions or require further information please do not hesitate to contact:**

Miss Katherine Fairhurst
MRC Clinical Research Fellow, MRC ConDuCT-II Hub PhD Student, General Surgery Registrar
Centre for Surgical Research, School of Social and Community Medicine, University of Bristol, Room 1.13, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS.
Tel: 0117 92 87386
Email: katherine.fairhurst@bristol.ac.uk

If you wise to raise a concern or complaint independently please contact:

Email: research-governance@bristol.ac.uk

## **Appendix VI** PEPSTAR Interview topic guide





Participant ID Number:

---

**Appendix 1: Semi structured interview topic guide**

---

**Exploring Perceptions and Experiences of Pilot work for Surgical Trials: A Qualitative Research Study (PEPSTAR)**

**1. Introduction**
Your name & role
Purpose and aims of the study
Test audio recorder
Explain will record interview
Written consent

**2. General Information**
Current & previous roles
e.g. surgeon, university, editor, funder, REC, trial team etc
Specific areas of research interest

**3. Topic areas** (Introduce each area as necessary)

A) General
What is (you/in general) pilot work (PW)?
What is (you/in general) feasibility work (FW)?
What role does it have?
What should its role be?
What issues should PW/FW consider?
Generic? (E.g. recruitment, retention, eligibility, protocol adherence, outcome assessment – selection/data collection & quality)
Specific to surgery? (E.g. developing innovation, standardisation, leaning curves, quality assurance, co-interventions, preferences/acceptability, blinding)
When would you do internal/external pilot?

B) Importance
How important is PW?
Why is it important?
Any examples of how/why important

C) Experience
Examples of specific studies
Why i.e. purpose/rationale
Objectives at start – did they change?
Why internal/external/both?
Participant's role
Design
How design decided? By who?
Did design change?
How? Why?
Conduct
Discussions with funders
When? Why? Outcome?

D) Impact
On the main trial
Design/Conduct/Funding

E) Challenges/difficulties
What challenges in general with PW/FW?
*Why* are they challenges?
Examples of specific challenges/difficulties
What were these?
Why did they occur?
How dealt with?
Learning points?
Do anything differently?
Any remaining uncertainties about a main trial?

F) Essential components
Any key components to the design?
Progression criteria?
Any key processes for conduct?
Who needs to be involved?

**4. Closing the interview**
Thank them for their time - Any further questions/comments?
Re-iterate may be in touch about inviting them to take part in a consensus process

Example questions:

**A) General**

- What do you/other people think PW/FW is?
- Can you expand on the meanings of the terms IP/EP and NRFW?
- What role should PW/FW have in the context of trial design?
- What role does PW/FW actually have currently?
- What sort of questions (about whether a MT is possible) should PW/FW be answering?
  OR
- What is important for PW/FW to show before you go ahead with a MT?
- How would you choose when to do an internal or external pilot?

**B) Importance**

- Do you think it is important to always do PW/FW before a MT?
  - Why do you think it is important?
  - Do you have any examples of when PW/FW might not be necessary?
- Are you able to give any examples of how important PW/FW is from your own/others work?

**C) Experience**

- Could we talk about some of your own experiences of designing/conducting/funding PW/FW?
  - What was the rationale for doing X study?
  - Why was it important to do a pilot/feasibility study?
  - Why did you choose to do an internal/external pilot?
- How was the design of X study decided?
  - Who was involved in the design?
  - Did the design change in the course of the study?
  - What happened?
  - Why did the design change?
- Could you describe any positive or negative experiences you had with funding bodies in the context of pilot/feasibility studies?
  - Any difficulties getting funding for X study?
  - Why do you think there were difficulties?

**D) Impact**

- How did the PW (IP/EP)/FW impact on the design of the MT?
- How did the PW (IP/EP)/FW impact on the conduct of the MT?
- Did the PW (IP/EP)/FW have any impact on getting/not getting funding for the MT?

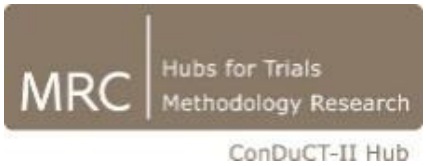
**E) Challenges**

- We have discussed some of these already, but could you summarise in your opinion, what you think the main challenges of doing (designing/conducting) PW/FW are?
  - Why do you think there are these challenges?
  - How did you/the trial team deal with these challenges?
  - Would you do anything differently if you were planning the PW/FW study again?
- Were there any remaining uncertainties about the MT once the PW/FW was completed?

**F) Essential components**

- Again we have talked about some important components to PW/FW already, but could you summarise/comment on what in your opinion, you think the key components to designing successful pilot/feasibility studies are?
- Do you have any experience of determining progression criteria for PW?
  - What do you think these should be? What are their purpose?
  - How should they be decided? How are they decided in your experience?
  - Do you have examples where an IP has progressed to a MT with PC being met/not met?
  - Do you think there are any specific challenges around selecting and using PC for IP studies?
- Do you think there are any key processes/personnel for the conduct of PW/FW?

## **Appendix VIa** PEPSTAR Revised interview topic guide





Participant ID Number:

**Appendix 1: Semi structured interview topic guide**

**Exploring Perceptions and Experiences of Pilot work for Surgical Trials: A Qualitative Research Study (PEPSTAR)**

**1. Introduction**
Your name & role
Purpose and aims of the study
Test audio recorder
Explain will record interview
Written consent

**2. General Information**
Current & previous roles
e.g. surgeon, university, editor, funder, REC, trial team etc
Specific areas of research interest

**3. Topic areas** (Introduce each area as necessary)

A) General
What is (you/in general) PW/FW?
What is (you/in general) IP/EP/NRFW?
When would you do internal/external pilot?
What role does it have?
What should its role be?
What issues should PW/FW consider?
Generic? (E.g. recruitment, retention, eligibility, protocol adherence, outcome assessment – selection/data collection & quality)
Specific to surgery? (E.g. developing innovation, standardisation, leaning curves, quality assurance, co–interventions, preferences/acceptability, blinding)
Progression criteria?
What/how decided/always met?/challenges

B) Importance
How important is PW?
Why is it important?
Any examples of how/why important
How impacts on design/conduct/funding

C) Experience
Examples of specific studies
Why i.e. purpose/rationale
Objectives at start – did they change?
Why internal/external/both?
Participant's role
Design
How design decided? By who?
Did design change?
How? Why?
Conduct
Discussions with funders
When? Why? Outcome?

D) Challenges/difficulties
What challenges in general with PW/FW?
*Why* are they challenges?
Examples of specific challenges/difficulties
What were these? Methodological/cultural?
Why did they occur?
How dealt with?
Learning points?
Do anything differently?
Any remaining uncertainties about a main trial?

E) Essential components
Any key components to the design? (Cost effectiveness?)
Any key processes for conduct?
Who needs to be involved?

**4. Closing the interview**
Thank them for their time - Any further questions/comments?
Re-iterate may be in touch about inviting them to take part in a consensus process

Example questions:

**F) General**

- What do you/other people think PW/FW is?
- Can you expand on the meanings of the terms IP/EP and NRFW?
- How would you choose when to do an internal or external pilot?
- What role should PW/FW have in the context of trial design?
- What role does PW/FW have currently?
- What are the reasons for doing PW/FW?
  OR What is important for PW/FW to show before you go ahead with a MT?
  OR What sort of questions (about whether a MT is possible) should PW/FW be answering?
- Where does your understanding of PAFS come from?
- Are you aware of any guidance on the design and conduct of PAFS?
  - What guidance would you find helpful? Who should create/endorse that guidance?
- Do you have any experience of determining progression criteria for PW?
  - What do you think these should be? What are their purpose?
  - How should they be decided? How are they decided in your experience?
  - Do you have examples where an IP has progressed to a MT with PC being met/not met?
  - Do you think there are any specific challenges around selecting and using PC for IP studies?

**G) Importance & Impact**

- Do you think it is important to always do PW/FW before a MT?
  - Why do you think it is important? Examples of importance is from your own/others work?
  - When might PW/FW might not be necessary? Examples?
  - Do you think a PW/FW showing a MT is not viable is valuable?
- How does PW/FW impact on MT design/conduct/funding?
- Should PAFS be published? Why?

**H) Experience**

- Could we talk about some of your own experiences of designing/conducting/funding PW/FW?
  - What was the rationale for doing X study?
  - Why did you choose to do an internal/external pilot?
- How was the design of X study decided?
  - Who was involved in the design?
  - Did the design change during the study? Why?
- Could you describe any positive or negative experiences you had with funding bodies in the context of pilot/feasibility studies?
  - Any difficulties getting funding for X study?
  - Why do you think there were difficulties?

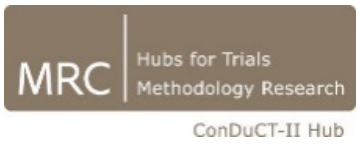
**I) Challenges**

- We have discussed some of these already, but could you summarise in your opinion, what you think the main challenges of doing (designing/conducting) PW/FW are?
  - Methodological/cultural?
  - Why do you think there are these challenges?
  - How did you/the trial team deal with these challenges?
  - Would you do anything differently if you were planning the PW/FW study again?
- Were there any remaining uncertainties about the MT once the PW/FW was completed?

**J) Essential components**

- We have talked about some important components to PW/FW already, but could you summarise/comment on what in your opinion, you think the key components to designing successful pilot/feasibility studies are?
- Do you think it is important to consider cost effectiveness in PW/FW?
- Do you think there are any key processes/personnel for the conduct of PW/FW?

## **Appendix VII** PEPSTAR Consent form





Participant ID Number:

**Appendix 6: CONSENT FORM A for study participants**

**Exploring Perceptions and Experiences of Pilot work for Surgical Trials: A Qualitative Research Study (PEPSTAR)**

Name of Researcher: **Katherine Fairhurst**

Contact details: **katherine.fairhurst@bristol.ac.uk, 0117 92 87386**

Please initial all boxes

1. I confirm that I have read and understand the participant information sheet A dated **August 2016** version **1.0** for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. I understand that I can decline from answering a question during the interview and that I am free to withdraw my consent for data to be used at a later date if I wish. ☐
3. I understand that I will be interviewed, that the interview will be audio-recorded and that notes will be documented by the researcher. ☐
4. I understand that all audio recordings, transcripts and copies of study documents will be stored on a secure computer and/or in a locked filing cabinet at the University of Bristol. ☐
5. I understand that all data collected will be treated as confidential and that where possible I will be anonymous in written reports from the research. In situations where I might potentially be identified, the research team will contact me to secure my agreement for a data extract (e.g. a verbatim quote) to be used. ☐
6. I understand that only members of the research team will have access to my personal data and that data from the study, excluding my personal data, may be seen and used by other researchers, for ethically approved research projects, on the understanding that confidentiality will be maintained. ☐
7. I agree to take part in the above study. ☐
8. I agree to being contacted about my participation in a consensus process in the future. ☐

| Name of participant | Date | Signature |
|---|---|---|
| Name of person taking consent | Date | Signature |

**Appendix VIII** Published paper in the British Journal of Surgery

Review

# Value of surgical pilot and feasibility study protocols

K. Fairhurst[1], J. M. Blazeby[1], S. Potter[1], C. Gamble[2], C. Rowlands[1] and K. N. L. Avery[1]

[1]Centre for Surgical Research and Medical Research Council (MRC) ConDuCT-II Hub for Trials Methodology Research, Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, and [2]MRC North West Hub for Trials Methodology Research, University of Liverpool, Liverpool, UK

*Correspondence to:* Miss K. Fairhurst, Centre for Surgical Research, Department of Population Health Sciences, Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK (e-mail: katherine.fairhurst@bristol.ac.uk)

**Background:** **RCTs in surgery are challenging owing to well established methodological issues. Well designed pilot and feasibility studies (PFS) may help overcome such issues to inform successful main trial design and conduct. This study aimed to analyse protocols of UK-funded studies to explore current use of PFS in surgery and identify areas for practice improvement.**

**Methods:** **PFS of surgical interventions funded by UK National Institute for Health Research programmes from 2005 to 2015 were identified, and original study protocols and associated publications sourced. Data extracted included study design characteristics, reasons for performing the work including perceived uncertainties around conducting a definitive main trial, and whether the studies had been published.**

**Results:** **Thirty-five surgical studies were identified, of which 29 were randomized, and over half (15 of 29) included additional methodological components (such as qualitative work examining recruitment, and participant surveys studying current interventions). Most studies focused on uncertainties around recruitment (32 of 35), with far fewer tackling uncertainties specific to surgery, such as intervention stability, implementation or delivery (10 of 35). Only half (19 of 35) had made their results available publicly, to date.**

**Conclusion:** **The full potential of pretrial work to inform and optimize definitive surgical studies is not being realized.**

Presented to the International Clinical Trials Methodology Conference/Annual Meeting of the Society for Clinical Trials, Liverpool, UK, May 2017, and the Annual Meeting of the Society for Clinical Trials, Portland, Oregon, USA, May 2018; published in abstract form as *Trials* 2017; **18**(Suppl 1): P16 and *Clinical Trials* 2018; **15**(Suppl 2): 118 (A82)

Paper accepted 12 February 2019

Published online 10 May 2019 in Wiley Online Library (www.bjs.co.uk). **DOI:** 10.1002/bjs.11167

## Introduction

High-quality RCTs are necessary to inform evidence-based surgical practice. Surgical trials are challenging to do owing to established methodological issues extending beyond those of RCTs in other areas[1–4]. Challenges specific to surgical RCTs include uncertainties around the stability and/or standardization of the intervention, selection and/or measurement of relevant clinical and patient-reported outcomes, and issues surrounding patient recruitment such as clinician and patient equipoise[5–7]. These specific challenges, combined with surgeons' lack of familiarity with RCTs and the perception that participation can be onerous, may mean that trials are not initiated, conducted efficiently or completed on time and on target.

Well designed pilot and feasibility studies (PFS) may help to overcome challenges associated with undertaking RCTs in surgery, by allowing uncertainties to be addressed, and optimal design of the main trial to be determined[8–10]. Previous methodological work has considered the use and misuse of PFS in general. Several checklists have been developed[11–16] to identify and categorize specific reasons for undertaking pilot work, with the aim of guiding researchers into considering these for their study. More recently, guidelines for improved reporting of PFS, in the form of an extension to the CONSORT statement[17], have been published. Integral to the development of these guidelines was the publication of a conceptual framework[18] to promote understanding by defining the purpose and scope of different types of PFS in preparation for RCTs. These guidelines were, however, mostly theoretical,

with no specific recommendations for the use of PFS in surgery.

Some published recommendations have provided more practical guidance regarding the use of PFS in surgery. The Medical Research Council (MRC) framework[19] for developing and evaluating all complex interventions (defined as interventions with multiple components acting both independently and interdependently) includes surgery. Within this framework, undertaking PFS before full-scale evaluation of surgical interventions in a definitive trial is considered vital preparatory work[19]. The IDEAL (Idea, Development, Exploration, Assessment and Long-term follow-up) framework[20–22] also provides recommendations specific to the evaluation of novel surgical interventions from first in man to long-term studies, and suggests study designs and issues to be considered at each stage of evaluation. Both the MRC and, more specifically for surgery, the IDEAL recommendations discuss PFS as part of a larger framework for the development and assessment of new complex interventions, and list several elements of the design of PFS.

Although PFS are thought to be beneficial and are endorsed as part of the strategic guidance discussed, there is uncertainty regarding exactly how they influence the design and conduct of successful RCTs[22]. At a fundamental level, published literature suggests that the wider surgical community may not understand the concept of PFS, with evidence of small, underpowered RCTs often mislabelled as pilot or feasibility studies[11,14,23]. Such studies often fail to address baseline feasibility issues such as considering whether a main trial is possible, and instead focus on formal hypothesis testing[24–28]. Further work is therefore needed to understand when and how PFS may be used optimally to inform future main trials in surgery.

Development of guidance that addresses the challenges associated with undertaking surgical PFS requires understanding of the current use of such studies in surgery. As PFS are often poorly reported, a traditional systematic review of surgical PFS is unlikely to be informative beyond what is already known. Major research funders are increasingly recognizing the importance of well designed PFS in informing main trial design. It was therefore hypothesized that protocols of competitively funded PFS may provide more informative insights into the current use of PFS in surgery, and into how pretrial work may be used to inform future definitive studies. The aim of this paper was to analyse protocols of successfully funded surgical PFS. The purpose of this review was to identify how PFS are currently used, consider whether their use is appropriate, and envisage how the use of PFS could be further improved.

## Methods

A systematic analysis of the protocols of PFS of surgical interventions funded by the UK National Institute for Health Research (NIHR) was undertaken.

### Characteristics of information sources, search strategy and screening

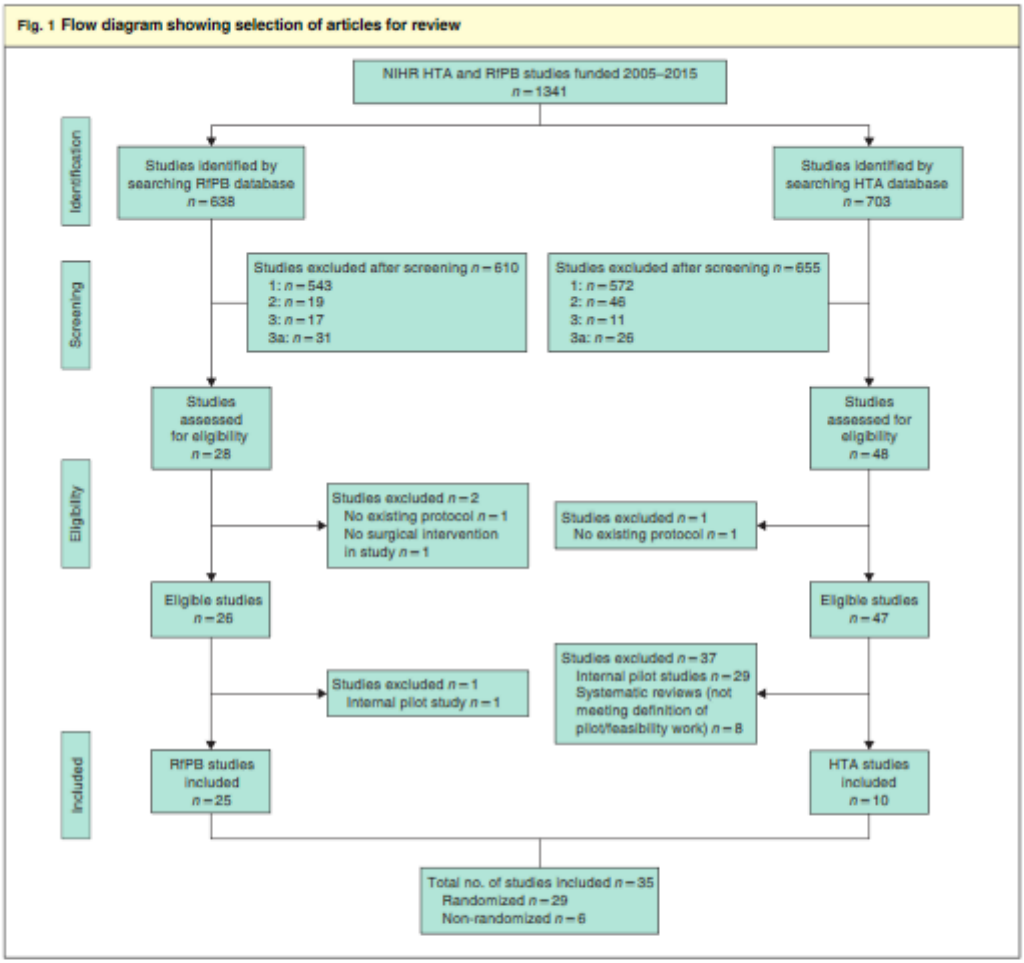
The UK NIHR Health Technology Assessment (HTA) and Research for Patient Benefit (RfPB) programmes were selected to identify PFS of surgical interventions. These programmes are established major national funders of high-quality patient-centred research, having funded trials for 25 and 12 years respectively. Both programmes fund definitive evaluations of the clinical and cost-effectiveness of interventions, as well as feasibility studies to inform future definitive trials. They have publicly available and searchable databases of funded studies[29,30]. Given the scope and longevity of both programmes, it was hypothesized that each would have funded surgical PFS, providing a sample of potentially well designed work from which to explore the role of PFS in surgery, and study their impact on main trial design and conduct.

The HTA and RfPB databases were searched for surgical PFS. Titles and abstracts were screened in duplicate, with any issues resolved by discussion and/or with senior input where necessary. Agreement was reached on inclusion or exclusion for 1283 studies (95·7 per cent) at the first attempt. Protocols for all included HTA studies were downloaded from the HTA website and those for all included RfPB studies (apart from 1 available online) were obtained by contacting the chief investigator of each study directly. Additional publications relating to included studies were identified by searching for links to published outputs on the NIHR website (HTA only), and using the study title, acronym and chief investigator name to search on PubMed, Google Scholar and the ISRCTN trials registry online.

### Inclusion and exclusion criteria

Protocols of all surgical PFS funded by the NIHR HTA and RfPB programmes between 1 January 2005 and 31 December 2015 were included. In the absence of universally adopted definitions of surgical interventions and PFS, for the purposes of this review, pilot/feasibility work was defined as: any research that is undertaken before a main study and is explicitly intended to inform the design and/or conduct of a future main study, where main study is defined as a definitive study or RCT of an intervention(s). A surgical intervention was defined as: a

**Fig. 1 Flow diagram showing selection of articles for review**




Reasons for exclusion: 1, not a surgical intervention; 2, surgical intervention, not pilot/feasibility work; 3, surgical intervention is a co-intervention; 3a, surgical intervention is a co-intervention and not pilot/feasibility work. NIHR, National Institute for Health Research; HTA, Health Technology Assessment; RfPB, Research for Patient Benefit.

diagnostic, therapeutic or adjunctive invasive intervention performed by a trained clinician, using hands, instruments and/or devices, and included operative, radiological and endoscopic procedures.

Internal pilot studies were excluded owing to growing opinion among trial methodologists that internal pilots do not meet the true definition of pilot studies[31]. This is because internal pilots are very distinct from external pilots in their methodology, being designed and funded as a part of a main trial with all data generated from this first phase contributing to the final analysis. Internal pilots are, therefore, most often used when no substantive changes to key components of the trial, such as the intervention or outcomes, are anticipated. In addition, study protocols of RCTs with an internal pilot phase usually include only limited detail regarding the internal pilot phase itself, such as a list of proposed progression criteria. It was therefore

considered that a review of trial protocols with an integrated internal pilot phase would be of limited value for the purpose of this work. Also excluded were funded systematic reviews that did not state any intention to inform a future definitive study, and studies that focused on the evaluation of co-interventions to surgery, for example the administration of anaesthetic drugs, and postoperative rehabilitation or enhanced recovery programmes. This was because the primary focus of this work was to explore the specific difficulties surrounding studies of surgical interventions.

### Data extraction and analyses

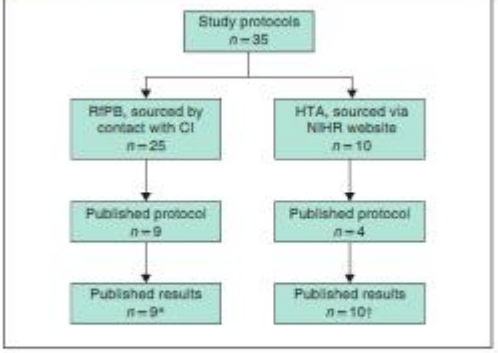
Data were extracted using a standard database developed in Microsoft Excel® (Microsoft, Redmond, Washington, USA), including general study characteristics, available data sources in addition to the study protocol (published papers) and the surgical specialty of the study. The data extraction form is available in *Appendix S1* (supporting information). Details of the study design (randomized or non-randomized; quantitative or qualitative) and conduct, including characteristics of the patient population, were extracted. A framework was developed for capturing the uncertainties and challenges regarding the viability of a future main trial, informed by expert knowledge and previous methodological work regarding the design, definitions and reporting of PFS[11–18,32–34], published MRC guidelines[19] and the IDEAL framework[20,21]. All possible reasons identified for undertaking PFS were grouped into five key domains: main trial design; logistics; recruitment; intervention; and outcomes. The domains were constructed and ordered according to how extraction progressed, with cross-checking between authors. Special consideration was given to uncertainties and challenges considered more specific and/or relevant to surgical trials. Results were analysed in Microsoft Excel®. Descriptive statistics are reported with comparison between the HTA and RfPB, and randomized and non-randomized cohorts where relevant.

## Results

### Screening

A total of 1341 funded studies were identified (703 HTA studies, 52·4 per cent), and 35 eligible studies (25 RfPB, 10 HTA) were included in the final analysis (*Fig. 1*). Additional data sources were available for 25 of the 35 studies; these included a published protocol paper for 13 of 35 studies and a paper reporting the PFS findings for 19 (*Fig. 2*). Publication rates of the study protocols were similar between

**Fig. 2 Data sources available for the pilot and feasibility studies included in the review**




*Published in peer-reviewed journal. †Health Technology Assessment (HTA) report +/– other paper(s) in peer-reviewed journal. RfPB, Research for Patient Benefit; CI, chief investigator; NIHR, National Institute for Health Research.

funders (HTA 4 of 10, RfPB 9 of 25), although the results of HTA studies were published more often than those of RfPB studies (HTA 10 of 10, RfPB 9 of 25).

### Study characteristics

Some 29 of 35 studies were randomized and six were non-randomized (*Fig. 1*), including qualitative work, national audit and cohort studies (*Table 1*). Randomized studies were more often multicentre and conducted by larger trial teams than non-randomized studies. Over half of the randomized studies (15 of 29) also included other types of pretrial work.

### Reasons for conducting pilot/feasibility studies

Reasons for performing PFS are summarized in *Table 2*. Addressing uncertainties around trial recruitment was cited as the most common reason (32 of 35 studies; RfPB 23 of 25, HTA 9 of 10), followed by the overarching aim to determine whether a main trial was possible or necessary (27 of 35 studies; RfPB 17 of 25, HTA 10 of 10). Logistical issues (such as trial paperwork, resources needed, running multicentre studies) were considered in two-thirds of studies (23 of 35 studies; RfPB 16 of 25, HTA 7 of 10) and outcomes (for example selecting the primary outcome, determining important outcomes for patients) in less than half (15 of 35 studies; RfPB 11 of 25, HTA 4 of 10). Around half of the studies considered issues regarding the sample size for the main trial (for example assessing variability in

**Table 1 Design characteristics of the 35 pilot/feasibility studies in the review**

| | Randomized studies (*n* = 29) | | | Non-randomized studies (*n* = 6) | | |
|---|---|---|---|---|---|---|
| | RfPB (*n* = 22) | HTA (*n* = 7) | Total (*n* = 29) | HTA (*n* = 3) | RfPB (*n* = 3) | Total (*n* = 6) |
| **Surgical specialty of study** | | | | | | |
| Gastrointestinal | 8 | 1 | 9 | 2 | 0 | 2 |
| Urology | 3 | 2 | 5 | 0 | 0 | 0 |
| Cardiothoracic | 3 | 0 | 3 | 0 | 0 | 0 |
| Orthopaedic | 2 | 2 | 4 | 0 | 0 | 0 |
| Obstetrics/gynaecology | 2 | 0 | 2 | 0 | 1 | 1 |
| Maxillofacial/ENT | 2 | 1 | 3 | 0 | 1 | 1 |
| Plastics | 1 | 0 | 1 | 0 | 0 | 0 |
| Paediatrics | 1 | 1 | 2 | 0 | 1 | 1 |
| Breast | 0 | 0 | 0 | 1 | 0 | 1 |
| **No. of centres*** | 3 (1–23) | 4 (2–10) | 3 (1–23) | 1 (1) | 2 (1–65) | 1 (1–65) |
| 1 | 7 | 0 | 7 | 3 | 1 | 4 |
| 2–20 | 14 | 7 | 21 | 0 | 1 | 1 |
| >20 | 1 | 0 | 1 | 0 | 1 | 1 |
| **Proposed no. of participants in study*** | 50 (30–200) | 70 (60–144) | 60 (30–200) | n.a. | n.a. | n.a. |
| **Patient characteristics** | | | | | | |
| Age | | | | | | |
| Adults | 19 | 5 | 24 | 2 | 2 | 4 |
| Children | 2 | 1 | 3 | 1 | 0 | 1 |
| Both | 1 | 1 | 2 | 0 | 1 | 1 |
| Sex | | | | | | |
| Male | 1 | 1 | 2 | 0 | 0 | 0 |
| Female | 3 | 0 | 3 | 1 | 1 | 2 |
| Both | 18 | 6 | 24 | 2 | 2 | 4 |
| Country | | | | | | |
| UK | 20 | 7 | 27 | 2 | 3 | 5 |
| Europe | 2 | 0 | 2 | 0 | 0 | 0 |
| Worldwide | 0 | 0 | 0 | 1 | 0 | 1 |
| **No. of personnel in trial team*** | 6·5 (4–19) | 9·5 (0–24) | 9 (0–24) | 1 (0–12) | 7 (2–25) | 5·5 (0–25) |
| **Non-randomized pretrial work** | 9 | 6† | 15† | 3 | 3 | 6‡ |
| Qualitative interviews | 8 | 6 | 14 | 2 | 2 | 4 |
| Participant/researcher survey | 1 | 1 | 2 | 2 | 3 | 5 |
| Economic modelling | 0 | 1 | 1 | 0 | 0 | 0 |
| Systematic review | 0 | 0 | 0 | 1 | 0 | 1 |
| National audit | 0 | 0 | 0 | 0 | 1 | 1 |
| Cohort study | 0 | 0 | 0 | 0 | 1 | 1 |

*Values are median (range). †Some studies planned more than one type of non-randomized work; ‡most studies planned more than one type of non-randomized work. RfPB, Research for Patient Benefit; HTA, Health Technology Assessment; ENT, ear, nose and throat; n.a., not applicable.

outcomes) (19 of 35 studies; RfPB 17 of 25, HTA 2 of 10) and costs/funding for the main trial (16 of 35 studies; RfPB 14 of 25, HTA 2 of 10).

Eleven of the 35 studies (all RfPB-funded) aimed to collect data regarding the safety or effectiveness of an intervention to inform the main trial and, of these, almost three-quarters (8 of 11) specified plans for formal hypothesis testing by comparing the intervention(s) and/or control groups to test effectiveness and/or safety, which is not recommended for PFS.

One-quarter (10 of 35 studies; RfPB 6 of 25, HTA 4 of 10) sought to explore uncertainties around the surgical intervention itself, such as intervention development, stability, delivery and the surgical learning curve. Of these ten studies, six were considering surgery *versus* no surgery, and four were considering a novel surgical technique *versus* an established method (*Table 3*). Of the ten studies specifically planning to evaluate a new surgical technique, only four aimed to address uncertainties surrounding the intervention. These uncertainties included: documenting the

**Table 2** Rationale as detailed in the protocol for the pilot and feasibility studies of surgical interventions included in the systematic analysis

| Area examined | Rationale | No. of studies stating each rationale in the study protocol | | | No. of studies stating examination of each area in the study protocol | | |
|---|---|---|---|---|---|---|---|
| | | RfPB (*n* = 25) (3 NR) | HTA (*n* = 10) (3 NR) | Total (*n* = 35) (6 NR) | RfPB (*n* = 25) | HTA (*n* = 10) | Total (*n* = 35) |
| **Main trial design** | | | | | | | |
| Main trial possible /necessary | To examine and test whether a main trial is possible | 14 (2) | 8 (1) | 22 (3) | | | |
| | To assess whether main trial is needed and/or produce a protocol | 3 (0) | 0 (2) | 3 (2) | **17** | **10** | **27** |
| | To test whether the protocol can be adhered to and modify it as necessary | 2 (0) | 2 (0) | 4 (0) | | | |
| Sample size | To estimate the variability in outcomes to help determine a sample size for the main trial | 15 (1) | 2 (0) | 17 (1) | **17** | **2** | **19** |
| | To determine a sample size for the main trial | 3 (0) | 0 (0) | 3 (0) | | | |
| Costs/funding | To assess/gather information on costs of performing the trial (direct and indirect) | 2 (0) | 0 (0) | 2 (0) | | | |
| | To perform/prepare for a cost-effectiveness analysis of the intervention(s) | 13 (1) | 2 (0) | 15 (1) | **14** | **2** | **16** |
| | To provide information/evidence to funders | 1 (0) | 0 (0) | 1 (0) | | | |
| Safety and effectiveness data | Preliminary data on safety to inform a main trial | 2 (0) | 0 (0) | 2 (0) | | | |
| | Information on adverse events | 4 (0) | 0 (0) | 4 (0) | | | |
| | Planned formal hypothesis testing of safety outcomes* | 3 (1) | 0 (0) | 3 (1) | **11** | **0** | **11** |
| | Preliminary data on effectiveness to inform a main trial | 0 (0) | 0 (0) | 0 (0) | | | |
| | Planned formal hypothesis testing of effectiveness outcomes* | 7 (1) | 0 (0) | 7 (1) | | | |
| **Logistics** | To test the logistics of multicentre studies | 5 (0) | 1 (0) | 6 (0) | | | |
| | To develop a research network as a resource for a future main trial | 1 (1) | 0 (0) | 1 (1) | | | |
| | To develop/test patient information content/forms/methods of delivery | 1 (0) | 3 (0) | 4 (0) | | | |
| | To develop/test data collection forms/methods | 13 (1) | 6 (1) | 19 (2) | | | |
| | To develop/test questionnaires/surveys | 5 (0) | 1 (0) | 6 (0) | | | |
| | To test response rates to questionnaires/surveys | 0 (0) | 0 (0) | 0 (0) | **16** | **7** | **23** |
| | To prepare/plan/assess monitoring procedures | 0 (0) | 1 (0) | 1 (0) | | | |
| | To determine what resources are needed for a main trial (e.g. funding/staff) | 3 (0) | 0 (0) | 3 (0) | | | |
| | To assess the logistics of delivering an intervention as part of a trial in the NHS | 1 (0) | 0 (0) | 1 (0) | | | |
| | To test (novel) methods of blinding | 1 (0) | 1 (0) | 2 (0) | | | |
| | To assess proposed data analysis techniques | 1 (0) | 1 (0) | 2 (0) | | | |
| | To learn about the day-to-day running of a trial | 1 (0) | 0 (0) | 1 (0) | | | |

**Table 2 Continued**

| Area examined | Rationale | No. of studies stating each rationale in the study protocol: RfPB (*n* = 25) (3 NR) | HTA (*n* = 10) (3 NR) | Total (*n* = 35) (6 NR) | No. of studies stating examination of each area in the study protocol: RfPB (*n* = 25) | HTA (*n* = 10) | Total (*n* = 35) |
|---|---|---|---|---|---|---|---|
| **Recruitment** | To test/modify inclusion/exclusion/eligibility criteria | 2 (0) | 0 (1) | 2 (1) | | | |
| | To estimate the expected prevalence or rate of incident cases in the population | 1 (1) | 1 (0) | 2 (1) | | | |
| | To estimate the number to be screened and proportions of eligible patients | 9 (0) | 3 (0) | 12 (0) | | | |
| | To assess numbers/rates of recruitment and consent | 17 (0) | 5 (0) | 22 (0) | **23** | **9** | **32** |
| | To test the randomization procedure | 5 (0) | 3 (0) | 8 (0) | | | |
| | To test the acceptability of randomization/trial design | 12 (1) | 5 (2) | 17 (3) | | | |
| | To determine the acceptability of the intervention to clinicians and patients | 12 (1) | 4 (2) | 16 (3) | | | |
| | To assess rates of retention in the study | 11 (0) | 2 (0) | 13 (0) | | | |
| **Intervention** | To assess and monitor the development of an intervention and/or its stability | 2 (1) | 1 (0) | 3 (1) | | | |
| | To develop and test the implementation and delivery of the intervention | 1 (0) | 3 (0) | 4 (0) | | | |
| | To train staff in delivery and assessment procedures | 1 (0) | 0 (0) | 1 (0) | | | |
| | To monitor the surgical learning curve | 2 (1) | 0 (0) | 2 (1) | **6** | **4** | **10** |
| | To test rates of crossover | 0 (0) | 1 (0) | 1 (0) | | | |
| | To examine reasons for non-adherence/crossover for the main trial | 2 (0) | 0 (0) | 2 (0) | | | |
| | To develop pathways and protocols for co-interventions | 0 (0) | 0 (0) | 0 (0) | | | |
| **Outcome** | To select the most appropriate primary outcome measure | 9 (1) | 0 (0) | 9 (1) | | | |
| | To develop and test a new outcome measure | 0 (0) | 0 (0) | 0 (0) | **11** | **4** | **15** |
| | To determine appropriate/important/suitability of outcome measures for patients/clinicians | 3 (0) | 4 (2) | 7 (2) | | | |

Values in parentheses are number of non-randomized (NR) studies. *Formal hypothesis testing to demonstrate the safety and/or effectiveness of an intervention is generally not recommended for pilot and feasibility studies because of the underpowered sample size (see discussion). RfPB, Research for Patient Benefit; HTA, Health Technology Assessment; NHS, National Health Service.

**Table 3 Type of trial for studies that examined details of the surgical intervention and those that did not**

| | Studies examining the intervention (*n* = 10) | Studies not examining the intervention (*n* = 25) | Total (*n* = 35) |
|---|---|---|---|
| Surgery *versus* no surgery | 6 | 9 | 15 |
| New/novel surgical technique *versus* surgery | 4 | 6 | 10 |
| Non-randomized pilot/feasibility work | 0 | 5 | 5 |
| Surgery *versus* surgery (both established techniques) | 0 | 4 | 4 |
| Surgery *versus* placebo and no surgery (2 arms) | 0 | 1 | 1 |

technical development of the new intervention to inform the development of a competency assessment tool for surgeon performance evaluation before participation in the main trial; considering the feasibility of training and implementing the new technique; determining the variation in the type of new procedure performed across the UK; and considering the impact of the learning curve on adverse outcomes to inform entry criteria for a main trial.

## Discussion

This study demonstrated that the full potential of PFS to address the uncertainties and challenges specific to undertaking surgical trials is yet to be realized. The reasons most often cited by authors for performing PFS reflect the targeting of uncertainties generic to trials in general, such as recruitment, and considering whether a main trial is possible. Less than one-third of surgical PFS explored challenges of specific relevance to designing and conducting trials in surgery, such as uncertainties around the stability or delivery of the surgical intervention itself. Notably, of the ten studies aiming to evaluate a novel surgical intervention, only four addressed uncertainties surrounding the procedure, such as development of the new intervention, implementation and delivery of the intervention, and the effect of the surgical learning curve. Of equal importance is the finding that the role of PFS in surgery is still often misunderstood, with nearly one-quarter of studies planning to conduct formal hypothesis testing. Results of PFS in surgery are frequently under-reported, with almost half not publishing the results to date, despite the majority having completed before 2018. These findings indicate that there is a need for guidance regarding the scope and optimal use of PFS to promote main trial success and prevent research waste.

There are several possible reasons for the findings observed in this study. Conceivably, there may be confusion among surgeons around the value of PFS. In addition, it is possible that the design of PFS in funding applications is skewed towards reasons perceived as important to funders, such as demonstrating adequate recruitment and the feasibility of completing a main trial to time and target. It is likely that most trialists would acknowledge that recruitment is paramount to study success. However, there may be a lack of awareness amongst applicants of the many other potential uncertainties that can compromise the success of a main trial, particularly those around the intervention.

Previous guidance[9,10,17–19,35,36] regarding the optimal design and conduct of PFS has been theoretical or generic. The IDEAL framework, for example, describes a pathway for new surgical interventions from first in man (stage 1) to long-term study (stage 4), with stage 2a (development) and 2b (exploration) studies considered to be PFS, focusing on addressing uncertainties before stage 3 assessment in a definitive RCT. The initial IDEAL publication[21], however, was largely theoretical with little practical guidance about how PFS should be performed. Recently published updated IDEAL recommendations[22] now provide some clarification regarding the role of PFS in surgery as a result of recognition that the original IDEAL guidance published in 2009[21] had little impact on the design and conduct of surgical PFS[37]. The updated IDEAL framework[22] suggests several feasibility issues to consider in stage 2a/2b studies, including estimating effect size, defining intervention quality and standards, evaluating learning curves, exploring subgroup differences, eliciting key stakeholder values and preferences, and analysis of adverse events. This list is far from exhaustive. The present analysis of protocols for NIHR-funded surgical PFS indicates that there are important additional issues regarding the design and conduct of surgical trials that may usefully be explored in PFS.

The findings of PFS are often not widely disseminated[11]. This may reflect journal editors' lack of appreciation of the value of pilot work[11,38] and concerns about the quality or methodological rigour of pilot work in general[11]. Introduction of open-access journals focused specifically on PFS such as *Pilot and Feasibility Studies* may address this[39]. Researchers themselves may also fail to prioritize the reporting and dissemination of findings from PFS, which may be fuelled further by the decision to pursue (or not) the main trial. Notably, however, all ten HTA-funded studies identified in this review published their results. This is because studies funded by HTA are required to publish a report in the peer-reviewed *Health Technology Assessment* journal.

The scope of PFS in general has historically often been narrow, focusing typically on issues relating to safety and efficacy[40,41] or recruitment[42]. A 2011 literature review[14] of 50 pilot RCTs demonstrated that only 56 per cent of the studies addressed methodological issues in any depth. Another literature review[11] of studies published between 2007 and 2008 found that up to 74 per cent of PFS performed and reported hypothesis testing for one or more variables (compared with 23 per cent (8 of 35) in the present study). There is now general acceptance that any suggestion of promise or significance should be reported with caution, given the underpowered sample size of most PFS[13,15,32,43,44], and the present examination of protocols perhaps demonstrates a growing, if not complete, understanding of this issue.

This study demonstrated some key limitations of PFS in surgery, but focused solely on studies funded by the NIHR HTA and RfPB funding streams. Other NIHR funding streams were considered but excluded as they do not commonly fund surgical research. Of the 29 studies funded via the 2012 NIHR-commissioned call for surgical research[45], for example, only four met the inclusion criteria of the present study. These four studies were funded by the RfPB programme and were therefore all included in this work. Although the NIHR HTA and RfPB programmes are the major funders of studies of surgical interventions in the UK, it is accepted that there are charities, for example the British Heart Foundation and Arthritis UK, which may fund such work. However, this would have been logistically challenging given that very few funders make study protocols publicly available. Conversely, the NIHR-funded protocols are likely to be of relatively high quality so may provide an overly positive perception of the quality of PFS in surgery.

Working collaboratively to design and perform pretrial work can deliver surgical PFS in a cost-effective and timely manner, and can advance the development of definitive studies[46]. High-quality PFS may also be resource- and cost-effective, preventing waste by averting futile main trials, or providing information to improve the design and conduct of the main trial[47]. The findings from this study indicate, however, that PFS in surgery are not currently used to their full potential.

Work is under way to incorporate the present study findings into a wider methodological project to develop guidelines to support surgical researchers in undertaking PFS. These are likely to build on existing guidance from the MRC, the IDEAL group and broader methodological work, but focus specifically on surgical PFS. It will be important to explore if and how PFS influence main trial design and conduct. This will be achieved by following these NIHR-funded PFS as they progress to main trials. The guidelines aim to provide researchers with clear and accessible information regarding how and when to undertake PFS, detail the key features to consider when designing and conducting PFS to inform a future main trial, and emphasize the importance of working collaboratively with trial methodologists to ensure that PFS address all uncertainties around future trial conduct accurately and wholly.

## Acknowledgements

As this is not a traditional systematic review, the study was not preregistered with an analysis plan in an independent, institutional registry.

The authors acknowledge the chief investigators of all the included RfPB studies who kindly supplied study protocols in order that this work could be completed: P. Bhandari, J. Blazeby, M. Birchall, J. de Caestecker, A. Care, M. Crundwell, T. Davis, A. Dhar, S. Eldabe, K. Franks, A. Goyal, D. Griffin, E. Hall, P. Hull, K. Ismail, C. Knowles, A. Marchbank, H. Mostafid, S. Pereira, T. Pinkney, S. Potter, R. Prasad, P. Roderick, A. Sutcliffe (protocol available online), M. Tickle and S. Verma. They also thank M. Ashton-Key and A. Blatch-Jones at the NIHR Evaluation Trials and Study Co-ordination Centre for assistance with the Microsoft Excel® database used for screening the HTA study cohort, and commenting on the protocol for this work.

This work was supported by a MRC Hub for Trials Methodology Research PhD studentship for K.F. J.M.B. and K.A. are supported by the NIHR Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol, and the MRC ConDuCT-II (Collaboration and innovation in Difficult and Complex randomized controlled Trials In Invasive procedures) Hub for Trials Methodology Research (MR/K0256/43/1). C.G. is supported by the MRC North West Hub for Trials Methodology Research at the University of Liverpool (MR/L004933/1). The views expressed are those of the authors and not necessarily those of the UK National Health Service, NIHR, Department of Health or MRC. S.P. is an NIHR Clinician Scientist and J.M.B. is an NIHR Senior Investigator.

*Disclosure:* The authors declare no conflict of interest.

## References

1 Ergina PL, Cook JA, Blazeby JM, Boutron I, Clavien PA, Reeves BC *et al.* Challenges in evaluating surgical innovation. *Lancet* 2009; **374**: 1097–1104.
2 Blencowe NS, Cook JA, Pinkney T, Rogers C, Reeves BC, Blazeby JM. Delivering successful randomized controlled trials in surgery: methods to optimize collaboration and study design. *Clin Trials* 2017; **14**: 211–218.
3 Cook JA. The challenges faced in the design, conduct and analysis of surgical randomised controlled trials. *Trials* 2009; **10**: 9.
4 McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *BMJ* 2002; **324**: 1448–1451.
5 Donovan JL, de Salis I, Toerien M, Paramasivan S, Hamdy FC, Blazeby JM. The intellectual challenges and emotional consequences of equipoise contributed to the fragility of recruitment in six randomized controlled trials. *J Clin Epidemiol* 2014; **67**: 912–920.
6 Donovan JL, Paramasivan S, de Salis I, Toerien M. Clear obstacles and hidden challenges: understanding recruiter

perspectives in six pragmatic randomised controlled trials. *Trials* 2014; **15**: 5.
7 Rooshenas L, Elliott D, Wade J, Jepson M, Paramasivan S, Strong S *et al.* ACST-2 study group; By-Band-Sleeve study group; Chemorad study group; CSAW study group; Optima prelim study group; POUT study group. Conveying equipoise during recruitment for clinical trials: qualitative synthesis of clinicians' practices across six randomised controlled trials. *PLoS Med* 2016; **13**: e1002147.
8 Arnold DM, Webert KE, Adhikari NK, Cook DJ. Pilot trials in transfusion medicine. *Transfusion* 2009; **49**: 1293–1295.
9 Eldridge S, Bond C, Campbell M, Hopewell S, Thabane L, Lancaster G *et al.* Defining feasibility and pilot studies in preparation for randomised controlled trials: using consensus methods and validation to develop a conceptual framework. *Trials* 2015; **16**(Suppl 2): 087.
10 Thabane L, Hopewell S, Lancaster GA, Bond CM, Coleman CL, Campbell MJ *et al.* Methods and processes for development of a CONSORT extension for reporting pilot randomized controlled trials. *Pilot Feasibility Stud* 2016; **2**: 25.
11 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 Methodol* 2010; **10**: 67.
12 Kistin C, Silverstein M. Pilot studies: a critical but potentially misused component of interventional research. *JAMA* 2015; **314**: 1561–1562.
13 Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. *J Eval Clin Pract* 2004; **10**: 307–312.
14 Shanyinde M, Pickering RM, Weatherall M. Questions asked and answered in pilot and feasibility randomized controlled trials. *BMC Med Res Methodol* 2011; **11**: 117.
15 Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP *et al.* A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol* 2010; **10**: 1.
16 Tickle-Degnen L. Nuts and bolts of conducting feasibility studies. *Am J Occup Ther* 2013; **67**: 171–176.
17 Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L *et al.*; PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ* 2016; **355**: i5239.
18 Eldridge SM, Lancaster GA, Campbell MJ, Thabane L, Hopewell S, Coleman CL *et al.* Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. *PLoS One* 2016; **11**: e0150205.
19 Medical Research Council. *Developing and Evaluating Complex Interventions: New Guidance*. Medical Research Council: London, 2008.
20 The IDEAL Collaboration. http://www.ideal-collaboration.net/ [accessed 28 November 2017].
21 McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC *et al.* No surgical innovation without evaluation: the IDEAL recommendations. *Lancet* 2009; **374**: 1105–1112.
22 Hirst A, Philippou Y, Blazeby J, Campbell B, Campbell M, Feinberg J *et al.* No surgical innovation without evaluation: evolution and further development of the IDEAL framework and recommendations. *Ann Surg* 2019; **269**: 211–220.
23 Loscalzo J. Pilot trials in clinical research: of what value are they? *Circulation* 2009; **119**: 1694–1696.
24 Frees SK, Aning J, Black P, Struss W, Bell R, Chavez-Munoz C *et al.* A prospective randomized pilot study evaluating an ERAS protocol *versus* a standard protocol for patients treated with radical cystectomy and urinary diversion for bladder cancer. *World J Urol* 2018; **36**: 215–220.
25 Natung T, Keditsu A, Shullai W, Goswami PK. Sutureless, glue-less conjunctival autograft *versus* conjunctival autograft with sutures for primary, advanced pterygia: an interventional pilot study. *J Clin Diagn Res* 2017; **11**: NC04–NC07.
26 Perera AP, Howell AM, Sodergren MH, Farne H, Darzi A, Purkayastha S *et al.* A pilot randomised controlled trial evaluating postoperative packing of the perianal abscess. *Langenbecks Arch Surg* 2015; **400**: 267–271.
27 Till SR, Hobbs KA, Moulder JK, Steege JF, Siedhoff MT. McCall culdoplasty during total laparoscopic hysterectomy: a pilot randomized controlled trial. *J Minim Invasive Gynecol* 2018; **25**: 670–678.
28 Wijayasinghe N, Andersen KG, Kehlet H. Analgesic and sensory effects of the pecs local anesthetic block in patients with persistent pain after breast cancer surgery: a pilot study. *Pain Pract* 2017; **17**: 185–191.
29 National Institute for Health Research. *Research for Patient Benefit*. https://www.nihr.ac.uk/research-and-impact/research/nihr-studies/research-for-patient-benefit.htm [accessed 28 November 2017].
30 National Institute for Health Research. *Health Technology Assessment*. https://www.nihr.ac.uk/funding-and-support/funding-for-research-studies/funding-programmes/health-technology-assessment/ [accessed 28 November 2017].
31 Cook JA, Beard DJ, Cook JR, MacLennan GS. The curious case of an internal pilot in a multicentre randomised trial – time for a rethink? *Pilot Feasibility Stud* 2016; **2**: 73.
32 Lancaster GA, Campbell MJ, Eldridge S, Farrin A, Marchant M, Muller S *et al.* Trials in primary care: statistical issues in the design, conduct and evaluation of complex interventions. *Stat Methods Med Res* 2010; **19**: 349–377.
33 Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. *J Psychiatr Res* 2011; **45**: 626–629.
34 Van Teijlingen ER, Rennie AM, Hundley V, Graham W. The importance of conducting and reporting pilot studies: the example of the Scottish Births Survey. *J Adv Nurs* 2001; **34**: 289–295.
35 Eldridge S, Bond C, Campbell M, Lancaster G, Thabane L, Hopewell S. Definition and reporting of pilot and feasibility studies. *Trials* 2013; **14**(Suppl 1): O18.

36 Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L *et al.*; on behalf of the PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Pilot Feasibility Stud* 2016; **2**: 64.
37 McCulloch P, Feinberg J, Philippou Y, Kolias A, Kehoe S, Lancaster G *et al.* Progress in clinical research in surgery and IDEAL. *Lancet* 2018; **392**: 88–94.
38 Avery KN, Williamson PR, Gamble C, O'Connell Francischetto E, Metcalfe C, Davidson P *et al.*; members of the Internal Pilot Trials Workshop supported by the Hubs for Trials Methodology Research. Informing efficient randomised controlled trials: exploration of challenges in developing progression criteria for internal pilot studies. *BMJ Open* 2017; **7**: e013537.
39 Lancaster GA. Pilot and feasibility studies come of age! *Pilot Feasibility Stud* 2015; **1**: 1.
40 Leuck AM, Johnson JR, Hunt MA, Dhody K, Kazempour K, Ferrieri P *et al.* Safety and efficacy of a novel silver-impregnated urinary catheter system for preventing catheter-associated bacteriuria: a pilot randomized clinical trial. *Am J Infect Control* 2015; **43**: 260–265.
41 Romero-Cerecero O, Zamilpa A, Tortoriello J. Pilot study that evaluated the clinical effectiveness and safety of a phytopharmaceutical elaborated with an extract of *Ageratina pichinchensis* in patients with minor recurrent aphthous stomatitis. *J Ethnopharmacol* 2015; **173**: 225–230.
42 Woodhall SC, Nichols T, Alexander S, da Silva FC, Mercer CH, Ison C *et al.* Can we use postal surveys with anonymous testing to monitor chlamydia prevalence in young women in England? Pilot study incorporating randomised controlled trial of recruitment methods. *Sex Transm Infect* 2015; **91**: 412–414.
43 Arnold DM, Burns KE, Adhikari NK, Kho ME, Meade MO, Cook DJ; McMaster Critical Care Interest Group. The design and interpretation of pilot trials in clinical research in critical care. *Crit Care Med* 2009; **37**(Suppl): S69–S74.
44 Kraemer HC, Mintz J, Noda A, Tinklenberg J, Yesavage JA. Caution regarding the use of pilot studies to guide power calculations for study proposals. *Arch Gen Psychiatry* 2006; **63**: 484–489.
45 National Institute for Health Research. *Applied Health Research in Surgery*; 2012. https://www.nihr.ac.uk/funding-and-support/documents/themed-calls/Surgery.pdf [accessed 14 January 2019].
46 Reinforcement of Closure of Stoma Site (ROCSS) Collaborative and the West Midlands Research Collaborative. Feasibility study from a randomized controlled trial of standard closure of a stoma site *vs* biological mesh reinforcement. *Colorectal Dis* 2016; **18**: 889–896.
47 Morgan B, Hejdenberg J, Hinrichs-Krapels S, Armstrong D. Do feasibility studies contribute to, or avoid, waste in research? *PLoS One* 2018; **13**: e0195951.

**Supporting information**

Additional supporting information can be found online in the Supporting Information section at the end of the article.