The cost-effectiveness of Studies Within A Trial (SWATs) for improving recruitment and retention in randomised trials

Athanasios Gkekas, B.Sc. (Hons.), M.Sc.

Doctor of Philosophy

University of York

Department of Health Sciences

December 2023

Abstract

Background: Randomised controlled trials (RCTs) constitute the gold-standard design for evaluating interventions. However, their design, conduct and validity can be threatened by slow recruitment of invited patients and/or attrition of already recruited participants. Given such challenges, the research community is increasingly interested in identifying effective recruitment and retention strategies via the conduct of Studies Within A Trial (SWATs).

Aims: This thesis applies economic techniques to demonstrate the significance of trial recruitment and retention, and to introduce economic methods for improving the evaluation of recruitment and retention strategies via SWATs.

Methods: The thesis employs a wide range of health economic methods, including decision modelling, costing analysis, systematic review and Value of Information (VoI) analysis.

Results: Chapter 2 presents the impact of slow recruitment to the RECOVERY trial generated opportunity costs due to the delayed dissemination of a more cost-effective, available treatment for hospitalised COVID-19 patients. Chapter 3 highlights that participant loss to follow-up from the Occupational Therapist Intervention Study (OTIS) trial generated significant financial costs to the trial team and funder, despite its low attrition rate. Chapter 4 critically appraises the evidence surrounding the cost-effectiveness of recruitment and retention strategies, concluding that no cost-effective strategy exists with high certainty of evidence, due to the limited availability of economic evaluations alongside most SWATs. Finally, considering the limited resources for funding future SWATs, Chapter 5 introduces and applies a Value of Information (VoI) analysis framework to telephone reminders (recruitment strategy) and pens (retention strategy), showing that such a methodology can be feasibly used as a tool, alongside Trial Forge Guidance 2, for prioritising research on recruitment and retention strategies.

Conclusion: Integrating economic evaluations into SWATs and using VoI analyses can strengthen future SWAT-related research, allowing trial methodologists to improve the conduct of their studies.

Η παρούσα διατριβή είναι αφιερωμένη στην οικογένεια και τους φίλους μου

This thesis is dedicated to my family and friends

Αθανάσιος Α. Γκέκας

«Υπάρχει μόνο ένα καλό, η γνώση, και ένα κακό, η άγνοια.»

“There is only one good, knowledge, and one evil, ignorance.”

Socrates

Acknowledgments

First and foremost, I would like to thank my supervisors, David Torgerson, Adwoa Parker and Sarah Ronaldson for their support, expertise, patience, flexibility and insights. From the first day of the PhD, they were by my side, ready to discuss my thoughts and insights and address my concerns. David has transformed to me his practicality and experience in academia, Adwoa her enthusiasm and time management, Sarah her integrity and attention to detail.

I would also like to thank my Thesis Advisory Panel (TAP) members, Catherine Hewitt, Belen Corbacho and Steve Parrott for their guidance. They all showed an interest in the progress of my thesis and contributed to making improvements in my academic skills.

It goes without saying that I am grateful to all my friends. I would like to thank them for the late-night calls, the support they provided during the lockdown, the places we have been to, the social life and their consistent belief in me throughout my studies. I will forever appreciate their presence and impact.

I would also like to thank the University of York for awarding me a prestigious studentship for my PhD, which enabled me to cover my living expenses throughout my studies.

I would like to thank my parents, Vasso Drepaniotou and Asteris Gkekas. They stimulated me to work hard and meticulously, believe in myself, show appreciation to others, travel, learn music and foreign languages, be an open-minded person yet with inherent moral values. I am deeply thankful and grateful to them for all the sacrifices they made. They have been my lifetime teachers and friends. I would also like to thank my other relatives, including Tassos, Harris, Tassoula, Vasilis, Nassos, Karline, Christina, Katerina, Stella and others.

Finally, I wish to memorialise Nikos Tsioltas, my beloved uncle, who died prematurely during the first year of my PhD. He always believed in me and taught me how important positive thinking is; I truly miss him.

Declaration

I declare that this thesis is a presentation of original work, and I am the sole author. This work has not previously been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references.

Chapter 4.2 has been peer-reviewed and published as GKEKAS, A., EVANS, A., PARKER, A., RONALDSON, S. J. & TORGERSON, D. J. 2023. A systematic review of economic evaluations alongside studies within a trial (SWATs) for improving recruitment and retention in randomised controlled trials. *Research Methods in Medicine & Health Sciences,* 4**,** 94-112. The version presented in the thesis is identical to the peer-reviewed publication.

Contents

List of figures

List of tables

Chapter 1: Introduction

1.1. Randomised Controlled Trials

Randomised Controlled Trials (RCTs) are experimental prospective "*comparative studies with an intervention group and a control group; the assignment of the participant to a group is determined by the formal procedure of randomisation*" (Friedman et al., 2015). Randomisation means that the enrolled participants are randomly, and equally likely to be, allocated to the control and intervention groups. In contrast to nonrandomised study designs, such as case-control, cohort, and cross-sectional studies, RCTs ensure that selection and allocation bias when allocating participants to the intervention or to the control groups are minimal. Therefore, any differences in outcome measures between the intervention and the control group(s) can be statistically reliable (Torgerson and Torgerson, 2008). In other words, randomisation ensures that the groups are comparable in the sense that both measurable and unmeasurable, as well as known and unknown, variables can all be balanced so that any detected differences observed in outcomes between the control and the intervention groups are due to the intervention(s) in question (Torgerson and Torgerson, 2008). If the sample size is sufficient, the statistical analysis of effect differences between the intervention and the control groups in RCTs should be straightforward, since the outcome differences between the two (or more) groups follow a parametric probability distribution, which can be used for estimating statistical measures such as the confidence intervals of risk differences. Therefore, if an RCT is not subject to other issues and biases- such as patient and clinician subversion, technical bias, attrition bias, recruitment bias, dilution bias, effect dilution due to a delay between intervention allocation and intervention, resentful demoralisation, exclusion bias, and unplanned subgroup analysis- it is highly likely that such an RCT will; 1) achieve high internal validity 2) reliably capture the effectiveness of a given intervention (Drummond et al., 2015), and; 3) identify any differences in the outcomes of the intervention and control groups due to the intervention itself, and not due to any other known or unknown confounder (Torgerson and Torgerson, 2008).

However, even if designed perfectly, RCTs are subject to some significant limitations. First, under a clinical research question explored via an RCT, it is likely that rare or longer-term adverse events not be captured (Saldanha et al., 2022). Second, there may be ethical concerns, especially in placebo-controlled RCTs where participants under a severe clinical condition may be denied receiving a potentially effective treatment (Saldanha et al., 2022, Resnik, 2008). Third, due to financial or practical reasons the exploration of longer-term outcomes with

regards to the effectiveness of an intervention or treatment under evaluation may be limited or unavailable. This can be a challenge particularly when longer-term outcomes are critical to assess the clinical or cost effectiveness of an intervention or treatment for a given clinical condition (Saldanha et al., 2022). Fourth, it may be challenging to recruit patients when the disease under exploration is rare in terms of prevalence and/or incidence. Finally, whereas RCTs can achieve sufficient internal validity, the requirements for doing so may impede on producing findings that could be generalisable for different patient populations in different sociographic contexts (Saldanha et al., 2022). Hence, alternative research methodologies such as case-control or cohort studies, which could explore longer-term outcomes and are rarely restricted from ethical considerations, could be explored in instances where a research question presents challenges to the conduct or the interpretation of findings from a relevant RCT.

Despite the aforementioned limitations, RCTs are commonly used in healthcare settings for assessing the effectiveness of a new or existing intervention, provided that the experimentation is feasible for the research question(s) being addressed and that there are no significant ethical, political, and legal obstacles (Black, 1996), because of their huge potential for achieving internal validity. Thus, major research funders such as the Medical Research Council (MRC) regard the RCT to be the "*most scientifically rigorous, unbiased way of comparing alternative healthcare interventions*" (Warlow, 2003). Depending upon the sample size, the existence of potential confounders that may affect an outcome in question, the likelihood of obtaining balanced groups, and the available clinical settings, RCTs can be designed in different forms of randomisation, including simple randomisation, matched randomisation, blocked randomisation, and pairwise randomisation (Torgerson and Torgerson, 2008). Within the healthcare research context, and when feasible, it is often recommended that RCTs are double-blind, meaning that neither the participants nor the trial team/data collectors are aware of the treatment a participant receives, for the purpose of minimising several biases such as ascertainment bias (Friedman et al., 2015).

1.2. Recruitment Challenges for RCTs

Despite the distinguishing design features of RCTs, recruitment of participants remains one of the major challenges in achieving strong internal and external validity in RCTs, as recruitment is usually poor. Systematic reviews that attempted to capture the recruitment rates in RCTs have concluded that, despite some recent improvements, approximately 50% of the trials fail to meet their original recruitment targets (Treweek et al., 2018b, Fletcher et al., 2012). In

addition, there seems to be a challenge in the speediness of recruiting participants in RCTs, with a review of 388 trials, which were funded by the National Institute for Health & Care Research (NIHR) from 1997 to 2020, having concluded that the recruitment period was extended in 33% of trials and the recruitment target was revised downward in 20% of trials (Jacques et al., 2022). Moreover, only 74% of trials managed to reach 80% of their original recruitment target (Jacques et al., 2022).

Several studies have attempted to identify the barriers to successful recruitment in RCTs. For instance, in a mental health trial (RefraMED), which evaluated the clinical and the cost-effectiveness of a radically open dialectical behaviour therapy for refractory depression (Lynch et al., 2020), the main barriers for trial rejecters were their pre-existing perceptions against participating in a trial, their disagreements with the way the clinical practitioners viewed their illness and clinical management, their beliefs that the suggested trial treatment would not help in treating their illness, and their concerns that the personal costs of participating in a trial exceeded the personal benefits (Parker et al., 2016). From the clinical staff's perspective, a qualitative study concluded that regional and national competition for participants, clinical workload, lack of clarity over regulation by the ethical committee's regulation and intrapersonal relationships among trialists all impinged on achieving successful recruitment to RCTs (Adams et al., 2015). In addition, there is evidence that barriers to recruitment in RCTs can be more sensitive to participants suffering from mental health disorders, e.g. depression (Parker et al., 2015).

To understand the implications of poor recruitment for RCTs, the principles of prior RCT sample size estimation should be introduced. In general, there are several multidimensional aspects that need to be considered when estimating a desirable sample size for an RCT, including:

1) The study design, which is usually a parallel group design in RCTs
2) The type of hypothesis testing (i.e. one-sided vs two-sided)
3) The primary endpoints (i.e. continuous vs discrete outcomes)
4) The appraisal of previous evidence on the expected response of the intervention compared to the control treatment that accumulates the effectiveness of the intervention treatment, the effectiveness of the control treatment and the standard deviation
5) The magnitude of the clinical significance, which specifies the extent to which the intervention in question should be effective for it to be approved

6) The significance level, which is inversely related to Type I error; Type I error occurs when an ineffective intervention is wrongly estimated to be effective
7) The power, which is inversely related to Type II error; Type II error occurs when it is not statistically shown that an effective intervention is effective
8) The expected retention rate
9) The potential for unequal treatment allocation (Sakpal, 2010)

In other words, sample size calculations are useful for determining the number of participants needed to observe whether the null hypothesis of no statistical difference in clinical outcomes between the intervention and the control treatment is rejected for a given significance level and clinical significance, whilst minimising Type II errors. Therefore, when the expected desirable sample size for a given RCT is not achieved, criterion 7 is directly affected, as the power of the trial is correlated with the recruited sample size (Sakpal, 2010). This implies that low recruitment leads to underpowered trials that are subject to an increased likelihood of mistakenly rejecting the clinical effectiveness of a given intervention. The implications of an increased probability of Type II errors in under-recruited RCTs consist of research waste, rejecting effective healthcare interventions, taking longer for meta-analyses to obtain the effectiveness of rejected yet effective interventions, ethical issues through exposing participants to uncertainty during and after the trial and, most importantly, extension of the length of a given trial which puts a huge strain on the existing budget of such trial (Treweek et al., 2018b). Alternatively, if trialists try to extend the recruitment period to achieve their recruitment target, the trial may need to be extended, thus increasing the protocol-driven costs, or delaying the introduction of an effective intervention/treatment into practice, or terminating the trial prematurely (McDonald et al., 2006). To highlight the magnitude of costs arising from poor recruitment an American study has estimated the aggregate uncompensated costs of poor recruitment in 837 clinical studies to be approximately $1 million after estimating a wide range of start-up, maintenance, and close-out costs (Kitterman et al., 2011). The clinical study preparation costs had the largest share of the total uncompensated costs, followed by the necessary study modification costs because of poor recruitment. More recently, a study modelled the impact on human lives lost due to poor recruitment in the COVID-19 RECOVERY trial, which showed that over 2,600 lives were lost due to poor recruitment (Knowlson and Torgerson, 2020). Evidently, poor recruitment in RCTs can lead to huge direct and indirect economic costs reflecting the high opportunity costs (i.e. in terms of not allocating resources forgone towards other research activities that could have a stronger impact on clinical

practice) and the poor economic and clinical benefits of running an under-recruited trial (Kitterman et al., 2011). The large magnitude of such costs could also be a primary concern for research councils which may be incentivised to invest in faster but less robust research designs instead of an RCT (Watson and Torgerson, 2006).

The thesis aims to take a broader perspective from that of trial teams in order to highlight the costs of slow and hence poor patient recruitment into randomised trials, related to national healthcare systems through a case study. This thesis examines the RECOVERY trial, and the aforementioned clinical findings with regards to the epidemiological costs of slow recruitment to the RECOVERY trial, in terms of excess mortality. In *Chapter 2* I develop a decision model to evaluate the cost-effectiveness of dexamethasone as a treatment for hospitalised COVID-19 patients, whose results are used to estimate the cost-effectiveness of increasing patient recruitment rate from 15% to 50% in the RECOVERY trial by hiring or redeploying two research nurses into each hospital participating in the study.

1.3. Challenges of attrition from RCTs

Another main challenge with RCTs, one which can cancel out the main statistical advantages of RCTs over non-experimental study designs, is attrition, meaning that already recruited participants fail to complete their participation in a trial. Despite the fact that recruitment and retention are both related to participation in RCTs, recruitment of participants occurs before the start of a trial, whereas retention of participants occurs after the start of the trial and, therefore, arguably, is a more potent threat to a successful RCT as it not only diminishes the power of the trial but also can introduce selection bias. This means that even if recruitment of participants into a trial has been successful, it is not guaranteed that the attrition rate will be low.

The implications of poor retention in RCTs can affect their internal and external validity, as well as the statistical analysis of RCTs. If the attrition rate is non-random (i.e. the attrition is related to the outcomes of the trial and/or to confounders), the recruited participants remaining in the trial may systematically differ from the recruited participants deciding to leave the trial prematurely, in terms of unknown covariates, implying that the retained participants may not necessarily be representative of the general population. Whereas randomisation can account for such unknown confounders, attrition might imply selection bias and therefore larger uncertainty in the statistical analysis of the intermediate and final outcomes (Torgerson and

Torgerson, 2008). For example, if a new drug has adverse effects that lead to the participants to miss clinical appointments, the attrition rate between the intervention and the control group will be different and the new drug could be biasedly favoured (Friedman et al., 2015). Therefore, selection bias induced by attrition can lead to a change in the direction of effect and hence to a wrong conclusion that an intervention of interest is beneficial (harmful), whereas in fact it is harmful (beneficial). Then, even if the attrition rates were similar between the control and the intervention groups, selection bias would never be eliminated (Torgerson and Torgerson, 2008). To disentangle attrition, there exist statistical methods which, however, are based on assumptions that are implausible with RCTs of healthcare interventions (Friedman et al., 2015). For example, regression models can act as imputation methods for handling missing data; however, their assumption is that the missing data are random, whereas it is usually the case that attrition is non-random in clinical trials. Another method is endpoint analysis, but the assumption of the missing future observations being constant with past observations is unfeasible within the healthcare context. The financial costs of poor attrition can be also significant, with the time costs of researchers dealing with follow- up being dominant (Peterson et al., 2012).

Many RCTs may struggle with achieving strong retention, i.e. the ability to retain participants in an RCT after it has started. A systematic review of 151 trials associated with the UK's NIHR HTA Programme has found the median retention rates to be 89%, with the lower interquartile range being 80% and the upper interquartile range being 97% (Jacques et al., 2022). In other words, 25% of NIHR-funded trials were reported to have attrition rates exceeding 20% (Jacques et al., 2022).

To identify the barriers to successful retention in RCTs, a qualitative study approached participants who dropped out from the EXercise for Type 1 Diabetes (EXTOD) trial, with time costs of participation, travel costs, long duration of trial, frequency of visits, change of residence and personal preferences acting as the main factors in participants prematurely leaving the trial (Henshall et al., 2018). Similar findings have been found in a qualitative study of participants experiencing depression and cancer simultaneously, with financial costs and familial relationships further acting as barriers to successful retention (Wells et al., 2015). In addition, retention rates might vary according to the participants' ethnic background (Villarruel et al., 2006). Finally, some participants may simply and for non-specific reasons disagree to

continue undertaking a treatment/intervention having previously consented to the trial (Friedman et al., 2015).

The thesis aims, through a case study, to adopt the economic perspective of trial teams in order to highlight the financial costs of participant attrition from randomised trials. By looking into the Occupational Therapist Intervention Study (OTIS), *Chapter 3* collects the direct economic costs of attrition in terms of recruitment and retention costs for recruited participants that were lost to follow-up. Such costs are protocol-driven and include the administrative, print and shipping costs of trial materials sent to participants who were randomised but eventually were lost to follow-up. In addition, this chapter considers the timepoint at which attrition occurred for each participant.

1.4. Evidence for preventing attrition and improving recruitment

Given the threats to internal validity and statistical reliability that poor recruitment and attrition pose to RCTs, there has been a growing literature on identifying strategies that could improve recruitment and/or retention in RCTs.

A review of 14 RCTs has concluded that telephone reminders, questionnaires rather than interviews, financial incentives, adopting an open instead of a placebo trial, and adjusting the trial materials favourably towards participants' cultural characteristics can improve recruitment rates in RCTs; however, few trials have been included in the review and therefore these results are uncertain (Watson and Torgerson, 2006). Another review of 13 trials failed to find strategies for improving recruitment in mental health RCTs, but it did identify strategies for improving retention, including financial incentives, follow-up, and shorter questionnaires (Liu et al., 2018). The most recent Cochrane recruitment review has found with high GRADE[1] certainty that an open trial design and use of telephone reminders can boost recruitment in randomised trials; however, telephone reminders were successful only in trials with very low recruitment rates, i.e. it is unknown whether reminders would be effective in trials with moderate recruitment rates, and many included trials were subject to a high risk of

1 A high-GRADE evidence means that "the authors have a lot of confidence that the true effect is similar to the estimated effect". The GRADE ratings are applicable to systematic reviews and can be used to make clinical practice recommendations. The five main factors that could affect the four GRADE certainty ratings (i.e. very low, low, moderate, high) are the extents of; 1) risk of bias; 2) imprecision; 3) inconsistency; 4) indirectness, and; 5) publication bias, that are found in the reviewed studies (Sieminiuk and Guyatt).

bias, implying that more needs to be done to identify recruitment strategies with higher certainty (Treweek et al., 2018b).

A review of 38 retention trials found that monetary and high-valued incentives boosted retention through postal questionnaires, and that monetary incentives boosted retention through electronic questionnaires (Brueton et al., 2014). Nevertheless, the effectiveness of other retention strategies remains unclear, implying that more evidence is needed before coming to conclusions on effective retention strategies (Brueton et al., 2014). The most recent Cochrane retention review has not found any intervention to be effective in improving retention with high GRADE certainty (Gillies et al., 2021). The only intervention that may reduce attrition is adding a diary to the follow-up process (Gillies et al., 2021). The latter review questioned the use of monetary incentives as means of improving retention, as two of the three included studies were assessed with high risk of bias.

Both Cochrane recruitment and retention reviews highlighted that the evaluations of additional novel recruitment and retention strategies, rather than the replication of trials with the existing recruitment and retention strategies in different settings, have implied poor progress on identifying effective recruitment and retention strategies. Further factors behind such uncertainty include the high or uncertain Cochrane risks of bias of many studies which explored the effectiveness of recruitment and retention interventions and, if more than one study explored the effectiveness of a recruitment or retention strategy, the low overall quality of the evidence (GRADE) for most strategies. Therefore, there is limited and uncertain evidence on strategies that could improve recruitment and/or retention in RCTs, or, more broadly, on how researchers could make well-informed decisions on administering and running RCTs effectively and with more certainty (Treweek et al., 2015).

1.5. Studies within a Trial (SWATs)

Due to the recognition of the lack of evidence of methods for improving the methodology and conduct of randomised trials, there is growing support in the research community for the use of ***Studies Within A Trial (SWATs)*** (Bower et al., 2014, Treweek et al., 2018a). The first initiative on SWATs was founded through a partnership involving the All-Ireland Hub for Trials Methodology Research, along with the MRC Network of Hubs in the UK, and various other collaborators (Clarke et al., 2015) . As a result, *Trial Forge*, a collaborative group led by Shaun Treweek of the University of Aberdeen, has been formed to improve and disseminate

rigorous evidence for improving the conduct of RCTs (Trial Forge, 2023). Trial Forge Guidance 1 outlines a comprehensive guidance on planning a SWAT, with respect to costs, randomisation, ethics, statistical analysis, implementation, and publication (Treweek et al., 2018a). According to this guidance (Treweek et al., 2018a):

A Study Within a Trial (SWAT) is a "self-contained research study that has been embedded within a host trial with the aim of evaluating or exploring alternative ways of delivering or organising a particular trial process."

The primary objective of SWATs is to improve trial methodology and efficiency (Treweek et al., 2018a). Thus, SWATs could also work as a useful study design for identifying strategies to improve recruitment and retention in RCTs. The key features of SWATs are that they should have their distinct trial protocol, can have the same or different randomisation process from the host trial, depending on the context, can be embedded within one or more trials, simultaneously or sequentially, may help inform decisions about the original host trial(s) and future trials, and should not affect the protocol, randomisation and statistical analysis of the host trial(s) (Treweek et al., 2018a).

Additionally, Trial Forge Guidance 2 outlines the five criteria which could be applied for deciding whether a further evaluation of a recruitment/retention intervention would be needed in a future SWAT; these are GRADE certainty, cumulated evidence, PICOT , balance of benefit and disadvantage to participants, and balance of benefit and disadvantage to the host trial (Treweek et al., 2020). More details about Trial Forge Guidance 2 can be found in *Table 1.1*.

Table 1.1: Trial Forge Guidance 2, adapted from (Treweek et al., 2020: p.3):

"The five proposed criteria for deciding whether the intervention needs another evaluation in a SWAT. The more criteria that are met, the more likely we are to conclude that further evaluation in a SWAT is appropriate.
1. *GRADE*: the GRADE certainty in the evidence for all key outcomes is lower than 'high'
2. *Cumulated evidence*: the cumulative meta-analysis shows that the effect estimate for each outcome essential to make an informed decision has not converged
3. *Context*: the range of host trial contexts evaluated to date does not translate easily to the context of the proposed SWAT.[d] For the proposed SWAT consider PICOT:
• P – is the population in the host trial so different from those already included that the current evidence does not provide sufficient certainty?
• I – are the health interventions in the host trial so different from those already included that the current evidence does not provide sufficient certainty?
• C – is the comparator in the host trial so different from those already included that the current evidence does not provide sufficient certainty?
• O – is the SWAT outcome(s) so different to those used in the existing evaluations that that the current evidence does not provide sufficient certainty?
• T – in the time since the existing evaluations were done, have regulatory, technological or societal changes made those evaluations less relevant?
4. *Balance – participants*: the balance of benefit and disadvantage to participants in the host trial and/or the SWAT is not clear
5. *Balance – host trial*: the balance of benefit and disadvantage to the new host trial is not clear"

There is a formal SWAT repository, developed by Professor Mike Clarke at Queen's University Belfast, which attempts to capture the existing SWATs for the purpose of avoiding effort duplication and for motivating researchers on how they could adopt SWATs in their host trials (Medical Research Council, 2020). Moreover, there are registered Trial Forge Centres across the UK, Switzerland, Australia, Canada and Ireland that have been conducting SWATs, including the York Trials Unit (YTU). Examples of recruitment interventions that have been evaluated in SWATs include leaflet advertisement, pen printed with the trial logo, financial incentives, social media advertisement, site visits by clinical team and video clip (Medical Research Council, 2020). Examples of retention interventions that have been evaluated in SWATs include financial incentives, sending pre-notification cards before outcome measurement, birthday cards with nudge (i.e. a behavioural intervention aiming to change people's decisions without altering their personal preferences and choices), Christmas cards, telephone follow-up and differential timing of offering incentives to participants (Medical Research Council, 2020). In general, the aforementioned recruitment and retention

interventions have been designed to act as facilitators to the aforementioned barriers to recruitment and retention, by affecting the trust and relationships between invited/recruited participants and trial units, and by lowering the tangible and intangible participant costs.

Currently, there is a funding stream for SWATs across the UK for the purpose of making embedded recruitment and retention trials a routine task in more RCTs. The NIHR has highlighted the ongoing uncertainty that exists in the preparation and conduct of a clinical trial, and therefore has introduced a new funding scheme of up to £30,000 for a newly registered SWAT. Such initiative signals the growing interest by funders and the research community more generally in finding ways that could significantly improve trial efficiency (National Institute for Health & Care Research, 2023). Also, the PROMETHEUS programme was a SWAT funding stream, funded by the UK Medical Research Council (MRC) and Clinical Trials Unit (CTU) infrastructure from the NIHR. It reimbursed main trials with up to £5,000 for undertaking an embedded trial of a recruitment or retention intervention (Clark et al., 2022). At the international level, the Health Research Board (HRB) in Ireland and the government-funded Accelerating Clinical Trials (ACT) Consortium in Canada have started funding SWATs (HRB-Trials Methodology Research Network, 2023, Accelerating Clinical Trials Consortium, 2023).

1.6. The motivation for doing economic evaluations of SWATs to improve recruitment and reduce attrition in RCTs

1.6.1. Economic evaluation of recruitment and retention strategies alongside SWATs

Given the direct and indirect costs of poor recruitment and retention rates, as well as the urgent need for improving trial efficiency, an economic evaluation of recruitment and retention interventions alongside SWATs is crucial. In this way, when evaluating a new recruitment and/or retention intervention, trialists, healthcare providers and research councils could make better-informed decisions on how to conduct trials more efficiently, whilst considering the direct and indirect incremental effects and costs of such an intervention to the researchers (or clinical units running the trial), the trial participants and the general population. More broadly, an *"economic evaluation offers an organised consideration of the range of possible alternative courses of action and the evidence of the possible effects of each. This is more likely to lead to better decisions that improve overall social value. It also requires that the scientific judgements needed to interpret evidence are made explicitly so they can be scrutinised and the impact of*

alternative but plausible views examined. More importantly, it can provide a clear distinction between these questions of fact and the unavoidable questions of value. Indeed, the main contribution of economic evaluation may not be in changing the decisions that are made but how they are made as it offers the opportunity for proper accountability for choices made on behalf of others" (Drummond et al., 2015).

The questions surrounding economic evaluation alongside SWATs should be similar to those of doing an economic evaluation alongside their host trials, as SWATs are also RCTs that are directly associated with their host trials. Challenges for a proper economic evaluation could be: the identification, measurement and valuation of costs and outcomes of recruitment/retention strategies; the proper attribution of direct and indirect health benefits, health costs, non-health benefits and non-health costs to recruitment/retention interventions; the determination of the relevant *and feasible* perspectives for economic evaluation of recruitment/retention strategies; and the decision of whether subgroup analysis should take place (Edwards et al., 2013). To date, there is no economic framework for SWATs that has been developed and shared with the research community.

In *Chapter 4,* the existing evidence around economic evaluations alongside SWATs for improving recruitment and retention in randomised trials is gathered through a comprehensive systematic review, which also makes recommendations about economic evaluations alongside future SWATs. This review has also been published as a peer-reviewed research article at Research Methods in Medicine and Health Sciences (Gkekas et al., 2023).

1.6.2. Value of Information (VoI) analysis related to recruitment and retention interventions

Another area of economic evaluation which could improve the embedded trial research on recruitment and retention to RCTs is the Value of Information (VoI) analysis that originates from statistical decision theory. By definition, a *"VoI is a means of valuing the expected gain from reducing uncertainty through some form of data collection exercise (e.g., a trial or epidemiological study). As such, it is a tool which can be used to assess the cost effectiveness of alternative research projects*" (Wilson, 2015). In trial research, a key measurement in VoI analysis is the expected value of sample information (EVSI), which equals *"the expected maximum expected net benefit with the new information yielded from a study of sample size n*

per arm less the maximum expected net benefit with current information, multiplied by the beneficial population less those enrolled in the study" (Wilson, 2015).

Within the context of SWATs, VoI analyses related to already evaluated recruitment and retention interventions are highly encouraged. The Cochrane recruitment and retention reviews have concluded high uncertainty in the effectiveness of several recruitment and retention interventions. As a result, trialists should cautiously determine whether further trials are needed for such interventions. Whereas the Trial Forge Guidance 2 is a comprehensive tool to enable researchers to determine whether undertaking additional SWATs of a given intervention is beneficial, the criteria it uses are mainly implicit (Treweek et al., 2020). Also, this guidance does not clarify how future SWATs of a given intervention should be undertaken, e.g. there is no framework for determining the financial impact of funding an additional SWAT of a recruitment or retention strategy. Therefore, whereas the five criteria specified in the Trial Forge Guidance 2 should be considered, an additional VoI analysis related to a given intervention could augment the quality of decision making by estimating the value of additional research. For instance, the Cochrane retention review (Gillies et al., 2021) concluded that the addition of a pen compared to no pen may increase retention by up to 5%; however, the certainty of the evidence (GRADE) is low. As a result, the review argued that further SWATs need to be undertaken to reduce the uncertainty in the effectiveness of the pen intervention. Similarly, Trial Forge Guidance 2 itself (Treweek et al., 2020) recommends that further SWATs on telephone reminders as a recruitment strategy be undertaken to enhance the certainty of evidence, even if such reminders have been estimated to be effective with high GRADE certainty of evidence in the Cochrane review of recruitment strategies (Treweek et al., 2018b)

In the context of SWATs, *Chapter 5* develops a VoI analysis framework which follows closely the methodology developed by (Claxton et al., 2015a) and (McKenna et al., 2016), to determine if it would be cost-effective to undertake and finance further SWATs of telephone reminders and pens from the economic perspective of a SWAT commissioner such as the NIHR, which currently offers up to £30,000 in financial support towards SWATs (National Institute for Health & Care Research, 2023). This study could work as a guidance for decision makers, alongside Trial Forge Guidance 2, for optimising future SWAT-related research.

1.7. Research questions and structure of the thesis

1.7.1. Research questions

As a result, given:

- randomised trials are the gold-standard research design for evaluating interventions
- poor patient recruitment and participant attrition from randomised trials threaten trial efficiency and can generate economic costs
- SWATs are widely used for evaluating the effectiveness of recruitment and retention strategies
- trial teams can benefit from economic evaluations alongside SWATs in terms of implementing effective recruitment and retention strategies in randomised trials while considering the costs of doing so
- the absence of evidence surrounding the cost-effectiveness of recruitment and retention strategies
- the main source of funding for SWATs comes from the NIHR (up to £30,000 per trial), and previously from the PROMETHEUS programme (up to £5,000 per trial)
- future SWAT-related research should be optimised given the existing budget constraints

This thesis aims to highlight the economic implications of slow patient recruitment and participant attrition from randomised trials from the perspective of trial teams, national healthcare systems and funders, by applying economic techniques such as decision modelling and cost analysis. The thesis also aims to critically appraise the existing evidence on the cost-effectiveness of existing recruitment and retention strategies via a novel systematic review of economic evaluations alongside SWATs. Finally, it aims to make recommendations for the direction of future SWAT-related research; it will show how Value of Information (VoI) analysis methods could demonstrate whether it would be cost-effective for a SWAT commissioner to fund additional SWATs of a given recruitment or retention intervention, given the existing evidence and budget constraints.

The research questions to be addressed by this thesis are the following:

1. *What is the economic impact of poor patient recruitment into RCTs?*
2. *What is the economic impact of participant attrition from RCTs?*
3. *What is the cost-effectiveness of existing recruitment and retention strategies?*

4. *How could VoI analyses related to retention or recruitment interventions inform decision makers on whether additional SWATs are needed for improving the evidence on the (cost) effectiveness of such interventions?*

1.7.2. Structure of the thesis

Chapter 2 highlights the opportunity costs of slow recruitment to RCTs from the National Health Service (NHS) perspective, using as a case study the RECOVERY trial which evaluated the effectiveness of dexamethasone in reducing the mortality incidence for UK patients admitted to hospital with severe symptoms of COVID-19. While dexamethasone proved to be efficacious in reducing the death rates in hospitalised patients (Horby et al., 2021), the recruitment rate was low (15%), meaning that over 2600 more lives could have been saved had the recruitment rate been 50% (Knowlson and Torgerson, 2020). This chapter goes on to estimate the cost-effectiveness of faster recruitment to the RECOVERY trial from the NHS perspective had a recruitment rate of 50% been achieved by recruiting or redeploying two research nurses to each hospital participating in the trial. To do so, it applies decision-model methods to evaluate the cost-effectiveness of dexamethasone, then of updated clinical practice following the dissemination of the results, and subsequently of faster recruitment to the RECOVERY trial.

Chapter 3, in a similar fashion to *Chapter 2*, demonstrates the financial impact of participant attrition from randomised trials, using as a case study the OTIS trial. Despite the low attrition rate reported in this trial (9.8%), there are still direct costs related to the recruitment and retention of participants lost to follow-up, which were incurred by the trial team. To identify these costs, the chapter applies a costing analysis which considers the protocol-driven costs of trial materials sent to participants lost to follow-up and the timepoint at which they withdrew from the study or died. In addition, the analysis presented could be used as a guidance for research teams to estimate the financial costs resulting from participant attrition in their randomised trials.

Given the threats of poor recruitment and attrition for trial teams, funders and national healthcare systems, *Chapter 4* discusses the existing evidence on SWATs by summarising the evidence from two Cochrane systematic reviews of recruitment and retention strategies (Treweek et al., 2018b; Gillies et al., 2021). Then, a systematic review of economic evaluations alongside SWATs for improving recruitment and retention in RCTs is undertaken for the first

time, in order to critically appraise the cost-effectiveness of recruitment and retention strategies. The review identifies the frequency at which, as well as the mechanisms under which, economic evaluations or costs of recruitment and retention strategies are reported in SWATs. Furthermore, it makes recommendations about economic evaluations alongside future SWATs.

Chapter 5 undertakes Value of Information (VoI) analyses related to a pre-existing recruitment strategy, i.e. telephone reminders to non-responders, and a retention strategy, i.e. addition of a pen in follow-up questionnaires. The meta-analysis data from two SWATs of telephone reminders from the Cochrane recruitment review (Treweek et al., 2018b) and five SWATs of pen retention studies included in the Cochrane retention review (Gillies et al., 2021) were considered. The purpose of this chapter is to show how a VoI analysis can be undertaken in SWATs, and to demonstrate why this economic and statistical approach should be considered when deciding on whether an additional SWAT of a given recruitment or retention intervention is needed, in addition to the Trial Forge Guidance 2 (Treweek et al., 2020). The methodology presented in this chapter could work as a guidance for decision makers on optimising future SWAT-related research.

Chapter 2: The cost-effectiveness of improving patient recruitment to RCTs: a case-study of dexamethasone from the RECOVERY trial

2.1. Abstract

Background: The RECOVERY trial assessed the effectiveness of medicinal treatments for preventing severe outcomes from COVID-19 disease in hospitalised patients from 176 NHS hospitals. Of the 9,355 eligible recruited COVID-19 patients, 6,425 participated in the comparison of dexamethasone plus usual care versus usual care only. Mortality benefits of dexamethasone were observed for COVID-19 patients who either stayed in acute hospital wards or were treated in intensive care units (ICUs) by receiving invasive mechanical ventilation and/or non-invasive ventilation. Despite the urgency for results, the average recruitment rate across the participating hospitals was only 15%.

Aim: The aim of this chapter is to estimate the cost-effectiveness of improving recruitment to the RECOVERY trial from 15% to 50%, related to the evaluation of dexamethasone as a COVID-19 treatment, by employing as a recruitment strategy two research nurses to each hospital affiliated with the RECOVERY trial.

Methods: A decision tree model of 28 clinical pathways is developed to estimate the cost-effectiveness of dexamethasone plus usual care (i.e. Dexamethasone) against usual care only (i.e. No Dexamethasone). Probability, utility, and cost inputs are estimated for each clinical pathway and treatment group. Given the dexamethasone cost-effectiveness findings, a cost-utility analysis of clinical practice post-RECOVERY trial versus previous clinical practice is undertaken; the latter analysis is aggregated at the population level and includes the cost of recruiting or redeploying two research nurses at each hospital, to estimate the incremental cost-effectiveness ratio (ICER) of faster recruitment to the RECOVERY trial and the associated incremental net benefit.

Results: Following probabilistic sensitivity analysis, faster recruitment to the RECOVERY trial could have generated an incremental net benefit of £13,955,476.42 (95% CI: £12,457,048.54, £15,453,904.30) thus highlighting the magnitude of the foregone population health benefits due to not having updated clinical practice earlier with faster recruitment. If recruiting two research nurses to each involved hospital increased recruitment rates from 15% to 50%, only £10,641 would need to be invested by a decision maker in order to generate an additional quality-adjusted life year (QALY).

Conclusion: The effect of poor recruitment to randomised controlled trials (RCTs) can have severe implications for national healthcare systems.

2.2. Introduction

When the COVID-19 pandemic struck the United Kingdom one of the initial responses was to rapidly implement the RECOVERY trial to evaluate the effectiveness of potential medicinal treatments on preventing severe outcomes from COVID-19 disease in hospitalised patients (Horby et al., 2021). Patients who were admitted to one of the 176 participating hospitals and had suspected or Polymerase Chain Reaction (PCR) -test-confirmed SARS-Cov-2 infection received one of the five treatments originally available, i.e. dexamethasone, hydroxychloroquine, lopinavir–ritonavir, azithromycin, or usual care only. RECOVERY was designed as a controlled, open-label, adaptive, multi-armed platform trial. The recruitment began on March 19th, 2020, and by June 8th, 2020, 11,303 COVID-19 patients consented and underwent randomisation, of whom 9,355 (83%) could receive dexamethasone, i.e. because the drug was available at a hospital and these patients did not have any contraindications to dexamethasone. Of the 9,355 eligible COVID-19 patients, 6,425 participated in the comparison of dexamethasone plus usual care versus usual care only for assessing their effectiveness in reducing 28-day mortality (Horby et al., 2021). The results of the dexamethasone arm from the RECOVERY trial were announced on June 16th, 2020 (Wise and Coombes, 2020).

Most benefits of dexamethasone were observed for COVID-19 patients who received invasive mechanical ventilation and non-invasive ventilation (Horby et al., 2021). The incidence of COVID-19-related death in the dexamethasone group was significantly lower compared to the usual care group, with the risk ratio being 0.64 (95% confidence interval (CI): 0.51, 0.81), for patients who received invasive ventilation, and 0.82 (95% CI: 0.72, 0.94), for patients who received non-invasive ventilation as their most intensive treatment for COVID-19.

Whereas the dexamethasone results of the trial were produced rapidly, the average recruitment rate across the participating hospitals was only 15%, with recruitment ranging from 3% to 80% per hospital (Wise and Coombes, 2020). Interestingly, despite the calls from the National Institute for Health & Care Research (NIHR) and Professor Chris Whitty, the Chief Medical Officer for England, to increase recruitment in the RECOVERY trial as "*useful evidence could be available within weeks*" (National Institute for Health & Care Research, 2020), recruitment rates remained low overall. Consequently, the RECOVERY trial, despite its flexible design, did not manage to escape from the potential consequences of slow (and hence poor) recruitment to randomised controlled trials (RCTs) (Treweek et al., 2018b), such as policymakers indirectly

rejecting an effective healthcare intervention (e.g. dexamethasone) through the delay of the dissemination of the RECOVERY trial's results, and the extension of the length of the RECOVERY trial, which may have put a strain on its allocated budget.

A study has estimated the potential clinical benefits of improving recruitment to the RECOVERY trial in lowering COVID-19-related mortality in the UK (Knowlson and Torgerson, 2020). Considering the original recruitment rate of 15% and the number of COVID-19 patients randomised in the RECOVERY trial (n=11,303), 75,353 COVID-19 hospitalised patients were identified between March 19th, 2020 (i.e. when the first COVID-19 patient was recruited to the trial) and June 8th, 2020 (i.e. when all COVID-19 patients associated with the dexamethasone trial groups of the RECOVERY trial had been randomised). Assuming a 50% recruitment rate instead, only 22,606 hospitalised COVID-19 patients would need to have been identified for recruitment to end, a target that could have been met by April 1st, 2020 (Knowlson and Torgerson, 2020). Therefore, as the RECOVERY results were originally disseminated eight days after the end of patient recruitment, the trial's results could have been disseminated on April 9th rather than June 16th. Consequently, since 77,310 COVID-19 patients were admitted between April 9th and July 15th (i.e. one month after the dissemination of the trial's results), and if 83% of the patients had had access to and were eligible for dexamethasone, 2,880 deaths from COVID-19 could have been prevented, with a 50% recruitment rate to the RECOVERY trial. This is a much higher figure than the estimate of 260 deaths that were actually prevented due to the RECOVERY trial following the dissemination of the results by press conference rather than waiting for journal publication (Knowlson and Torgerson, 2020). Therefore, at least 2,620 more lives could have been saved with more rapid recruitment.

By closely following both the RECOVERY Collaborative Group's findings (Horby et al., 2021) and the paper on the effectiveness of improving recruitment to the RECOVERY trial on preventing COVID-19-related mortality (Knowlson and Torgerson, 2020), this chapter developed a comprehensive decision tree model and used appropriate probability, utility and cost inputs, in order to assess the *cost-effectiveness* of improving recruitment to the RECOVERY trial from 15% to 50% by employing two full-time research nurses, to assist with recruitment, at each hospital involved in the study. The robustness of the findings was assessed by applying deterministic sensitivity analysis and probabilistic sensitivity analysis (PSA). The aim of this chapter is to demonstrate that poor recruitment to RCTs may put a huge strain on a

national healthcare system in terms of foregone health benefits, due to delays in the dissemination of a potentially cost-effective intervention, or treatment such as dexamethasone plus usual care for reducing the incidence of COVID-19-related death for hospitalised COVID-19 patients.

2.3. Methods

2.3.1. Hypothesis and rationale for using a decision tree model for cost-utility analysis

Improving patient recruitment to the RECOVERY trial from a rate of 15% to 50% would have been an *effective* thing to do due to the foregone health benefits of not having updated the clinical practice with a more *clinically effective* treatment (i.e. dexamethasone plus usual care, instead of usual care only) in the shortest possible time during a global pandemic (Knowlson and Torgerson, 2020). A recruitment strategy which could potentially achieve such an increase in the recruitment rates could be the hiring or redeployment of two Band 5 research nurses to each National Health Service (NHS) hospital affiliated with the study. Assuming it would take an hour on average to speak to each patient about the study, it would be feasible for each research nurse to recruit up to six patients per day. Given a 17% exclusion rate in the study (Jolly et al., 2021) and an assumed refusal rate of 20%, this figure would fall to 3.78 patients a day. Given the proposed recruitment strategy of 352 research nurses, this implies they would be able to recruit 1,331 patients a day, thus 10.5 days (including five working days in one week and two days off work, followed by another 3.5 working days in the following week) would be needed to reach the desirable recruitment figure of 11,303 patients. Therefore, it could be feasible for the proposed recruitment strategy to have accelerated the recruitment rate from 15% to 50%.

Given the low prescription costs and the remarkable effectiveness of dexamethasone in reducing the mortality risk from COVID-19 for ventilated hospitalised patients (National Institute for Health and Care Excellence, 2023, Horby et al., 2021), hypothesis of the chapter is that improving patient recruitment to the RECOVERY trial from a rate of 15% to 50%, by employing two research nurses to each NHS hospital affiliated with the study, would also have been a *cost-effective strategy.* The mechanisms that could confirm such a hypothesis are the following: 1) given the remarkably low prescription costs of dexamethasone and its reported clinical effectiveness in the RECOVERY trial, it is expected that dexamethasone will be a *cost-effective* treatment for the clinical management of hospitalised patients with COVID-19; 2) if dexamethasone is a cost-effective treatment, there will be foregone health benefits in terms of

not having updated in the shortest possible time the clinical practice with a more *cost-effective* treatment earlier than June 2020 (i.e. dexamethasone plus usual care, instead of usual care only during a global pandemic); 3) the incremental costs of the recruitment strategy (i.e. the annual wages of two newly-hired Band 5 full-time nurses) may be low enough to justify the earlier dissemination of the RECOVERY trial's results, for the population to have gained the 'foregone health benefits' of the more cost-effective treatment, thus demonstrating faster recruitment to the RECOVERY trial to be a cost-effective strategy.

To determine whether a faster recruitment to the RECOVERY trial would have been a cost-effective strategy, the cost-effectiveness of dexamethasone is assessed in advance. For the latter, a cost-utility analysis framework is adopted throughout the chapter. Before proceeding to such an analysis, the health benefits of treatments corresponding to COVID-19 disease are defined. These are restricted to COVID-19-related deaths averted for SARS-Cov-2 positive hospitalised patients in the original paper of the RECOVERY trial and the study which explored the potential clinical effects of improving recruitment to the RECOVERY trial from 15% to 50% (Horby et al., 2021, Knowlson and Torgerson, 2020). The chapter, however, uses another dimension of health benefits, which has a broader perspective and considers not only the deaths averted from/with COVID-19 due to a proposed treatment (e.g. dexamethasone plus usual care) but also the associated gains in the expected quality-of-life following COVID-19, for survived hospitalised patients, due to this treatment. Such a dimension is referred to as Quality-Adjusted Life Years (QALYs), which is estimated by summing up an individual patient's yearly health utility weights whose range, in a given year, lies between 0 (i.e. corresponding to death) and 1 (i.e. corresponding to a year of perfect, disease-free health). For any condition reducing the quality of life for an individual patient, there is an associated disutility value that reduces their health utility weight and hence their QALYs. A proposed treatment reducing (increasing) the incidence of such a condition, implies that the probability of a patient experiencing disutility in their health as a result of this condition falls (increases), and hence their *expected QALYs* increase (decrease). In the comparison of two or more treatments, the corresponding expected QALYs for an individual patient are usually estimated for up to the period when the associated treatments would have an effect on their quality of life; such a period may also be dependent upon the condition itself.

Therefore, if dexamethasone plus usual care reduces the incidence of death from COVID-19, it is highly likely that such treatment generates additional expected QALYs for the population

that could not be gained under the baseline treatment of usual care only. However, the expected QALYs gained from dexamethasone plus usual care are also dependent upon the type of usual care a hospitalised SARS-Cov-2-positive patient receives and the likelihood of them experiencing long COVID for up to six months post-infection. Long COVID could potentially be a disease for some patients who survived from COVID-19 and were discharged from hospital, with similar severity for myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) patients (Taquet et al., 2021). By aggregating the individual-level gains in expected QALYs, the population-level gains in expected QALYs, i.e. the incremental health benefit used for undertaking the cost-utility analysis, are estimated. The time horizon is of one year, assuming that a patient would not be reinfected with Sars-Cov-2 within that time period.

To determine whether dexamethasone plus usual care is a cost-effective treatment for reducing the incidence of death from COVID-19 and/or increasing the population QALYs during the COVID-19 pandemic, a cost-utility analysis is undertaken accordingly, to also account for the incremental costs associated with dexamethasone, such as prescription costs. The outcome of interest is the incremental cost per additional QALY gained with dexamethasone plus usual care. The comparator treatment is usual care only. If the ratio of incremental costs to incremental benefits, i.e. the incremental cost-effectiveness ratio (ICER), is lower than a prespecified cost-effectiveness threshold, dexamethasone plus usual care is a cost-effective treatment and the incremental gain in a unit of health outcome, i.e. QALYs gained, can be obtained at a cost less than the maximum value a decision maker is willing to pay to obtain it. The cost-effectiveness threshold is set at £20,000, the minimum value the National Institute for Health and Care Excellence (NICE) uses when it evaluates the cost-effectiveness of proposed treatments or interventions (McCabe et al., 2008). An alternative outcome describing the cost-effectiveness of dexamethasone (and subsequently of faster recruitment to the RECOVERY trial) is the incremental net benefit, defined as:

$$\text{Incremental net benefit}_{A,B} = \text{Incremental benefit}_{A,B} * \text{cost-effectiveness threshold} - \text{incremental cost}_{A,B} \quad \textit{(Equation 2.1)}$$

where A is the proposed treatment, i.e. dexamethasone plus usual care, and B is the baseline treatment, i.e. usual care only. By the definition of *Equation 2.1*, if the incremental net benefit is positive (negative), then dexamethasone is more (less) cost-effective compared to usual care.

Finally, determining which perspective should be adopted in the cost-utility analysis for including appropriate benefits and costs is a crucial question, as there can be several different decision-makers, informed by economic evaluation and interested in different benefits and costs related to a proposed treatment or intervention; these could include patients, clinicians, hospital administrators, national governments, etc. (Drummond et al., 2015). Given the focus of the chapter i.e. clinical management of hospitalised COVID-19 patients, the NHS perspective is followed, which considers the health benefits for the population, such as expected gains in population QALYs, but not non-health benefits such as potential improvements in labour productivity following the dissemination of dexamethasone for hospitalised COVID-19 patients. All costs are health-related and associated with the delivery of COVID-19-related hospital care and the administration of dexamethasone, but do not include non-health-related ones such as compensation payments for workers, e.g. the Furlough Scheme that was introduced by the UK government as a financial support to the self-employed, employees and employers following the imposition of national lockdown (Stuart et al., 2021). A type of economic evaluation that considers both health and non-health benefits and costs, by adopting a societal perspective, is the cost-benefit analysis, which is not considered in the chapter due to lack of relevant data (Drummond et al., 2015). Failure to undertake an analysis from the societal perspective may, in this instance, make the results more conservative as some savings from productivity losses, such as time lost from work due to illness by the patient, are not included.

To undertake such a cost-utility analysis, economic modelling was applied as a tool to develop a model that could represent the real-world healthcare options and the corresponding health outcomes hospitalised patients face with a SARS-Cov-2 positive infection. The design of the modelling is a decision tree. A decision tree is an analytical model that visually represents all possible outcomes for a cohort of patients, via distinct branches. It consists of nodes representing choices or probabilities, which are combined to generate branches. Each node may represent a decision or a chance event, with probabilities totalling to one. Costs and outcomes are assigned to each branch, and these are combined to evaluate different alternatives. Finally, the tree is analysed to compare expected outcomes and costs for each decision point (York Health Economics Consortium, 2016a).

In the chapter's decision tree, the "choice node" refers the assignment of dexamethasone plus usual care or usual care only as available treatments at the point of admission of a SARS-Cov-

2-positive patient to hospital. For each treatment, there are seven "probability nodes" that consider the likelihood of all possible combinations of types of usual care during a patient's hospitalisation with COVID-19; their probabilities are obtained from the findings of the RECOVERY trial (Horby et al., 2021). For each of these probability nodes, there are also two further "probability nodes" of survival and death, with their probabilities also obtained from the findings of the RECOVERY trial (Horby et al., 2021). Therefore, by considering all choice and probability nodes, there are 28 pathways, 14 related to dexamethasone plus usual care and 14 related to usual care only. These are presented in *Table 2.2*, along with their respective probabilities. In line with the definition of decision trees (York Health Economics Consortium, 2016a), expected health outcomes and expected costs were estimated for each of the 28 pathways. The estimation of health outcomes and costs is presented in *Section 2.3.3* and *Section 2.3.4*.

Such a framework allows for undertaking a cost-utility analysis of dexamethasone plus usual care versus usual care only. Since 83% of patients are assumed to be eligible for receiving dexamethasone (Horby et al., 2021), a subsequent cost-utility analysis was undertaken, of updated clinical practice following the dissemination of the RECOVERY trial's results, i.e. 83% of hospitalised COVID-19 patients receive dexamethasone plus usual care and the remaining patients receive usual care only. The incremental QALYs and costs of faster recruitment from 15% to 50%, by employing two full-time nurses to each hospital, are calculated by multiplying the expected QALYs and costs of updated practice by the number of patients that would have benefited from it with a 50% recruitment rate or the number of patients that benefited from it with a 15% recruitment rate, until mid-July 2020 (Knowlson and Torgerson, 2020). In the former case, the labour costs of 352 research nurses are the incremental costs for improving patient recruitment to the study and are considered in the respective cost-utility analysis. More details about the cost-utility analysis of faster recruitment to the RECOVERY trial, from 15% to 50%, are available in *Section 2.3.5*. The incremental net benefit of faster recruitment is an outcome of interest, as defined in *Equation 2.1*, where A refers to 50% recruitment rate by hiring or redeployment of two research nurses to each participating hospital, and B refers to the achieved 15% recruitment rate.

It should be noted that the main alternative type of decision modelling, i.e. Markov model, is not considered in this chapter. A Markov model involves "*recurring events at uncertain times over the lifetime of an individual*" (Kuntz et al., 2013), where 'lifetime' may be restricted to a

specific time period. Such recurrent events can be represented concisely in a Markov model, which is structured around mutually exclusive disease states and can therefore evaluate the cost-effectiveness of treatments associated with chronic diseases, such as gastroesophageal reflux disease (GERD). The present study recognises that, due to the recent emergence of the COVID-19 pandemic, there is noticeable uncertainty with respect to several long-term disease-related outcomes; hence outcomes such as reinfection with Sars-Cov-2, rehospitalisation with COVID-19, and mortality following rehospitalisation with COVID-19, are not considered at this time. The nature of this uncertainty arises from the uncertain effectiveness of future COVID-19 vaccines in reducing the incidence of infection, hospitalisation and death from COVID-19 disease, the potential emergence of immunity-resistant (from infection or vaccination) variants of SARS-Cov-2 and the future public health response to the trajectory of the COVID-19 pandemic. At the time of the RECOVERY trial, when the original strain of SARS-Cov-2 was transmitted across the UK, there were no COVID-19 vaccines available for use and the UK Government implemented a nationwide lockdown. However, the study still considers longer-term outcomes such as recovery from COVID-19 disease for a patient having received invasive mechanical ventilation versus a patient not having received invasive ventilation, and the risk of long COVID for up to six months following infection with SARS-Cov-2. Such outcomes are feasibly modelled using a decision tree, within the model's time horizon of one year. The choice of one year as a time horizon also implies that the cost-effectiveness findings presented may be conservative, as improved recruitment to the trial may have longer-term effects, which however remain uncertain to be identified given the existing evidence. Moreover, peer-reviewed, published studies that assessed the cost-effectiveness of dexamethasone and other treatments evaluated from the RECOVERY trial have also applied decision tree methods (Águas et al., 2021, Carta and Conversano, 2021). Finally, using a Markov model for the decision problem presented in this chapter would imply the development of a model consisting of variable daily/weekly cycles, something which would be computationally demanding whereas it would not significantly alter the cost-effectiveness findings presented in this chapter.

2.3.2. Decision problem

A decision tree model is set up in Microsoft Excel to evaluate the cost-effectiveness of improving recruitment to the RECOVERY trial, from the economic perspective of the UK's National Health Service (NHS). A patient enters the model with suspected or PCR-test-confirmed SARS-Cov-2 infection and is admitted to hospital. At the point of entry, the patient

either receives 6mg of dexamethasone per day for up to 10 days following hospital admission or does not receive dexamethasone. In addition, regardless of whether they are prescribed with dexamethasone, they receive hospital treatment, i.e. usual care. Then, there are several probabilities for which this patient may receive different types of usual care during their hospitalisation, such as admission to an acute hospital ward with no ventilation support, or to an Intensive Care Unit (ICU) with non-invasive (oxygen only but no mechanical ventilation) or invasive ventilation (mechanical ventilation). These probabilities are also dependent on whether or not the patient is treated with dexamethasone at the point of entry, for up to 10 days. Such a set of choice nodes describes the potential different types of usual care a patient hospitalised with COVID-19 may receive. This setting aligns with the RECOVERY trial's design (Horby et al., 2021). Thus, the patient may enter one of the following decision nodes, i.e. treatments:

1) Dexamethasone arm[2]: 6mg of dexamethasone per day plus usual care for up to 10 days, followed by usual care only if hospitalised for longer, or;
2) No Dexamethasone arm: Usual care only for as long as needed.

As already mentioned, these treatments are the choice nodes of the decision tree, whereas the choice nodes are related to the different combinations of types of usual care a COVID-19 patient may receive during their hospitalisation.

In the setting of the trial, a COVID-19 inpatient receiving no ventilation support at the start of their treatment, and in case they have been prescribed with dexamethasone, may either: survive (Pathway 1; P1) or die (P2) without any ventilation support; survive (P3) or die (P4) following subsequent admission to ICU with non-invasive ventilation, but without invasive ventilation; survive (P5) or die (P6) following subsequent admission to ICU with non-invasive ventilation and another admission to ICU and invasive ventilation; survive (P7) or die (P8) following admission to ICU with invasive ventilation, but without non-invasive ventilation. A COVID-19 inpatient admitted to ICU with non-invasive ventilation at the start of their treatment, in case they have been prescribed with dexamethasone, may either: survive (P9) or die (P10) without invasive ventilation; survive (P11) or die (P12) following subsequent admission to ICU with invasive ventilation. A COVID-19 inpatient admitted to ICU with invasive ventilation at

[2] Note: hereafter this will be referred as the Dexamethasone arm.

the start of their treatment, in case they have been prescribed with dexamethasone, may either: survive (P13) or die (P14).

Then, a COVID-19 inpatient receiving no ventilation support at the start of their treatment, and in case they have not been prescribed with dexamethasone, may either: survive (P15) or die (P16) without any ventilation support; survive (P17) or die (P18) following subsequent admission to ICU with non-invasive ventilation, but without invasive ventilation; survive (P19) or die (P20) following subsequent admission to ICU with non-invasive ventilation and another admission to ICU and invasive ventilation; survive (P21) or die (P22) following admission to ICU with invasive ventilation, but without non-invasive ventilation. A COVID-19 inpatient admitted to ICU with non-invasive ventilation at the start of their treatment, in case they have not been prescribed with dexamethasone, may either: survive (P23) or die (P24) without invasive ventilation; survive (P25) or die (P26) following subsequent admission to ICU with invasive ventilation. A COVID-19 inpatient admitted to ICU with invasive ventilation at the start of their treatment, in case they have not been prescribed with dexamethasone, may either: survive (P27) or die (P28).

These 28 pathways are developed according to the clinical outcomes reported in the RECOVERY trial (Horby et al., 2021). The pathways that are associated with survival from COVID-19 disease also capture the possibility of a patient experiencing long COVID for up to 6 months after SARS-Cov-2 infection. The probability inputs for long COVID are differentiated for patients who received invasive ventilation from those who did not (Taquet et al., 2021). Long COVID is incorporated in the model as being characterised by persistent fatigue, with severity assumed to be similar to myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). Due to the uncertain trajectory of the pandemic with respect to the emergence of future antibody-resistant variants (Servellita et al., 2022), the case where a COVID-19 inpatient becomes reinfected with SARS-Cov-2 in the future is not considered. Consequently, outcomes such as reinfection with Sars-Cov-2, rehospitalisation with COVID-19 and mortality following rehospitalisation with COVID-19 are not considered. The pathways' corresponding probabilities, shown in *Table 2.2*, are differentiated by treatment arms and have been computed using the probabilities of types of usual care received at admission and clinical outcomes (e.g. death, transition to a more invasive type of care) (see *Table 2.1* and *Supplemental Material 2.1* for more details) (Horby et al., 2021). The corresponding decision tree is shown in *Figure 2.1*.

Table 2.1: Probability inputs with respect to decision nodes at admission, clinical outcomes and long COVID

Probabilities related to decision nodes available at admission	**Source**	**All groups**	
P(acute hospital ward at admission)	(Horby et al., 2021)	0.239	
P(non-invasive ventilation received at admission)	(Horby et al., 2021)	0.604	
P(invasive ventilation received at admission)	(Horby et al., 2021)	0.157	
Probabilities of clinical outcomes during admission with COVID-19	**Source**	**Dexamethasone**	**No Dexamethasone**
P(death|acute hospital ward)	(Horby et al., 2021)	0.166	0.132
P(non-invasive ventilation|acute hospital ward)	(Horby et al., 2021)	0.032	0.044
P(invasive ventilation|acute hospital ward)	(Horby et al., 2021)	0.018	0.029
P(survival|acute hospital ward)	(Horby et al., 2021)	0.784	0.794
P(death| non-invasive ventilation)	(Horby et al., 2021)	0.206	0.217
P(invasive ventilation|non-invasive ventilation)	(Horby et al., 2021)	0.079	0.107
P(survival|non-invasive ventilation)	(Horby et al., 2021)	0.715	0.676
P(death| invasive ventilation)	(Horby et al., 2021)	0.293	0.414
P(survival| invasive ventilation)	(Horby et al., 2021)	0.707	0.586
Probabilities related to long COVID	**Source**	**All groups**	
P(long COVID; population total)	(Taquet et al., 2021)	0.128	
P(long COVID, ICU all)	(Taquet et al., 2021)	0.262	

Figure 2.1: Decision tree

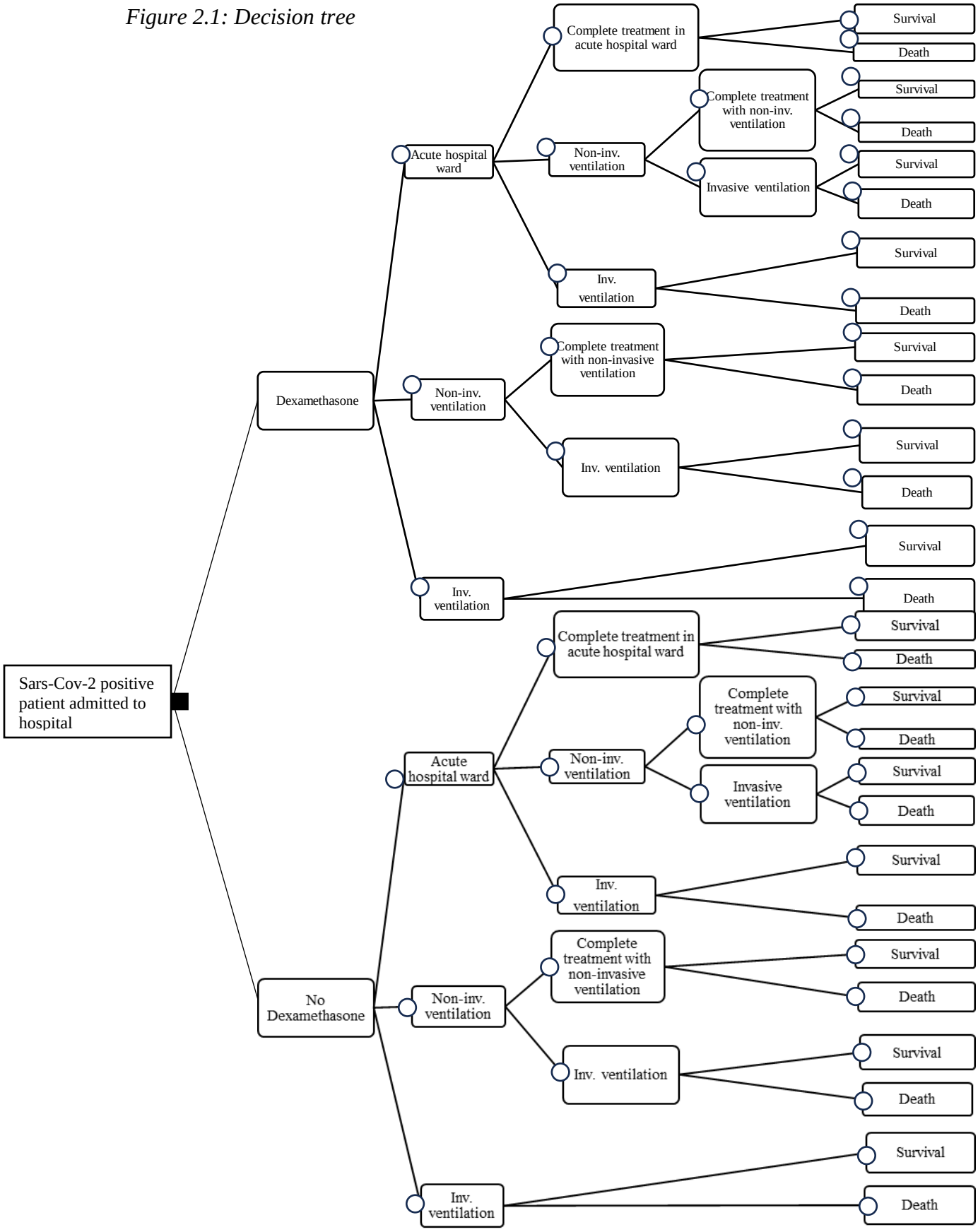


Table 2.2: Pathways and their associated probabilities in the two arms

Pathway number	Clinical pathway	Treatment	Probability
1	Acute hospital ward (survival)	Dexamethasone	0.187
2	Acute hospital ward (death)	Dexamethasone	0.040
3	Acute hospital ward, non-invasive ventilation (survival)	Dexamethasone	0.005
4	Acute hospital ward, non-invasive ventilation (death)	Dexamethasone	0.002
5	Acute hospital ward, non-invasive ventilation, invasive ventilation (survival)	Dexamethasone	0.003
6	Acute hospital ward, non-invasive ventilation, invasive ventilation (death)	Dexamethasone	0.001
7	Acute hospital ward, invasive ventilation (survival)	Dexamethasone	0.000
8	Acute hospital ward, invasive ventilation (death)	Dexamethasone	0.000
9	Non-invasive ventilation (survival)	Dexamethasone	0.432
10	Non-invasive ventilation (death)	Dexamethasone	0.125
11	Oxygen, invasive ventilation (survival)	Dexamethasone	0.034
12	Oxygen, invasive ventilation (death)	Dexamethasone	0.014
13	Invasive ventilation (survival)	Dexamethasone	0.111
14	Invasive ventilation (death)	Dexamethasone	0.046
15	Acute hospital ward (survival)	No Dexamethasone	0.190
16	Acute hospital ward (death)	No Dexamethasone	0.032
17	Acute hospital ward, non-invasive ventilation (survival)	No Dexamethasone	0.007
18	Acute hospital ward, non-invasive ventilation (death)	No Dexamethasone	0.002
19	Acute hospital ward, non-invasive ventilation, invasive ventilation (survival)	No Dexamethasone	0.004
20	Acute hospital ward, non-invasive ventilation, invasive ventilation (death)	No Dexamethasone	0.003
21	Acute hospital ward, invasive ventilation (survival)	No Dexamethasone	0.001
22	Acute hospital ward, invasive ventilation (death)	No Dexamethasone	0.000
23	Non-invasive ventilation (survival)	No Dexamethasone	0.408
24	Non-invasive ventilation (death)	No Dexamethasone	0.131
25	Oxygen, invasive ventilation (survival)	No Dexamethasone	0.038
26	Oxygen, invasive ventilation (death)	No Dexamethasone	0.027
27	Invasive ventilation (survival)	No Dexamethasone	0.092
28	Invasive ventilation (death)	No Dexamethasone	0.065

2.3.3. Quality-Adjusted Life Years (QALYs)

To estimate the Quality-Adjusted Life Years (QALY) gained for each pathway, data on the utility and day weights for each (combination of) decision node and health outcome (i.e. survival, death) are collected. Each pathway's appropriate utility weights are then multiplied by the corresponding day weights (relative to a year of 365.25 days) to generate their yearly QALYs. A time horizon of one year is used to estimate the aggregate individual QALYs associated with each of the 28 pathways.

Population-wide and age-specific utility weights related to the average health state of the population residing in England are collected (Szende et al., 2014). Stratified QALYs for the 45-64, 65-74 and 75+ age groups are estimated, as these groups had the highest risk of hospital admission along with/due to SARS-Cov-2 infection (St Sauver et al., 2021). Following this, age-independent disutility weights are obtained for SARS-Cov-2 infection, hospitalisation in an acute ward, and hospitalisation along with non-invasive and invasive ventilation support (Sheinson et al., 2021). In addition, an age-independent disutility weight for long COVID, identical to that of ME/CFS (Hvidberg et al., 2015), is incorporated. For each decision node and outcome, population-wide and age-specific utility weights are generated by subtracting from the initial utility weights the corresponding disutility weights, for the population as a whole and the age groups. The corresponding inputs are shown in *Table 2.3*.

To obtain the appropriate day weights, estimates of the duration of Sars-Cov-2 infection before admission to hospital (Sutherland et al., 2021), hospitalisation days associated with each pathway (Sheinson et al., 2021, Vekaria et al., 2021), and the duration of long COVID following survival (Taquet et al., 2021) are collected, all of which are divided by 365.25 days to obtain the corresponding weights. The remaining days each year are treated as COVID-free days, and the corresponding day weights are attached and multiplied by the appropriate utility weight of the average health state (Sutherland et al., 2021). All patients are assumed to be directly discharged from hospital following the relapse of their COVID-related symptoms, as there is an absence of relevant data from the RECOVERY trial and the literature. Long COVID is incorporated into the estimation of QALYs gained by multiplying the likelihood of it (not) occurring (see *Table 2.1*), by the corresponding disutility weight.

Hospitalisation days are estimated, for each combination of decision nodes, by the best study to date regarding the length of stay (LoS) in NHS hospitals due to COVID-19 disease (Vekaria et al., 2021). However, since there is no distinction between non-invasive ventilation and invasive ventilation in the study's estimates, estimates from an American study are additionally used to differentiate the LoS between non-invasive ventilation and invasive ventilation (Sheinson et al., 2021). For clinical pathways, where a clinical course involves a less invasive and a more invasive choice node, the assumption is that the maximum expected number of days with the less invasive node are needed before a patient is transferred to a more invasive node. Finally, due to lack of sufficient evidence, there is no difference in the hospitalisation days for each clinical scenario by the type of treatment COVID-19 inpatients received (i.e. Dexamethasone versus No Dexamethasone). It was possible to derive differentiated hospitalisation days for those who survived and those who died, with those dying from COVID staying in hospital for fewer days (Vekaria et al., 2021). Due to absence of empirical evidence, the number of days spent in non-invasive ventilation for patients dying from COVID-19, are assumed to be equal to the lowest bound estimate provided by (Sheinson et al., 2021). All QALYs gained for each of the 28 pathways, stratified by age group, are shown in *Table 2.4.*

2.3.4. Costs

Daily costs per capita are included for: treatment in an acute hospital ward (code: XC07Z) (Department of Health and Social Care, 2019); supplying non-invasive ventilation (code: XC06Z, DZ37A) (Department of Health and Social Care, 2019); supplying invasive ventilation (code: XC06Z, DZ37A) (Department of Health and Social Care, 2019); and providing 6mg of dexamethasone, according to NICE Tariff Prices (National Institute for Health and Care Excellence, 2023). As the NHS Reference costs make no distinction between non-invasive and invasive ventilation (Department of Health and Social Care, 2019), the cost of supplying invasive ventilation is calculated through a multiplication of the cost of non-invasive ventilation reported from the NHS Reference Costs by a multiplier, i.e. 1.258, from a review that has estimated the impact of invasive ventilation on the daily costs of ICU care (Kaier et al., 2019). Where costs are only available in 2018 price levels, such as the daily costs of different types of usual care, the UK GDP deflator is used to adjust the figures to 2020 prices (World Bank, 2020). Note that daily costs related to hospitalisation were multiplied by the

expected number of hospitalisation days corresponding to each pathway (Sheinson et al., 2021, Vekaria et al., 2021).

Table 2.3: Inputs associated with pathways' QALYs

Input	Source	Utility weight (45 to 64 age group)	Utility weight (65 to 74 age group)	Utility weight (75+ age group)	Utility weight (Population total)	Disutility weight (all age groups)
Average healthy state	(Szende et al., 2014)	0.849	0.785	0.734	0.856	N/A
COVID-19 Infection	(Szende et al., 2014, Sheinson et al., 2021)	0.579	0.515	0.464	0.586	0.27
Hospitalisation (Acute ward, no ventilation received)	(Szende et al., 2014, Sheinson et al., 2021)	0.469	0.405	0.354	0.476	0.11 (plus SARS-Cov-2 infection)
Hospitalisation (Non-invasive ventilation received)	(Szende et al., 2014, Sheinson et al., 2021)	0.219	0.155	0.104	0.226	0.36(plus SARS-Cov-2 infection)
Hospitalisation (Invasive ventilation received)	(Szende et al., 2014, Sheinson et al., 2021)	0.019	0	0	0.026	0.56 (plus SARS-Cov-2 infection)
Death	N/A	0	0	0	0	N/A
Long COVID (post-COVID syndrome)	(Szende et al., 2014, Hvidberg et al., 2015)	0.559	0.495	0.444	0.566	0.29

Input	Source	Figure
Days of COVID-19 infection	(Sutherland et al., 2021)	6.7
Days of long COVID (from the beginning of infection with SARS-Cov-2)	(Taquet et al., 2021)	182.63
Maximum hospitalisation days (acute hospital ward, survival)	(Vekaria et al., 2021)	9.4
Maximum hospitalisation days (acute hospital ward, death)	(Vekaria et al., 2021)	8.3
Maximum hospitalisation days (non-invasive ventilation, survival)	(Sheinson et al., 2021)	12.58
Maximum hospitalisation days (non-invasive ventilation, death)	(Sheinson et al., 2021, Vekaria et al., 2021)	9.41
Maximum hospitalisation days (invasive ventilation, survival)	(Sheinson et al., 2021, Vekaria et al., 2021)	24.5
Maximum hospitalisation days (non-invasive ventilation, death)	(Sheinson et al., 2021, Vekaria et al., 2021)	15.8

Table 2.4: QALYs gained by clinical pathway and age group

Pathway number	Clinical pathway	Quality-Adjusted Life Years (QALYs) gained			
		45-64 age group	65-74 age group	75+ age group	Population
1,15	Acute hospital ward (survival)	0.816	0.748	0.697	0.824
2, 16	Acute hospital ward (death)	0.021	0.019	0.017	0.022
3, 17	Acute hospital ward, non-invasive ventilation (survival)	0.806	0.737	0.686	0.814
4, 18	Acute hospital ward, non-invasive ventilation (death)	0.018	0.015	0.013	0.018
5, 19	Acute hospital ward, non-invasive ventilation, invasive ventilation (survival)	0.755	0.691	0.640	0.762
6, 20	Acute hospital ward, non-invasive ventilation, invasive ventilation (death)	0.014	0.010	0.006	0.014
7, 21	Acute hospital ward, invasive ventilation (survival)	0.761	0.740	0.646	0.768
8, 22	Acute hospital ward, invasive ventilation (death)	0.018	0.014	0.010	0.019
9, 23	Non-invasive ventilation (survival)	0.805	0.736	0.685	0.813
10, 24	Non-invasive ventilation (death)	0.016	0.013	0.011	0.017
11, 25	Oxygen, invasive ventilation (survival)	0.764	0.700	0.649	0.771
12, 26	Oxygen, invasive ventilation (death)	0.017	0.013	0.010	0.017
13, 27	Invasive ventilation (survival)	0.757	0.693	0.642	0.764
14, 28	Invasive ventilation (death)	0.011	0.008	0.004	0.012

If a patient survives COVID-19 disease, they have a chance of getting long COVID. In this case, it is assumed they would need up to two GP consultations in a 6-month period for screening and treatment purposes. Therefore, the unit costs of two GP appointments are also included, reported in 2020 price levels (Curtis and Burns, 2020). The dexamethasone arm faces the same costs as the No Dexamethasone arm, but with the addition of the daily cost of receiving 6mg of dexamethasone for up to 10 days. 28 Dexamethasone 2mg tablets have an NHS indicative price of £2.10, 50 tablets £3.43 and 100 tablets £8.93 (National Institute for Health and Care Excellence, 2023), implying that three tablets of 2mg (leading to the daily intake of 6mg of dexamethasone) cost £0.23, £0.21, and £0.27 correspondingly. The median value of £0.23 is considered for the baseline cost-effectiveness analysis, whereas the other two prices are considered for sensitivity analyses. Note that dexamethasone prices are reported in 2023 levels, as it was not possible to obtain accurate pricing figures of dexamethasone reported in 2020. Since all relevant costs are expected to occur within the first year of a patient being hospitalised with/due to SARS-Cov-2 infection, there is no discounting of future costs. In

addition, all costs are independent of age groups. The corresponding cost inputs are presented in *Table 2.5*. Each pathway's cumulative costs, by treatment arm, are shown in *Table 2.6*.

The annual salary cost of a Band 5 research nurse is also reported, although it will not be considered as a cost input for the cost-utility analysis of dexamethasone; instead, it will be considered for the subsequent cost-utility analysis of faster recruitment to the RECOVERY trial as it is considered to be the strategy which would have raised the recruitment rate from 15% to 50%. The salary cost (i.e. £39,841) reflects the annual gross wage of a Band 5 research nurse (i.e. £30,615 (National Health Service, 2020)) plus the salary oncosts (i.e. employer's national insurance contributions plus 20.68% of gross salary for employer's contribution to superannuation (Curtis and Burns, 2020)). The consideration of additional overhead costs for research nurses is excluded due to uncertainty about the extent of potential employment of new nurses by the NHS for patient recruitment purposes to the RECOVERY trial. This scenario would involve assessing overhead costs, which may differ from those reported in the Personal Social Services Research Unit (PSSRU) unit cost estimates given the tasks a research nurse would undertake for the study at the onset of a global pandemic, which may be different to those a Band 5 nurse normally undertakes (Kaier et al., 2019).

Table 2.5: Cost inputs

Input	Source	Cost (£)
Unit cost of GP appointment (long COVID patients)	(Curtis and Burns, 2020)	£39.23
Daily cost of staying in acute hospital ward	(Department of Health and Social Care, 2019)	£748.41
Daily cost of non-invasive ventilation (oxygen)	(Department of Health and Social Care, 2019)	£1,394.23
Daily cost of invasive ventilation	(Department of Health and Social Care, 2019, Kaier et al., 2019, World Bank, 2020)	£1,753.94
Daily cost of providing 6mg of dexamethasone	(National Institute for Health and Care Excellence, 2023)	£0.23
Annual cost of research nurse (Band 5 maximum)	(National Health Service, 2020, Curtis and Burns, 2020)	£39,840.77

Table 2.6: Costs by clinical pathway and treatment arm

Pathway number	Clinical pathway	Costs (£)	
		Dexamethasone	No Dexamethasone
1,15	Acute hospital ward (survival)	£7,047.23	£7,045.11
2, 16	Acute hospital ward (death)	£6,213.67	£6,211.80
3, 17	Acute hospital ward, non-invasive ventilation (survival)	£16,260.08	£16,257.83
4, 18	Acute hospital ward, non-invasive ventilation (death)	£11,830.18	£11,828.06
5, 19	Acute hospital ward, non-invasive ventilation, invasive ventilation (survival)	£44,480.69	£44,478.44
6, 20	Acute hospital ward, non-invasive ventilation, invasive ventilation (death)	£29,211.34	£29,209.09
7, 21	Acute hospital ward, invasive ventilation (survival)	£40,674.95	£40,672.70
8, 22	Acute hospital ward, invasive ventilation (death)	£26,545.88	£26,543.63
9, 23	Non-invasive ventilation (survival)	£17,551.72	£17,549.47
10, 24	Non-invasive ventilation (death)	£13,121.82	£13,119.70
11, 25	Oxygen, invasive ventilation (survival)	£38,458.70	£38,456.45
12, 26	Oxygen, invasive ventilation (death)	£24,329.64	£24,327.39
13, 27	Invasive ventilation (survival)	£42,983.87	£42,981.62
14, 28	Invasive ventilation (death)	£27,714.52	£27,712.27

2.3.5. Cost-effectiveness of faster recruitment to the RECOVERY Trial

A multi-step approach is followed so as to estimate the cost-effectiveness of increasing recruitment to the RECOVERY trial from 15% to 50%. The increased recruitment would not be targeted only at the dexamethasone plus usual care, i.e. Dexamethasone, and usual care only comparison, i.e. No Dexamethasone, since at the time there were multiple treatments under evaluation in the RECOVERY as it was a 'platform' trial. One ineffective treatment comparison (hydroxychloroquine) was reported before the dexamethasone comparison; the arm was closed early due to treatment 'futility'.

Step 1: For each treatment arm and age group, each pathway's QALYs (*Table 2.4*) and costs (*Table 2.6*) are multiplied by the corresponding probabilities from *Table 2.2*, to find each pathway's expected QALYs and costs.

Step 2: The incremental QALYs and costs of Dexamethasone against No Dexamethasone are calculated, to estimate the cost-effectiveness of dexamethasone on reducing the incidence of COVID-related mortality, for different age groups and the population. Results are presented in terms of the incremental cost-effectiveness ratio (ICER), i.e. the incremental cost per QALY, and as incremental net benefit, using a £20,000 cost-effectiveness threshold, the threshold used for all analyses.

Step 3: As dexamethasone was (cost-) effective for reducing the incidence of COVID-related mortality, the updated clinical practice encouraged the 10-day provision of dexamethasone to hospitalised COVID patients. However, only 83% of hospitalised COVID patients can receive Dexamethasone; hence the remaining (17%) hospitalised COVID patients still receive No Dexamethasone. Therefore, the cost-effectiveness of updated practice (83% Dexamethasone, 17% No Dexamethasone) against previous clinical practice (100% No Dexamethasone) is estimated, by weighting appropriately the expected QALYs and costs from Step 1, for different age groups and the population, with findings presented as ICERs and incremental net benefit estimates.

Step 4: With a 15% recruitment rate, 6,980 hospitalised patients originally benefited from the updated practice up to mid-July 2020 (15% recruitment rate) (Knowlson and Torgerson, 2020). With a 50% recruitment rate, 77,310 hospitalised patients would have benefited from the updated practice up to mid-July 2020 (Knowlson and Torgerson, 2020). The main incremental cost of increasing recruitment would be a corresponding recruitment strategy; the assumption is that employing two research nurses for each NHS hospital involved in the study (176 participating hospitals) would have been a sufficient recruitment strategy. The maximum 2020/21 annual gross salary for a Band 5 nurse is £30,615(National Health Service, 2020),with the estimated annual salary cost being £39,841 (NHS 2020; Curtis and Burns 2020). To break down the number of benefited patients and the costs of nurses into age groups, estimates for the proportion of COVID-19 inpatients at NHS hospitals is used, according to the daily averages of the age distribution from 12/10/2020 to 04/01/2022 (National Health Service, 2023).

Step 5: The incremental QALYs and costs of increasing recruitment from 15% to 50%, by employing two full-time nurses to each hospital, are calculated by multiplying the expected QALYs and costs of updated practice (Step 3), by the number of patients that would have benefited from it with a 50% recruitment rate or the number of patients that benefited from it with a 15% recruitment rate (Step 4). In the former case, the annual labour costs of 352 nurses (i.e. two nurses per site) are additional costs for improving patient recruitment to the study. The outcome is the incremental net benefit of faster recruitment to the RECOVERY trial, and the ICER. The cost-utility analyses are stratified by age group, in addition to the presentation of findings for the population overall.

2.3.6. Sensitivity analysis

2.3.6.1. Deterministic sensitivity analysis

In decision tree models, there is usually a suspicion of uncertainty about the input values used for interpreting the results of the cost-effectiveness of a given strategy (Gray et al., 2011). To evaluate the impact of parameter uncertainty on the cost-effectiveness of faster recruitment to the RECOVERY trial from 15% to 50% by hiring or redeploying two research nurses to each hospital, ranges of several probability, utility and cost inputs in *Table 2.1, Table 2.3 and Table 2.5* were collected from the literature. Such ranges were used to observe the variations in the primary outcome of interest, i.e. the incremental net benefit of faster recruitment to the RECOVERY trial. This process is defined as sensitivity analysis, and more specifically as *deterministic* sensitivity analysis, where a single (one-way sensitivity analysis) or double (two-way sensitivity analysis) input values are changed to estimate the variations, assuming everything else constant, in the incremental net benefit of faster recruitment to the RECOVERY trial. In this way the deterministic sensitivity analysis can evaluate the impact of input parameter uncertainty on the robustness of the chapter's findings.

The input parameter ranges, alongside the sources that were used, are shown in *Table 2.7*. Note that no sensitivity analysis of the incremental net benefit of faster recruitment to the RECOVERY trial is undertaken, with respect to the utility weight of the average health state, as there is no recommended range associated with this figure (Szende et al., 2014). Furthermore, the sensitivity analysis with respect to costs of hospitalisation in an acute hospital ward, non-invasive ventilation, and invasive ventilation is undertaken manually due to lack of confidence intervals from the original studies and datasets. One-way sensitivity analyses are undertaken with respect to all inputs shown in *Table 2.7*. Two-way sensitivity analyses are undertaken with respect to the mortality risk from COVID-19, for different combinations of final decision nodes and treatments, to explore whether increasing recruitment to the RECOVERY trial would have been a cost-effective strategy even if Dexamethasone were an ineffective, or a less effective, treatment compared to No Dexamethasone for lowering the incidence of COVID-19-related deaths for COVID-19 hospitalised patients. No deterministic sensitivity analysis stratified by age groups is undertaken. The inputs used for the deterministic sensitivity analyses are shown in *Table 2.7*.

Table 2.7: Ranges of inputs for deterministic sensitivity analysis

Input	Source	Range
Probability inputs		
All probability inputs except long COVID	(Horby et al., 2021, Taquet et al., 2021)	(0.0, 1.0)
P (long COVID; population total)	(Taquet et al., 2021)	(0.1256, 0.1309)
P (long COVID; ICU all)	(Taquet et al., 2021)	(0.249, 0.2754)
Disutility of COVID-19 Infection	(Sheinson et al., 2021, Vekaria et al., 2021)	(0.0, 0.95)
Disutility of Hospitalisation (Acute ward, no ventilation received)	(Sheinson et al., 2021, Vekaria et al., 2021)	(0.0, 1.0)
Disutility of Hospitalisation (Non-invasive ventilation received)	(Sheinson et al., 2021, Vekaria et al., 2021)	(0.0, 0.96)
Disutility of Hospitalisation (Invasive ventilation received)	(Sheinson et al., 2021, Vekaria et al., 2021)	(0.03, 0.99)
Cost inputs		
Unit cost of GP appointment (long COVID patients) (5-15 minutes of GP consultation; £4.3 per minute)	(Curtis and Burns, 2020)	(£21.5, £64.5)
Daily cost of staying in acute hospital ward	Manual, (Department of Health and Social Care, 2019)	(£500, £1,250)
Daily cost of non-invasive ventilation (oxygen)	Manual, (Department of Health and Social Care, 2019)	(£800, £1,800
Daily cost of invasive ventilation	Manual, (Department of Health and Social Care, 2019)	(£1,400, £4,800)
Daily cost of providing 6mg of dexamethasone	(National Institute for Health and Care Excellence, 2023)	(£0.21, £0.27)
Annual salary cost of a research nurse (Band 5 minimum to Band 6 maximum)	(National Health Service, 2020, Curtis and Burns, 2020)	(£35,097, £43,958)

2.3.6.2. Probabilistic Sensitivity Analysis (PSA)

Whereas a deterministic sensitivity analysis can capture the level of confidence in the estimate of the incremental net benefit of faster recruitment, with respect to variations in one (or two) of the model's input values, it cannot wholly and simultaneously capture the findings' robustness with respect to the uncertainty of inputs of all models. Consequently, another technique in health economic modelling is needed to capture such an uncertainty; this is called probabilistic sensitivity analysis (PSA).

Under PSA, all input parameters "*are represented as distributions around the point estimate*", for which statistical measures such as the mean and the standard error were collected from other studies (York Health Economics Consortium, 2016b). Such distributions reflect the domain of the input values, implying that different distributions were considered for utility (Beta), probability (Beta) and cost (Gamma) inputs.

The Beta distribution Beta (α, β) is a continuous probability distribution which is defined on the interval [0,1] in terms of two parameters α>0 and β>0, where α-1 reflects the number of successes and β-1 the number of failures. If a random variable, such as the probability of death following invasive ventilation and treatment with dexamethasone or the disutility weight of hospitalisation with invasive ventilation, is defined as X ~ Beta (α, β), then its mean $E(X)=\frac{\alpha}{\alpha+\beta}$ and variance $Var(X)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ which implies that: $\alpha=\left(\frac{1-E(X)}{Var(X)}-\frac{1}{E(X)}\right)E(X)^2$ and $\beta=\alpha\left(\frac{1}{E(X)}-1\right)$, where E(X) is the reported value of the parameter from other studies, such as the results from the RECOVERY Trial (Horby et al., 2021), and Var(X) is approximated as the standard error of E(X). Given E(X) and Var(X), the alpha and beta values of all probability and utility inputs were obtained using Excel to generate their corresponding Beta distributions used for PSA, as shown in *Table 2.8*.

The Gamma distribution Gamma (α, β) is a continuous probability distribution which is strictly positive in terms of two parameters α>0 and β>0, where α reflects the shape of the distribution and β is a strictly positive rate parameter. Such distribution accounts for the strictly positive figures of the costs, as well as for the skewness observed with cost data. If a random variable, such as the daily cost of a patient staying in an acute hospital ward, is defined as X ~ Gamma (α, β), then its mean is $E(X)=\frac{\alpha}{\beta}$ and variance is $Var(X)=\frac{\alpha}{\beta^2}$, which implies that:

$\alpha=\frac{E(X)}{\beta}$ and $\beta=\frac{Var(X)}{E(X)}$, where E(X) is the reported value of the parameter from other studies, such as the NHS Reference Costs (Department of Health and Social Care, 2019) and Var(X) is approximated as the standard error of E(X). Given E(X) and Var(X), the alpha and beta values of all probability and utility inputs were obtained using Excel to generate their corresponding Gamma distributions used for PSA as shown in *Table 2.8*. In the case of an effectiveness input (such as risk ratio (RR)), which is not applicable in this chapter, its chosen distribution would have been a log-normal distribution.

Following these distributions, a random sampling of 10,000 iterations of input parameter values across all the generated distributions is applied to yield distributions of incremental QALYs and incremental costs, and therefore a distribution of the incremental net benefit of faster

recruitment to the RECOVERY trial. Under this distribution, it also becomes feasible to estimate the 95% confidence interval of the incremental net benefit of faster recruitment to the RECOVERY trial, an important estimate to confirm whether the estimated mean value of the incremental net benefit of faster recruitment is statistically robust and not significantly influenced by the underlying parameter uncertainty. The distributions of the incremental QALYs and the incremental costs are also graphically presented on a cost-effectiveness plane, showing the random joint values of incremental QALYs and incremental costs under which faster recruitment to the RECOVERY trial is (or is not) a cost-effective strategy, in relation to the £20,000 cost-effectiveness threshold. Finally, a cost-effectiveness acceptability curve (CEAC) is also generated from the PSA findings, representing the probability of faster recruitment to the RECOVERY trial being cost-effective under different values of the cost-effectiveness threshold (which was ranged from £0 to £50,000). No PSA stratified by age groups is undertaken. The PSA is undertaken in Excel, through the conversion of the baseline deterministic decision model to a probabilistic one and the use of Excel Visual Basic for Applications (VBA) to generate random values of input parameter values and consequently of incremental QALYs and incremental costs.

The 32 inputs used for PSA, together with their mean, standard error, alpha and beta values, are shown in *Table 2.8*.

Table 2.8: Inputs for probabilistic sensitivity analysis (PSA)

Input	Source	Mean	Standard error	Alpha	Beta	Distribution
Probability inputs						
Proportion of patients who can have dexamethasone	(Horby et al., 2021)	0.8277	0.0036	9354.17	1947.83	Beta
P (acute hospital ward)	(Horby et al., 2021)	0.2389	0.0053	1534.76	4889.24	Beta
P (non-invasive ventilation)	(Horby et al., 2021)	0.6044	0.0061	3882.40	2541.60	Beta
P (invasive ventilation)	(Horby et al., 2021)	0.1567	0.0045	1006.84	5417.16	Beta
P (death|acute hospital ward, Dexamethasone)	(Horby et al., 2021)	0.1657	0.0166	82.83	417.17	Beta
P (non-invasive ventilation|acute hospital ward, Dexamethasone)	(Horby et al., 2021)	0.0319	0.0079	15.97	484.03	Beta
P (invasive ventilation|acute hospital ward, Dexamethasone)	(Horby et al., 2021)	0.0180	0.0059	8.98	491.02	Beta
P (death|acute hospital ward, No Dexamethasone)	(Horby et al., 2021)	0.1325	0.0105	136.87	896.13	Beta
P (non-invasive ventilation|acute hospital ward, No Dexamethasone)	(Horby et al., 2021)	0.0445	0.0064	45.96	987.04	Beta
P (invasive ventilation|acute hospital ward, No Dexamethasone)	(Horby et al., 2021)	0.0290	0.0052	29.97	1003.03	Beta
P (death| non-invasive ventilation, Dexamethasone)	(Horby et al., 2021)	0.2064	0.0113	263.79	1014.21	Beta
P (invasive ventilation|non-invasive ventilation, Dexamethasone)	(Horby et al., 2021)	0.0790	0.0075	100.92	1177.08	Beta
P (death| non-invasive ventilation, No Dexamethasone)	(Horby et al., 2021)	0.2170	0.0081	564.78	2038.22	Beta
P (invasive ventilation|non-invasive ventilation, No Dexamethasone)	(Horby et al., 2021)	0.1071	0.0061	278.89	2324.11	Beta
P (death| invasive ventilation, Dexamethasone)	(Horby et al., 2021)	0.2932	0.0089	763.23	1839.77	Beta
P (death| invasive ventilation, No Dexamethasone)	(Horby et al., 2021)	0.4143	0.0097	1078.55	1524.45	Beta
P (long COVID; population total)	(Taquet et al., 2021)	0.1282	0.0006	35077.87	238539.13	Beta
P (long COVID; ICU all)	(Taquet et al., 2021)	0.2623	0.0008	71769.74	201847.26	Beta
Utility inputs						
Disutility of COVID-19 Infection	(Sheinson et al., 2021, Vekaria et al., 2021)	0.27	0.3	0.32	0.87	Beta
Disutility of Hospitalisation (Acute ward, no ventilation received)	(Sheinson et al., 2021, Vekaria et al., 2021)	0.11	0.3	0.01	0.08	Beta
Disutility of Hospitalisation (Non-invasive ventilation received)	(Sheinson et al., 2021, Vekaria et al., 2021)	0.36	0.3	0.56	1.00	Beta
Disutility of Hospitalisation (Invasive ventilation received)	(Sheinson et al., 2021, Vekaria et al., 2021)	0.56	0.3	0.97	0.76	Beta
Disutility of long COVID	(Szende et al., 2014, Hvidberg et al., 2015)	0.29	0.0306	63.43	155.29	Beta
Cost inputs						
Unit cost of GP appointment (long COVID patients)	(Curtis and Burns, 2020)	£39.23	10.9694	12.79	3.07	Gamma
Daily cost of staying in acute hospital ward	Manual, (Department of Health and Social Care, 2019)	£748.41	191.3265	15.30	48.91	Gamma
Daily cost of non-invasive ventilation (oxygen)	Manual, (Department of Health and Social Care, 2019)	£1,394.23	255.1020	29.87	46.68	Gamma
Daily cost of invasive ventilation	Manual, (Department of Health and Social Care, 2019)	£1,753.94	867.3469	4.09	428.91	Gamma
Daily cost of providing 6mg of dexamethasone	(National Institute for Health and Care Excellence, 2023)	£0.23	0.0362	38.63	0.01	Gamma
Annual salary cost of research nurse	(National Health Service, 2020, Curtis and Burns, 2020)	£39,840.77	2260.4592	310.64	128.25	Gamma

2.4. Results

2.4.1. Cost-effectiveness of faster recruitment to the RECOVERY Trial

2.4.1.1. Step 1: Expected QALYs and costs

For both the Dexamethasone and the No Dexamethasone arms, their pathway QALYs (as shown in *Table 2.4*) and costs (as shown in *Table 2.6*) are multiplied by their pathway probabilities (as shown in *Table 2.2*) to estimate the expected QALYs and costs. The expected QALYs, stratified by age group and the population in total, are shown in *Table 2.9.*1 for the Dexamethasone arm and *Table 2.9.2* for the No Dexamethasone arm. The expected costs are shown in *Table 2.10* for the Dexamethasone and the No Dexamethasone arms.

Table 2.9.1: Expected QALYs (Dexamethasone arm)

Pathway number	Clinical pathway	Expected QALYs (Dexamethasone arm)			
		45-64 age group	65-74 age group	75+ age group	Population
1	Acute hospital ward (survival)	0.153	0.140	0.131	0.154
2	Acute hospital ward (death)	0.001	0.001	0.001	0.001
3	Acute hospital ward, non-invasive ventilation (survival)	0.004	0.004	0.004	0.004
4	Acute hospital ward, non-invasive ventilation (death)	0.000	0.000	0.000	0.000
5	Acute hospital ward, non-invasive ventilation, invasive ventilation (survival)	0.002	0.002	0.002	0.002
6	Acute hospital ward, non-invasive ventilation, invasive ventilation (death)	0.000	0.000	0.000	0.000
7	Acute hospital ward, invasive ventilation (survival)	0.000	0.000	0.000	0.000
8	Acute hospital ward, invasive ventilation (death)	0.000	0.000	0.000	0.000
9	Non-invasive ventilation (survival)	0.348	0.318	0.296	0.351
10	Non-invasive ventilation (death)	0.002	0.002	0.001	0.002
11	Oxygen, invasive ventilation (survival)	0.026	0.024	0.022	0.026
12	Oxygen, invasive ventilation (death)	0.000	0.000	0.000	0.000
13	Invasive ventilation (survival)	0.084	0.077	0.071	0.085
14	Invasive ventilation (death)	0.001	0.000	0.000	0.001

Table 2.9.2: Expected QALYs (No Dexamethasone arm)

Pathway number	Clinical pathway	Expected QALYs (No Dexamethasone arm)			
		45-64 age group	65-74 age group	75+ age group	Population
15	Acute hospital ward (survival)	0.155	0.142	0.132	0.156
16	Acute hospital ward (death)	0.001	0.001	0.001	0.001
17	Acute hospital ward, non-invasive ventilation (survival)	0.006	0.005	0.005	0.006
18	Acute hospital ward, non-invasive ventilation (death)	0.000	0.000	0.000	0.000
19	Acute hospital ward, non-invasive ventilation, invasive ventilation (survival)	0.003	0.003	0.003	0.003
20	Acute hospital ward, non-invasive ventilation, invasive ventilation (death)	0.000	0.000	0.000	0.000
21	Acute hospital ward, invasive ventilation (survival)	0.001	0.000	0.000	0.001
22	Acute hospital ward, invasive ventilation (death)	0.000	0.000	0.000	0.000
23	Non-invasive ventilation (survival)	0.329	0.301	0.280	0.332
24	Non-invasive ventilation (death)	0.002	0.002	0.001	0.002
25	Oxygen, invasive ventilation (survival)	0.029	0.027	0.025	0.029
26	Oxygen, invasive ventilation (death)	0.000	0.000	0.000	0.000
27	Invasive ventilation (survival)	0.069	0.064	0.059	0.070
28	Invasive ventilation (death)	0.001	0.000	0.000	0.001

Table 2.10: Expected costs (Dexamethasone and No Dexamethasone arm)

Pathway number	Clinical pathway	Costs (£)	
		Dexamethasone	No Dexamethasone
1,15	Acute hospital ward (survival)	£1,320.71	£1,336.43
2, 16	Acute hospital ward (death)	£245.94	£196.63
3, 17	Acute hospital ward, non-invasive ventilation (survival)	£88.66	£116.79
4, 18	Acute hospital ward, non-invasive ventilation (death)	£18.63	£27.28
5, 19	Acute hospital ward, non-invasive ventilation, invasive ventilation (survival)	£134.93	£180.56
6, 20	Acute hospital ward, non-invasive ventilation, invasive ventilation (death)	£36.76	£83.89
7, 21	Acute hospital ward, invasive ventilation (survival)	£17.32	£27.13
8, 22	Acute hospital ward, invasive ventilation (death)	£4.69	£12.52
9, 23	Non-invasive ventilation (survival)	£7,580.66	£7,168.53
10, 24	Non-invasive ventilation (death)	£1,636.90	£1,720.39
11, 25	Oxygen, invasive ventilation (survival)	£1,297.27	£1,458.36
12, 26	Oxygen, invasive ventilation (death)	£340.45	£652.71
13, 27	Invasive ventilation (survival)	£4,761.59	£3,945.29
14, 28	Invasive ventilation (death)	£1,273.63	£1,799.68

2.4.1.2. Step 2: Cost-utility analysis of Dexamethasone versus No Dexamethasone

Following the aggregation of each pathway's expected QALYs and costs for both the Dexamethasone and the No Dexamethasone arms, it is possible to undertake a cost-utility analysis of Dexamethasone against No Dexamethasone, stratified by age groups and the population in total.

Dexamethasone is found to be cost-effective relative to No Dexamethasone, with the ICER being £1,236, which is significantly lower than the £20,000 cost-effectiveness threshold recommended by the National Institute for Health and Care Excellence (NICE) (McCabe et al., 2008) and below the £12,936 per QALY of existing NHS treatments according to an empirical analysis by Claxton et al. (2015b). Under the same threshold, the incremental net benefit of Dexamethasone, against No Dexamethasone, is £480.67. Dexamethasone remains cost-effective across all age groups. The summary of the cost-utility analysis of Dexamethasone versus No Dexamethasone, for the population as a whole and by age group, is shown in *Table 2.11*. Note this analysis assumes 100% patients receive Dexamethasone versus 100% patients receiving No Dexamethasone.

Table 2.11: Cost-utility analysis of Dexamethasone versus No Dexamethasone

Treatment Group	**Expected QALYs**	**Expected costs**	**Incremental QALYs**	**Incremental costs**	**Incremental net benefit (£)**	**ICER**
Dexamethasone	0.621	£18,757.84	0.025	£31.66	£476.02	£1,247.27
No Dexamethasone	0.596	£18,726.18				
Patients aged 45-64 years old						
Treatment Group	**Expected QALYs**	**Expected costs**	**Incremental QALYs**	**Incremental costs**	**Incremental net benefit (£)**	**ICER**
Dexamethasone	0.568	£18,757.84	0.023	£31.66	£434.44	£1,358.53
No Dexamethasone	0.544	£18,726.18				
Patients aged 65-74 years old						
Treatment Group	**Expected QALYs**	**Expected costs**	**Incremental QALYs**	**Incremental costs**	**Incremental net benefit (£)**	**ICER**
Dexamethasone	0.529	£18,757.84	0.022	£31.66	£403.12	£1,456.39
No Dexamethasone	0.507	£18,726.18				
Patients aged 75+ years old						
Treatment Group	**Expected QALYs**	**Expected costs**	**Incremental QALYs**	**Incremental costs**	**Incremental net benefit (£)**	**ICER**
Dexamethasone	0.627	£18,757.84	0.026	£31.66	£480.67	£1,235.94
No Dexamethasone	0.601	£18,726.18				
Population (total)						

*All Incremental net benefit and ICER calculations used a cost-effectiveness threshold of £20,000.

2.4.1.3. Step 3: Cost-utility analysis of updated versus old clinical practice

The updated clinical practice following the dissemination of the results from the RECOVERY trial (i.e. where 83% of patients receive Dexamethasone and 17% of patients receive No Dexamethasone) is found to be cost-effective compared with previous clinical practice (i.e. 100% of patients received No Dexamethasone), with the ICER being £1,236, which is significantly lower than the £20,000 cost-effectiveness threshold recommended by NICE (McCabe et al., 2008) and the £12,936 threshold recommended by Claxton et al. (2015b). The ICER figure remains identical to that of the cost-utility analysis of Dexamethasone, versus No Dexamethasone, as 17% of the incremental QALYs and costs associated with Dexamethasone were simultaneously foregone with the updated clinical practice. Using the same threshold, the incremental net benefit of updated clinical practice is £397.83. The updated clinical practice remains cost-effective across all age groups. The summary of the cost-utility analysis of updated clinical practice versus previous clinical practice, for the population as a whole and by age group, is shown in *Table 2.12*.

Table 2.12: Cost-utility analysis of updated clinical practice against previous clinical practice

Clinical practice	**Expected QALYs**	**Expected costs**	**Incremental QALYs**	**Incremental costs**	**Incremental net benefit (£)**	**ICER**
Updated clinical practice	0.617	£18,752.38	0.021	£26.20	£393.98	£1,247.27
Old clinical practice	0.596	£18,726.18				
Patients aged 45-64 years old						
Clinical practice	**Expected QALYs**	**Expected costs**	**Incremental QALYs**	**Incremental costs**	**Incremental net benefit (£)**	**ICER**
Updated clinical practice	0.564	£18,752.38	0.019	£26.20	£359.57	£1,358.53
Old clinical practice	0.545	£18,726.18				
Patients aged 65-74 years old						
Clinical practice	**Expected QALYs**	**Expected costs**	**Incremental QALYs**	**Incremental costs**	**Incremental net benefit (£)**	**ICER**
Updated clinical practice	0.525	£18,752.38	0.018	£26.20	£333.65	£1,456.39
Old clinical practice	0.507	£18,726.18				
Patients aged 75+ years old						
Clinical practice	**Expected QALYs**	**Expected costs**	**Incremental QALYs**	**Incremental costs**	**Incremental net benefit (£)**	**ICER**
Updated clinical practice	0.623	£18,752.38	0.021	£26.20	£397.83	£1,235.94
Old clinical practice	0.601	£18,726.18				
Population (total)						

*All Incremental net benefit and ICER calculations used a cost-effectiveness threshold of £20,000. Updated clinical practice: 83% receive Dexamethasone and 17% receive No Dexamethasone; previous clinical practice: 100% receive No Dexamethasone.

2.4.1.4. Steps 4 and 5: The cost-effectiveness of faster recruitment to the RECOVERY trial

With a 15% recruitment rate, 6,980 UK COVID-19 patients benefited from the updated clinical practice by mid-July 2020 (Knowlson and Torgerson, 2020). Had the recruitment rate been 50%, 77,310 UK COVID-19 patients would have benefited from the updated clinical practice by mid-July 2020 (Knowlson and Torgerson, 2020). Following the aggregation of the figures from *Table 2.12*, achieving a 50% recruitment rate by employing two research nurses at each NHS hospital is found to be a cost-effective strategy compared to the 15% recruitment rate, with the ICER being £10,641, which is below the cost-effectiveness threshold ranges of £20,000 and £30,000 suggested by NICE (McCabe et al., 2008) and the £12,936 threshold recommended by Claxton et al. (2015b). Under the £20,000 cost-effectiveness threshold, the incremental net benefit of faster recruitment to the RECOVERY trial is £13,955,476.42. Improving the recruitment rate remains cost-effective across all age groups. The summary of the cost-utility analysis of 50% recruitment rate against 15% recruitment rate, for the population as a whole and by age group, is shown in *Table 2.13*.

Table 2.13: Cost-utility analysis of faster recruitment to the RECOVERY Trial

Recruitment Status	**Incremental QALYs of updated practice**	**Incremental costs of updated clinical practice**	**Incremental QALYs of faster recruitment**	**Incremental costs of faster recruitment**	**Incremental net benefit of faster recruitment (£)**	**ICER**
50% Recruitment Rate to the RECOVERY Trial	410.364	£4,055,041.46	373.314	£4,008,829.90	£3,457,453.67	£10,738.49
15% Recruitment Rate to the RECOVERY Trial	37.050	£46,211.56				
Patients aged 45-64 years old						
Recruitment Status	**Incremental QALYs of updated practice**	**Incremental costs of updated clinical practice**	**Incremental QALYs of faster recruitment**	**Incremental costs of faster recruitment**	**Incremental net benefit of faster recruitment (£)**	**ICER**
50% Recruitment Rate to the RECOVERY Trial	232.007	£2,497,095.64	211.060	£2,468,638.55	£1,752,561.72	£11,696.38
15% Recruitment Rate to the RECOVERY Trial	20.947	£28,457.09				
Patients aged 65-74 years old						
Recruitment Status	**Incremental QALYs of updated practice**	**Incremental costs of updated clinical practice**	**Incremental QALYs of faster recruitment**	**Incremental costs of faster recruitment**	**Incremental net benefit of faster recruitment (£)**	**ICER**
50% Recruitment Rate to the RECOVERY Trial	456.333	£5,265,326.73	415.1329	£5,205,322.66	£3,097,335.57	£12,538.93
15% Recruitment Rate to the RECOVERY Trial	41.200	£60,004.07				
Patients aged 75+ years old						

Recruitment Status	Incremental QALYs of updated practice	Incremental costs of updated clinical practice	Incremental QALYs of faster recruitment	Incremental costs of faster recruitment	Incremental net benefit of faster recruitment (£)	ICER
50% Recruitment Rate to the RECOVERY Trial	1639.106	£16,049,788.75	1491.118	£15,866,884.13	£13,955,476.42	£10,640.93
15% Recruitment Rate to the RECOVERY Trial	147.988	£182,904.62				
Population (total)						

*All Incremental net benefit and ICER calculations used a cost-effectiveness threshold of £20,000.

2.4.2. Deterministic (one-way and two-way) sensitivity analysis

For the one-way sensitivity analysis, the ranges of the incremental net benefit of faster recruitment to the RECOVERY trial with respect to variations in each of the input parameters from *Table 2.7*, assuming everything else constant, are presented in *Table 2.14*.

With respect to the probability inputs obtained from the original study of the RECOVERY trial (Horby et al., 2021), the incremental net benefit of faster recruitment would remain positive even if the proportion of patients eligible for dexamethasone fell to a low level, that of 43%; in the original study this figure was much higher, at 83%. The findings would remain robust when the probabilities of the any type of usual care a COVID-19 patient initially receives varied. Interestingly, in both arms the estimate of the incremental net benefit is more sensitive to variations in the probabilities of death when the most intensive type of usual care received during hospitalisation is either a stay in an acute hospital ward (no oxygen supply) or non-invasive ventilation (oxygen supply but no mechanical ventilation), rather than to variations in the probabilities of death in both arms when invasive ventilation is the most intensive type of usual care received during hospitalisation for which the incremental net benefit remains positive. The incremental net benefit turns negative when the probability of death in the Dexamethasone (No Dexamethasone) arm following treatment in acute hospital wards is 0.33 and above (0.07 and below); however, the reported probability from the trial is significantly lower (higher), at 0.17 (0.13) (Horby et al., 2021).Similarly, the incremental net benefit turns negative when the probability of death in the Dexamethasone (No Dexamethasone) arm following non-invasive ventilation in an ICU is 0.24 and above (0.19 and below); however, the reported probability from the trial is lower (higher), at 0.21 (0.22) (Horby et al., 2021). The findings are robust with respect to variations in the probability of long COVID following discharge from hospital. Nevertheless, uncertainty as to the incremental net benefit of faster recruitment persisted with respect to variations in the probabilities of a hospitalised patient

being transferred to a more invasive type of care, for all possible (combinations of) types of usual care delivered for patients receiving Dexamethasone during their hospitalisation. However, the corresponding cut-off values remained well above the reported probability inputs from *Table 2.1*.

The findings have remained robust with respect to variations in the ranges of all cost and utility inputs. This implies that the effectiveness of dexamethasone in reducing the incidence of COVID-19 mortality is a key factor behind the cost-effectiveness findings with respect to faster recruitment to the RECOVERY trial.

The findings from the two-way sensitivity analyses are presented in *Table 2.15.1*, *Table 2.15.2*, and *Table 2.15.3*. Regardless of the probability range and the final decision node, when the probability of a COVID-19 inpatient dying remained the same across both treatments, the net benefit of faster recruitment to the RECOVERY trial was positive for two types of usual care received during hospitalisation with COVID-19 (i.e. invasive ventilation (when probability of death is equal or above 0.10) and acute hospital ward) assuming everything else remains constant. Therefore, even if Dexamethasone had the same clinical effectiveness on reducing the incidence of COVID-19-related mortality as No Dexamethasone after mechanical ventilation (when the probability of death is equal to or above 0.10) or stay in acute hospital wards, improving patient recruitment would still generate positive incremental net benefits, assuming everything else constant. Nevertheless, the same conclusion cannot be drawn with respect to the probability of death following non-invasive ventilation, as the incremental net benefit of faster recruitment would remain negative if dexamethasone was not effective in reducing the incidence of COVID-19-related mortality.

Table 2.14: One-way sensitivity analysis of the Incremental Net Benefit of faster recruitment to the RECOVERY Trial

Input	Source	Range for one-way sensitivity analysis	Range of INB>0	Input range	Mean value
Probability inputs					
Proportion of patients who can have dexamethasone	(Horby et al., 2021)	(-£14,023,949.72, £19,781,657.30)	(0.42,1)	(0,1)	0.8277
P (acute hospital ward)	(Horby et al., 2021)	(£12,154,661.16, £19,692,275.53)	(0,1)	(0,1)	0.2389
P (non-invasive ventilation)	(Horby et al., 2021)	(-£12,497,944.39, £31,273,167.82)	(0.29,1)	(0,1)	0.6044
P (invasive ventilation)	(Horby et al., 2021)	(£13,704,272.31, £14,230,286.36)	(0,1)	(0,1)	0.1567
P (death|acute hospital ward, Dexamethasone)	(Horby et al., 2021)	(-£53,136,227.60, £27,277,513.35)	(0, 0.33)	(0,1)	0.1657
P (non-invasive ventilation|acute hospital ward, Dexamethasone)	(Horby et al., 2021)	(-£44,893,838.32, £15,896,897.11)	(0, 0.26)	(0,1)	0.0319
P (invasive ventilation|acute hospital ward, Dexamethasone)	(Horby et al., 2021)	(-£384,072,187.05, £21,236,470.26)	(0, 0.05)	(0,1)	0.0180
P (death|acute hospital ward, No Dexamethasone)	(Horby et al., 2021)	(-£14,091,727.31, £197,593,007.38)	(0.07,1)	(0,1)	0.1325
P (non-invasive ventilation|acute hospital ward, No Dexamethasone)	(Horby et al., 2021)	(£4,883,303.71, £208,809,968.45)	(0,1)	(0,1)	0.0445
P (invasive ventilation|acute hospital ward, No Dexamethasone)	(Horby et al., 2021)	(-£1,597,023.47, £534,445,805.91)	(0.01,1)	(0,1)	0.0290
P (death| non-invasive ventilation, Dexamethasone)	(Horby et al., 2021)	(-£310,956,948.52, £98,464,717.98)	(0,0.24)	(0,1)	0.2064
P (invasive ventilation|non-invasive ventilation, Dexamethasone)	(Horby et al., 2021)	(-£710,112,659.75, £76,036,021.20)	(0,0.0.09)	(0,1)	0.0790
P (death| non-invasive ventilation, No Dexamethasone)	(Horby et al., 2021)	(-£75,314,635.50, £336,117,880.33)	(0.19,1)	(0,1)	0.2170
P (invasive ventilation|non-invasive ventilation, No Dexamethasone)	(Horby et al., 2021)	(-£71,191,645.91, £723,514,829.16)	(0.09,1)	(0,1)	0.1071
P (death| invasive ventilation, Dexamethasone)	(Horby et al., 2021)	(£13,618,182.74, £14,095,401.74)	(0,1)	(0,1)	0.2932
P (death| invasive ventilation, No Dexamethasone)	(Horby et al., 2021)	(£13,378,645.55, £14,770,785.07)	(0,1)	(0,1)	0.4143
P (long COVID; population total)	(Taquet et al., 2021)	(£13,947,184.74, £13,963,464.80)	(0.1256, 0.1309)	(0.1256, 0.1309)	0.1282
P (long COVID; ICU all)	(Taquet et al., 2021)	(£13,930,641.09, £13,980,690.89)	(0.249, 0.2754)	(0.249, 0.2754)	0.2623
Utility inputs					
Disutility of COVID-19 Infection	(Sheinson et al., 2021, Vekaria et al., 2021)	(£13,861,844.96, £13,967,048.85)	(0, 0.95)	(0.0, 0.95)	0.27
Disutility of Hospitalisation (Acute ward, no ventilation received)	(Sheinson et al., 2021, Vekaria et al., 2021)	(£13,893,405.95, £13,992,718.70)	(0,1)	(0.0, 1.0)	0.11
Disutility of Hospitalisation (Non-invasive ventilation received)	(Sheinson et al., 2021, Vekaria et al., 2021)	(£13,916,162.97, £13,987,372.23)	(0.0, 0.96)	(0.0, 0.96)	0.36
Disutility of Hospitalisation (Invasive ventilation received)	(Sheinson et al., 2021, Vekaria et al., 2021)	(£13,775,105.11, £14,135,847.72)	(0.03, 0.99)	(0.03, 0.99)	0.56
Disutility of long COVID	(Szende et al., 2014, Hvidberg et al., 2015)	(£13,943,060.13, £13,964,187.96)	(0.23,0.35)	(0.23, 0.35)	0.29
Cost inputs					
Unit cost of GP appointment (long COVID patients)	(Curtis and Burns 2020)	(£13,943,060.13, 13,964,187.97)	(£21.5, £64.5)	(£21.5, £64.5)	£39.23
Daily cost of staying in acute hospital ward (no oxygen)	Manual, (Department of Health and Social Care, 2019)	(£12,991,778.60, £14,432,743.06)	(£500, £1250)	(£500, £1250)	£748.41
Daily cost of non-invasive ventilation (oxygen)	Manual, (Department of Health and Social Care, 2019)	(£13,078,867.38, £15,078,137.26)	(£800, £1800)	(£800, £1800)	£1,394.23
Daily cost of invasive ventilation	Manual, (Department of Health and Social Care, 2019)	(£13,476,011.57, £18,081,804.39)	(£800, £4,800)	(£800, £4,800)	£1,753.94
Annual salary cost of a research nurse	(National Health Service, 2020)	(£12,506,210.14, £15,340,162.14)	(£35097, £43958)	(£35097, £43958)	£39,840.77

Table 2.15.1: Two-way sensitivity analysis of the Incremental Net Benefit of faster recruitment to the RECOVERY Trial (with respect to the risk of COVID-19-related death following invasive ventilation)

P (death| invasive ventilation, Dexamethasone)							
P (death| invasive ventilation, No Dexamethasone)		**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
	0	-£769,690.39	£20,398,783.08	£41,567,256.55	£62,735,730.02	£83,904,203.49	£105,072,676.96
	0.1	-£8,811,064.48	£12,357,408.99	£33,525,882.46	£54,694,355.93	£75,862,829.40	£97,031,302.87
	0.2	-£16,852,438.57	£4,316,034.90	£25,484,508.36	£46,652,981.83	£67,821,455.30	£88,989,928.77
	0.3	-£24,893,812.67	-£3,725,339.20	£17,443,134.27	£38,611,607.74	£59,780,081.21	£80,948,554.68
	0.4	-£32,935,186.76	-£11,766,713.29	£9,401,760.18	£30,570,233.64	£51,738,707.11	£72,907,180.58
	0.5	-£40,976,560.86	-£19,808,087.39	£1,360,386.08	£22,528,859.55	£43,697,333.02	£64,865,806.49
	0.6	-£49,017,934.95	-£27,849,461.48	-£6,680,988.01	£14,487,485.46	£35,655,958.93	£56,824,432.39
	0.7	-£57,059,309.05	-£35,890,835.58	-£14,722,362.11	£6,446,111.36	£27,614,584.83	£48,783,058.30
	0.8	-£65,100,683.14	-£43,932,209.67	-£22,763,736.20	-£1,595,262.73	£19,573,210.74	£40,741,684.21
	0.9	-£73,142,057.23	-£51,973,583.76	-£30,805,110.30	-£9,636,636.83	£11,531,836.64	£32,700,310.11
	1	-£81,183,431.33	-£60,014,957.86	-£38,846,484.39	-£17,678,010.92	£3,490,462.55	£24,658,936.02

P (death| invasive ventilation, No Dexamethasone)		**0.6**	**0.7**	**0.8**	**0.9**	**1**
	0	£126,241,150.43	£147,409,623.90	£168,578,097.37	£189,746,570.84	£210,915,044.31
	0.1	£118,199,776.33	£139,368,249.80	£160,536,723.27	£181,705,196.74	£202,873,670.21
	0.2	£110,158,402.24	£131,326,875.71	£152,495,349.18	£173,663,822.65	£194,832,296.12
	0.3	£102,117,028.15	£123,285,501.62	£144,453,975.08	£165,622,448.55	£186,790,922.02
	0.4	£94,075,654.05	£115,244,127.52	£136,412,600.99	£157,581,074.46	£178,749,547.93
	0.5	£86,034,279.96	£107,202,753.43	£128,371,226.90	£149,539,700.36	£170,708,173.83
	0.6	£77,992,905.86	£99,161,379.33	£120,329,852.80	£141,498,326.27	£162,666,799.74
	0.7	£69,951,531.77	£91,120,005.24	£112,288,478.71	£133,456,952.18	£154,625,425.65
	0.8	£61,910,157.67	£83,078,631.14	£104,247,104.61	£125,415,578.08	£146,584,051.55
	0.9	£53,868,783.58	£75,037,257.05	£96,205,730.52	£117,374,203.99	£138,542,677.46
	1	£45,827,409.49	£66,995,882.96	£88,164,356.42	£109,332,829.89	£130,501,303.36
P (death| invasive ventilation, Dexamethasone)						

Table 2.15.2: Two-way sensitivity analysis of the Incremental Net Benefit of faster recruitment to the RECOVERY Trial (with respect to the risk of COVID-19-related death following non-invasive ventilation)

P (death| non-invasive ventilation, Dexamethasone)							
P (death| non-invasive ventilation, No Dexamethasone)		**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
	0	-£9,111,101.13	£70,359,546.38	£149,830,193.88	£229,300,841.39	£308,771,488.90	£388,242,136.40
	0.1	-£87,725,969.22	-£8,255,321.72	£71,215,325.79	£150,685,973.30	£230,156,620.80	£309,627,268.31
	0.2	-£166,340,837.32	-£86,870,189.81	-£7,399,542.31	£72,071,105.20	£151,541,752.71	£231,012,400.22
	0.3	-£244,955,705.41	-£165,485,057.91	-£86,014,410.40	-£6,543,762.89	£72,926,884.61	£152,397,532.12
	0.4	-£323,570,573.51	-£244,099,926.00	-£164,629,278.49	-£85,158,630.99	-£5,687,983.48	£73,782,664.03
	0.5	-£402,185,441.60	-£322,714,794.10	-£243,244,146.59	-£163,773,499.08	-£84,302,851.58	-£4,832,204.07
	0.6	-£480,800,309.70	-£401,329,662.19	-£321,859,014.68	-£242,388,367.18	-£162,917,719.67	-£83,447,072.16
	0.7	-£559,415,177.79	-£479,944,530.29	-£400,473,882.78	-£321,003,235.27	-£241,532,587.77	-£162,061,940.26
	0.8	-£638,030,045.89	-£558,559,398.38	-£479,088,750.87	-£399,618,103.37	-£320,147,455.86	-£240,676,808.35
	0.9	-£716,644,913.98	-£637,174,266.48	-£557,703,618.97	-£478,232,971.46	-£398,762,323.96	-£319,291,676.45
	1	-£795,259,782.08	-£715,789,134.57	-£636,318,487.06	-£556,847,839.56	-£477,377,192.05	-£397,906,544.54

P (death| non-invasive ventilation, No Dexamethasone))		**0.6**	**0.7**	**0.8**	**0.9**	**1**
	0	£467,712,783.91	£547,183,431.42	£626,654,078.92	£706,124,726.43	£785,595,373.94
	0.1	£389,097,915.82	£468,568,563.32	£548,039,210.83	£627,509,858.34	£706,980,505.84
	0.2	£310,483,047.72	£389,953,695.23	£469,424,342.74	£548,894,990.24	£628,365,637.75
	0.3	£231,868,179.63	£311,338,827.13	£390,809,474.64	£470,280,122.15	£549,750,769.65
	0.4	£153,253,311.53	£232,723,959.04	£312,194,606.55	£391,665,254.05	£471,135,901.56
	0.5	£74,638,443.44	£154,109,090.94	£233,579,738.45	£313,050,385.96	£392,521,033.46
	0.6	-£3,976,424.66	£75,494,222.85	£154,964,870.36	£234,435,517.86	£313,906,165.37
	0.7	-£82,591,292.75	-£3,120,645.25	£76,350,002.26	£155,820,649.77	£235,291,297.27
	0.8	-£161,206,160.85	-£81,735,513.34	-£2,264,865.83	£77,205,781.67	£156,676,429.18
	0.9	-£239,821,028.94	-£160,350,381.44	-£80,879,733.93	-£1,409,086.42	£78,061,561.09
	1	-£318,435,897.04	-£238,965,249.53	-£159,494,602.02	-£80,023,954.52	-£553,307.01
P (death| non-invasive ventilation, Dexamethasone)						

Table 2.15.3: Two-way sensitivity analysis of the Incremental Net Benefit of faster recruitment to the RECOVERY Trial (with respect to the risk of COVID-19-related death following admission to an acute hospital ward)

P (death| acute hospital ward, Dexamethasone)							
P(death| acute hospital ward, No Dexamethasone))		**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
	0	£36,870,920.10	£72,178,949.77	£107,486,979.43	£142,795,009.10	£178,103,038.76	£213,411,068.42
	0.1	£4,372,239.86	£39,680,269.52	£74,988,299.19	£110,296,328.85	£145,604,358.52	£180,912,388.18
	0.2	-£28,126,440.39	£7,181,589.28	£42,489,618.94	£77,797,648.61	£113,105,678.27	£148,413,707.94
	0.3	-£60,625,120.63	-£25,317,090.97	£9,990,938.70	£45,298,968.36	£80,606,998.03	£115,915,027.69
	0.4	-£93,123,800.87	-£57,815,771.21	-£22,507,741.54	£12,800,288.12	£48,108,317.78	£83,416,347.45
	0.5	-£125,622,481.12	-£90,314,451.45	-£55,006,421.79	-£19,698,392.12	£15,609,637.54	£50,917,667.21
	0.6	-£158,121,161.36	-£122,813,131.70	-£87,505,102.03	-£52,197,072.37	-£16,889,042.70	£18,418,986.96
	0.7	-£190,619,841.60	-£155,311,811.94	-£120,003,782.28	-£84,695,752.61	-£49,387,722.95	-£14,079,693.28
	0.8	-£223,118,521.85	-£187,810,492.18	-£152,502,462.52	-£117,194,432.86	-£81,886,403.19	-£46,578,373.53
	0.9	-£255,617,202.09	-£220,309,172.43	-£185,001,142.76	-£149,693,113.10	-£114,385,083.43	-£79,077,053.77
	1	-£288,115,882.34	-£252,807,852.67	-£217,499,823.01	-£182,191,793.34	-£146,883,763.68	-£111,575,734.01

P(death| acute hospital ward, No Dexamethasone))		**0.6**	**0.7**	**0.8**	**0.9**	**1**
	0	£248,719,098.09	£284,027,127.75	£319,335,157.42	£354,643,187.08	£389,951,216.75
	0.1	£216,220,417.85	£251,528,447.51	£286,836,477.17	£322,144,506.84	£357,452,536.50
	0.2	£183,721,737.60	£219,029,767.27	£254,337,796.93	£289,645,826.60	£324,953,856.26
	0.3	£151,223,057.36	£186,531,087.02	£221,839,116.69	£257,147,146.35	£292,455,176.02
	0.4	£118,724,377.11	£154,032,406.78	£189,340,436.44	£224,648,466.11	£259,956,495.77
	0.5	£86,225,696.87	£121,533,726.53	£156,841,756.20	£192,149,785.86	£227,457,815.53
	0.6	£53,727,016.63	£89,035,046.29	£124,343,075.96	£159,651,105.62	£194,959,135.28
	0.7	£21,228,336.38	£56,536,366.05	£91,844,395.71	£127,152,425.38	£162,460,455.04
	0.8	-£11,270,343.86	£24,037,685.80	£59,345,715.47	£94,653,745.13	£129,961,774.80
	0.9	-£43,769,024.11	-£8,460,994.44	£26,847,035.22	£62,155,064.89	£97,463,094.55
	1	-£76,267,704.35	-£40,959,674.68	-£5,651,645.02	£29,656,384.64	£64,964,414.31
P (death| acute hospital ward, Dexamethasone)						

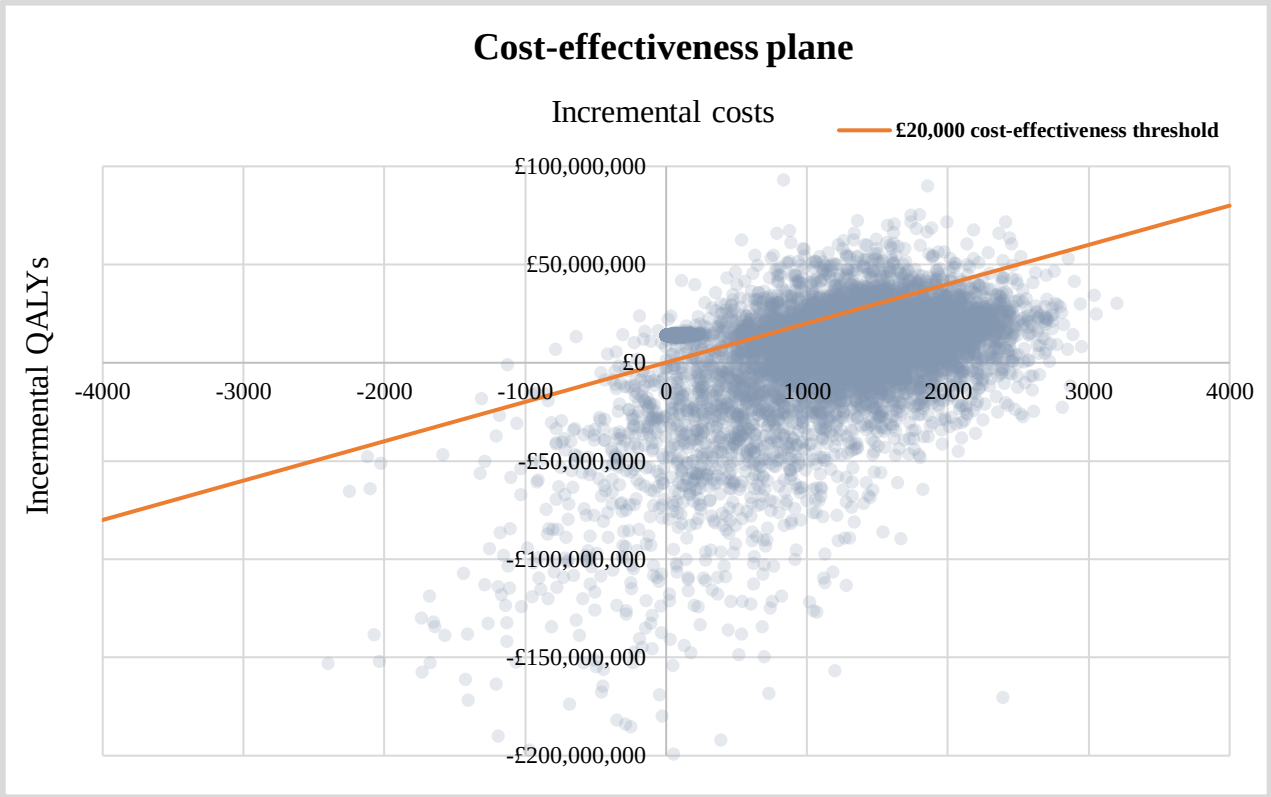
2.4.3. Probabilistic sensitivity analysis (PSA)

Following 10,000 random iterations on the distributions of all input parameters, as specified from *Table 2.8*, distributions of the incremental QALYs and costs related to faster recruitment to the RECOVERY trial were generated, both of which led to the generation of distributions of the incremental net benefit at the £20,000 cost-effectiveness threshold, and the ICER.

The generated 10,000 combinations of incremental QALYs and incremental costs are visualised in *Figure 2.2*, which represents a cost-effectiveness plane. Combinations of incremental QALYs and incremental costs on the top left quadrant would indicate that faster recruitment to the RECOVERY trial would strictly be dominated by prior recruitment, as the incremental QALYs associated with faster recruitment would be negative and the incremental costs would be positive. Combinations of incremental QALYs and incremental costs on the bottom right quadrant would indicate that faster recruitment to the RECOVERY trial would strictly dominate prior recruitment, as the incremental QALYs associated with faster recruitment would be positive but also the incremental costs would be negative. Combinations of incremental QALYs and incremental costs on the bottom left and top right quadrants,

however, require the consideration of a cost-effectiveness threshold, because in the former case faster recruitment would be associated with less QALYs but also with less costs, whereas in the latter case faster recruitment would be associated with more QALYs but also with more costs; in other words, no intervention or strategy strictly dominates or is strictly dominated by another one. For these quadrants, any combinations of incremental QALYs and incremental costs positioned on the right-hand side of the gradient (£20,000 threshold), would imply that their ICER was less than the cost-effectiveness threshold and therefore faster recruitment to the trial would dominate prior recruitment. On the contrary, any combinations of incremental QALYs and incremental costs positioned on the left-hand side of the gradient (£20,000 threshold), would imply that their ICER was less than the cost-effectiveness threshold and therefore faster recruitment would be weakly dominated by prior recruitment to the trial. As the figure demonstrates, the vast majority of the random combinations of incremental QALYs and incremental costs lie to the right-hand side of the gradient on the top right and bottom left quadrants, with combinations also observed on the bottom right quadrant that imply dominance for faster recruitment. Therefore, by considering the overall parameter uncertainty related to the decision model for estimating the cost-effectiveness of faster recruitment, it is highly likely that increasing recruitment to the RECOVERY trial by hiring or redeploying research nurses would be a cost-effective strategy, compared to not increasing recruitment to the trial.

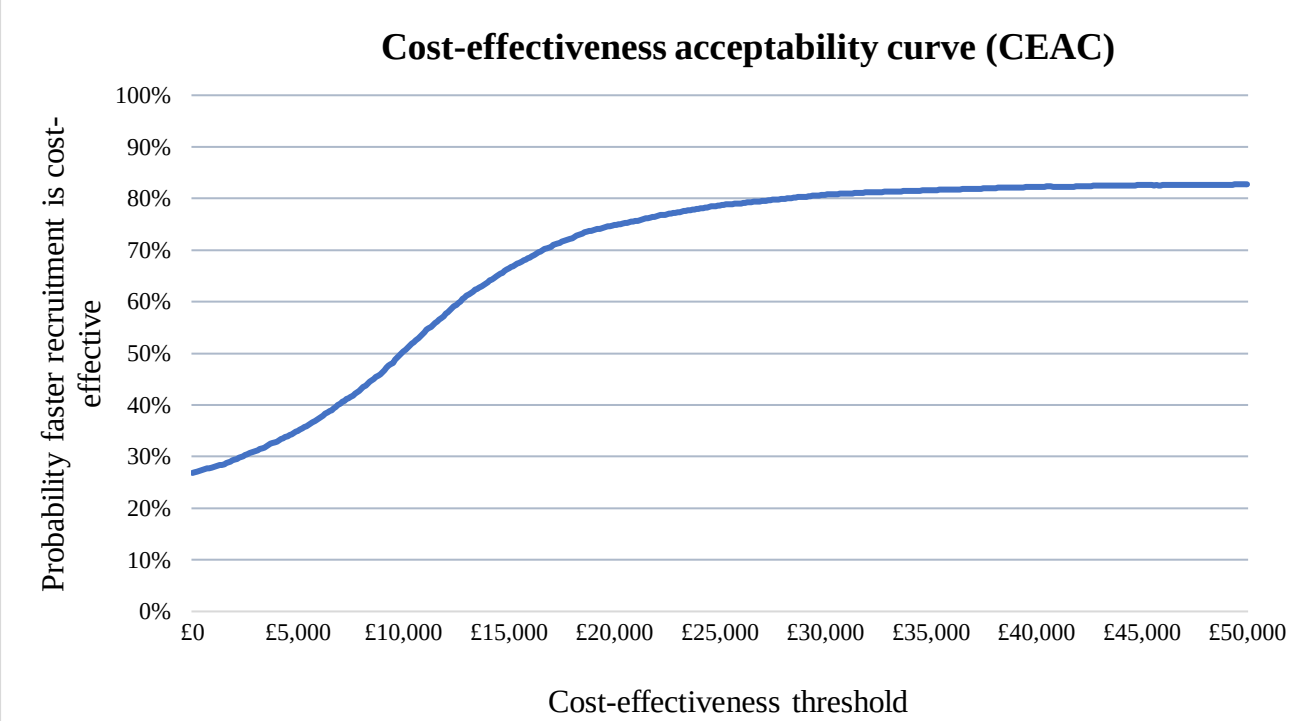
Figure 2.2: Cost-effectiveness plane



The mean of the distribution of the incremental net benefit is equal to the estimated incremental net benefit figure of £13,955,476.42, as reported in *Table 2.13*. Under the assumption of normality, it was also possible to generate the 95% confidence intervals (CI) of the incremental net benefit to faster recruitment to the RECOVERY trial from its generated distribution. Thus, by considering all the distributions of input parameters and accounting for parameter uncertainty, the incremental net benefit of faster recruitment to the RECOVERY trial was estimated to be: £13,955,476.42 (95% CI: £12,457,048.54, £15,453,904.30). Such an estimate confirms that the model is robust with respect to parameter uncertainty and raises the confidence in the statistical validity and significance of the baseline cost-effectiveness findings. Under the same distribution, it was also possible to estimate the probability of faster recruitment being a cost-effective strategy, by counting the proportion of the incremental net benefits that were above zero under the 10000 iterations. At the £20,000 cost-effectiveness threshold, this probability was 0.75, which also means that 75% of the combinations of incremental QALYs and incremental costs shown on the cost-effectiveness plane in *Figure 2.2* are positioned on the right-hand side of the threshold gradient. It was also possible to vary the cost-effectiveness thresholds, to observe the probability of faster recruitment being cost-effective under different thresholds and following the 10000 iterations. At the £30,000 cost-effectiveness threshold, which is the upper limit that NICE recommends for the evaluation of

health technologies (McCabe et al., 2008), the probability increases to 81%. At the £12,936 threshold (Claxton et al. (2015b)), the probability falls to 61%, which remains a significant figure. The combinations of the probabilities of faster recruitment being cost-effective and cost-effectiveness thresholds are graphically represented through a cost-effectiveness acceptability curve (CEAC), available in *Figure 2.3*.

Figure 2.3: Cost-effectiveness acceptability curve (CEAC)



2.5. Discussion

2.5.1. Summary of findings

The chapter's findings highlight the importance of improving recruitment of patients to RCTs, from the economic perspective of a healthcare system, by considering a noticeable case study related to the COVID-19 pandemic. By the time the results of the RECOVERY trial were originally disseminated, 39,961 deaths and 122,638 hospitalisations (UK Health Security Agency, 2022) from COVID-19 had already been reported across the UK, thus signalling the detrimental impact of the COVID-19 pandemic on public health and the provision of healthcare services. Therefore, it is crucial to ensure that cutting edge adaptive trials, such as the RECOVERY trial, reach their full potential through faster recruitment so as to be beneficial for all affected patients, the healthcare system, society overall and the national economy.

Improving the recruitment rate of the RECOVERY trial from 15% to 50% would have saved at least 2,620 more lives, in the UK, by mid-July 2020 (Knowlson and Torgerson, 2020). In addition, this improvement could have generated an incremental net monetary benefit of £14 million, thus highlighting the magnitude of the foregone population health benefits due to the absence of implementing a more effective recruitment strategy in the beginning of a trial. If employing two research nurses to each involved hospital were to constitute such an effective strategy, less than £10,641 would need to be invested to generate an incremental QALY for COVID-19 hospitalised patients, a figure significantly below the minimum value of the cost-effectiveness threshold NICE recommends, i.e. £20,000 (McCabe et al., 2008), and below the lower cost-effectiveness threshold of £12,936 recommended by Claxton et al. (2015b). As the NHS is a fixed budget national healthcare system, it is essential for policy makers to prioritise research and resources into improving the recruitment of participants to future RCTs that evaluate healthcare treatments or public health interventions.

The principal method of sensitivity analysis, the probabilistic sensitivity analysis (PSA), which considers simultaneously the ranges of all parameter inputs to estimate the 95% confidence interval of the incremental net benefit, has found the incremental net benefit of faster recruitment to be £13,955,476.42 (95% CI: £12,457,048.54, £15,453,904.30). This means that improving recruitment to the RECOVERY trial would have been a statistically significant, cost-effective strategy, accounting for parameter uncertainty. Also, the cost-effectiveness acceptability curve (CEAC) shows that the likelihood of faster recruitment being a cost-effective strategy is 75% and 81% under the £20,000 and £30,000 cost-effectiveness thresholds respectively ; these are the thresholds NICE recommends for use for the economic evaluation of health technologies (McCabe et al., 2008). From a different angle, the one-way sensitivity analyses confirm that even if a very low share of COVID-19 inpatients were treated with Dexamethasone, increasing recruitment to the RECOVERY trial would still generate positive incremental net benefits. Such findings highlight benefits that improved trial efficiency could have generated for the healthcare system during the COVID-19 pandemic and demonstrate the robustness of this decision tree with respect to variations in its input parameter values.

2.5.2. Strengths and limitations of the study

This is the first study that evaluates how costly poor patient recruitment could be for a healthcare system, by using conventional economic evaluation methods such as a decision tree

model. The study attempts to estimate such costs in the context of a global pandemic, which caused significant mortality and deterioration in the health-related quality of life for patients exposed to Sars-Cov-2 virus. The study also highlights how crucial clinical evidence, as well as its timely dissemination, are for improving health outcomes for national populations and making the treatment decisions of national healthcare systems more cost-effective. Whereas the impact of poor recruitment is usually discussed for statistical reasons such as reduced power and selection bias, or financial reasons related to a potentially extended trial period, this study demonstrates that the adoption of appropriate recruitment strategies would be a cost-effective strategy for national healthcare systems. In fact, it is highly likely that the chapter's estimate of the net benefit of faster recruitment to the RECOVERY trial is underestimated, since it is based solely upon the UK hospitalisation and death data. As the COVID-19 pandemic was significantly affecting the operations of multiple healthcare systems around the globe from March 2020 onwards, a lot more than 2,620 lives would have been saved globally had the results of the RECOVERY trial been disseminated in April rather than mid-June and had all countries updated their clinical practices accordingly.

Improving recruitment to randomised trials can be shown to be cost-effective for a national healthcare system in three distinct scenarios. Firstly, since the use of dexamethasone as a treatment for COVID-19 patients leads to an increase in QALYs, more QALYs could have been gained had its implementation been faster through faster recruitment to the RECOVERY trial. Alternatively, if a treatment under evaluation was deemed ineffective, there would be an indirect QALY gain from the earlier identification of an alternative, efficacious treatment. This is facilitated via faster recruitment to trials assessing the (cost-) effectiveness of the former treatment, leading to its premature cessation. Finally, if a treatment under evaluation was found to be harmful, an overall decrease in QALYs would occur. In this case, there could be a potential cost gain, as well as a direct QALY gain from the avoidance of a prolonged use of such a treatment. For instance, the RECOVERY trial observed an increase in mortality (albeit not statistically significant) among patients randomised to the hydroxychloroquine arm, which would have led to a reduction in the worldwide prescription of this treatment for COVID-19; thus, hydroxychloroquine falls into the last scenario. The current analysis did not take these additional benefits into account. Thus, speeding up knowledge about which treatments are ineffective is also important as many of these are costly (e.g., plasma) as well as having no clinical value.

The disadvantage of the study lies in the model's structural assumptions, which had to be made because of the statistical uncertainty associated with the pandemic. For instance, it is assumed that no recovered COVID-19 patient would catch the SARS-Cov-2 infection within a year. Given the observed re-emergence in COVID-19 infections as a result of the immunity-resistant variants, this assumption seems not to be realistic *ex post*. However, the trajectory of an ongoing pandemic remains uncertain, and each assumption has its own expected benefits and risks. Another assumption is that survived ventilated patients would fully recover outside of hospital; whereas this is true for some patients, many would remain in acute hospital beds for a smooth recovery. However, neither the findings from the RECOVERY trial nor from any other COVID-related study, did not provide any information on the likelihood and the length of stay of survived ventilated COVID-19 patients in acute hospital beds. A further assumption was around the costs incurred by long COVID patients, namely that they would incur two GP visits on average; this could be an over-estimate, or indeed an under-estimate, of the costs encountered.

Despite the strength of the findings, employing additional research nurses is not an evidence-based recruitment strategy; the most recent Cochrane review of recruitment strategies to RCTs has not included any studies associated with this strategy (Treweek et al., 2018b). Theoretically, however, research nurses could actively support hospitalised COVID-19 patients during their participation in the RECOVERY trial and effectively communicate the benefits and risks of participating in the trial. Assuming it would take an hour on average to speak to each patient about the study, it would be feasible for each research nurse to recruit up to six patients per day, which implies it would take approximately 10.5 days (assuming five working days a week) to reach the desirable recruitment figure of 11,303 patients. Therefore, it would be feasible for the proposed recruitment strategy to have accelerated the recruitment rate from 15% to 50%. Such support would be valuable given the RECOVERY chief investigator's remarks, which related slow recruitment to the poor willingness of some patients to enter a trial, and the lack of promotion of the RECOVERY trial to some patients (Wise and Coombes, 2020). Furthermore, an increase in the recruitment rate of 15% to 50% can be realistic, as several NHS hospitals had already achieved overall recruitment rates above 50% (Wise and Coombes, 2020); recruiting research nurses could have significantly improved the recruitment performance of the worst-affected hospitals. Moreover, the cost attached to such a recruitment strategy may have been strict, as it was assumed that the NHS would incur the costs of hiring or redeploying such research nurses for an entire year. Thus, the cost-

effectiveness estimates of faster recruitment to the RECOVERY are likely to be conservative. However, this approach also signals that if there was a significantly effective recruitment strategy, which would also be considerably costly, accelerating recruitment to the RECOVERY trial would have remained a cost-effective strategy from the NHS perspective. A disadvantage of hiring or redeploying research nurses for recruitment purposes is the concern that, at the onset of a global pandemic, it might not have been feasible to employ 352 research nurses across the country. However, as many elective operations and primary care appointments were cancelled due to the emergence of the pandemic, it may also have been feasible for such a recruitment strategy to have been implemented feasibly due to potential redirection of duties for research nurses. Interestingly, a senior NHS research nurse reported that during the COVID-19 pandemic, the intensive care units (ICUs) were adequately staffed, implying that research nurses could swift from clinical duties to research responsibilities (Great Ormond Street Hospital, 2021), an argument that supports the feasibility of considering research nurses as a potential recruitment strategy in the RECOVERY trial.

2.5.3. Direction for future research

The model of the cost-effectiveness of improving patient recruitment to a real-life RCT was based upon an *ex-post* analysis where dissemination of a cost-effective treatment was delayed because of slow recruitment. Such an analysis was undertaken to demonstrate that the detrimental effects of poor recruitment to trials are not limited to trial teams and statisticians but can also impact an entire healthcare system. Alternatively, if a treatment was not cost-effective, as was the case with hydroxychloroquine in the RECOVERY trial, faster recruitment could still be cost-effective but would have to be modelled alternatively by considering aspects such as: QALYs saved from rejecting earlier a harmful treatment and switching research to an alternative, cost-effective treatment; resource use costs saved from finishing the trial earlier than originally. A useful tool, which could determine *ex ante* whether improving recruitment is cost-effective, could be the Expected Value of Perfect Information (EVPI) or the Expected Value of Sample Information (EVSI), the value of which could be subtracted by the incremental costs of a recommended recruitment strategy to generate the *ex-ante* incremental net benefit of faster recruitment to a corresponding RCT. In the case of the RECOVERY trial, simulations of different effectiveness scenarios of Treatments A and B could be undertaken *ex ante* to estimate (and update) the EVPI in real time; the EVPI could then be subtracted by the incremental cost of hiring or redeploying two research nurses to each hospital in order to

estimate the *ex-ante* incremental net benefit of faster recruitment to the RECOVERY trial. There is no study that has estimated *ex ante* the cost-effectiveness of accelerating patient recruitment to a randomised trial.

With respect to improving patient recruitment to randomised trials, the conduct of Studies Within A Trial (SWATs) is highly suggested, in order for effective and cost-effective patient recruitment strategies to be identified. According to Trial Forge Guidance 1, a SWAT is a *"self-contained research study that has been embedded within a host trial with the aim of evaluating or exploring alternative ways of delivering or organising a particular trial process."* (Treweek et al., 2018a). A Cochrane review, which critically appraised the evidence surrounding the effectiveness of existing recruitment strategies, included studies whose study design was predominantly identical or similar to that of a SWAT (Treweek et al., 2018b).

2.5.4. Concluding remarks

This is the first study that presents, using conventional economic evaluation methods, the importance of achieving sufficient recruitment rates in randomised trials, from the economic perspective of a national healthcare system. In the case of the RECOVERY trial, a randomised study related to the treatment of hospitalised patients with COVID-19, improving recruitment to the RECOVERY trial could have generated an incremental net benefit of £13,955,476.42 (95% CI: £12,457,048.54, £15,453,904.30). Thus, it is imperative that effective and cost-effective recruitment strategies be identified, so that national healthcare systems could benefit from faster and improved patient recruitment to randomised trials. The conduct of SWATs for novel and existing recruitment strategies could positively contribute to the aim of improving trial efficiency, with benefits for patients, trial teams and national healthcare systems.

Chapter 3: Economic costs of participant attrition from RCTs: a case study from the OTIS trial

3.1. Abstract

Background: Attrition of participants can threaten the statistical validity of randomised trials. In addition, it can have financial consequences for trial teams and their funders. This chapter explores the Occupational Therapist Intervention Study (OTIS) where, despite a withdrawal rate of less than 10%, the trial team incurred opportunity costs related to participants who were initially recruited but subsequently decided to withdraw from the trial.

Aim: To estimate the cost of participant losses to follow-up in the OTIS trial and thus introduce a costing framework to research teams on how they could estimate the direct financial costs of attrition in their randomised trials.

Methods: The participants lost to follow-up were differentiated by: 1) the time point at which they were lost to follow-up; 2) the treatment group they were allocated to; 3) their response patters to follow-up questionnaires These factors were used to produce the relevant types and subtypes of attrition; the numbers of participants belonging to each type of attrition were obtained by the trial team. Data regarding protocol-driven costs of each trial materials were collected from the trial team, including administration, print and shipping costs. The trial materials corresponding to each type of attrition were determined, following which unit costs by type of attrition were estimated. For each type of attrition, the unit costs were multiplied by the corresponding number of participants to obtain the aggregate figures. Following the summation of the aggregate cost figures by type of attrition, it became possible to obtain the aggregate and the average cost of participant attrition to the trial team.

Results: The average cost per participant loss to follow-up in the OTIS trial is £98.41. The aggregate cost of participant loss to follow-up to the trial team, and hence to the trial funder, is £10,234.90.

Conclusion: Despite the low attrition rate of the OTIS trial, which is lower than the median attrition rate usually observed in randomised trials, loss to follow-up has still generated considerable financial costs for the trial team. It is strongly recommended that decision makers focus on identifying strategies which could improve participant retention in randomised trials.

3.2. Introduction

3.2.1. Participant loss to follow-up in randomised trials

Attrition remains one of the key challenges to the statistical validity of randomised controlled trials (RCTs) (Leon et al., 2006), which naturally benefit from the feature of randomisation holding other things constant (Hariton and Locascio, 2018). Attrition occurs when already recruited participants withdraw from a trial during follow-up, die during the study or are lost to follow-up while they receive the intervention they were randomly allocated thus generating missing data for trial teams (Leon et al., 2006). This phenomenon is capable of introducing selection bias if the dropout rate is differential by trial group (e.g. dependent upon the intervention or treatment received) and/or participants' baseline characteristics (e.g. health status, age, sex, etc.), thus introducing uncertainty in a trial's findings and its dissemination (Leon et al., 2006). Thus, attrition may reduce the external validity of the study's results, as the sample size can be skewed so that the final sample differs significantly from the original sample in terms of baseline characteristics (Torgerson and Torgerson, 2008). It is also likely that the patient population be indirectly affected by trial attrition as the treatment they will be receiving after the dissemination of the findings may not necessarily improve their health outcomes as much as other available treatments could, as the treatment effects of proposed (rejected) interventions or treatments may have been wrongly estimated (Akl et al., 2012).

Loss to follow-up in a trial is also expected to have economic consequences for trial teams, clinical and health research funders, and healthcare systems overall. With respect to trial teams, economic costs may be related to the original costs of recruiting patients who are eventually lost to follow-up, the time costs of delay in an affected trial, as well as the protocol-driven costs for preparing and sending reminders to those lost to follow-up. With respect to clinical and health research funders, investing in a trial affected by attrition may generate a lower return on investment due to the direct opportunity costs arising from financial resources used for recruiting and retaining participants who eventually were lost to follow-up. Finally, a resource-limited healthcare system could be affected by trial attrition. If no statistical method was valid for addressing missing data such as multiple imputation (MI) (Heymans and Twisk, 2022), it would be challenging to determine which interventions or treatments could maximise population health outcomes within the existing financial constraints.

3.2.2. The OTIS trial

Although it is believed that "*the cost of loss to follow-up…is very hard to assess*" (Grimsby and Jacobs, 2016) , this chapter introduces a framework for estimating the direct financial costs of participant losses to follow-up by using a case study from the Occupational Therapist Intervention Study (OTIS), a cohort, pragmatic, two-arm, open RCT (Cockayne et al., 2021). This trial was completed in December 2019 and explored the efficacy of a home environmental assessment and modification tool administrated by occupational therapists (OTs), versus usual care for patients subject to a higher risk of falling. The primary outcome was the number of falls per recruited participant within a year of randomisation (Cockayne et al., 2021). With respect to patient recruitment, 19308 packs were distributed to patients from GP surgeries, as well as from previous trials. 3100 patients returned their screening questionnaire and an appropriate screening form. Following eligibility checks, falls calendar packs, including falls calendars and a baseline questionnaire, were sent to 1496 patients, of whom 1410 returned a baseline questionnaire. 1331 patients were eventually randomised to the trial. 901 participants were allocated to the usual care group, whereas 430 were allocated to the intervention group following 2:1 randomisation. The number of participants recruited was slightly above the desired sample size of n=1299, allowing for an attrition rate of 10% and 90% power (two-sided 5% significance level).

With respect to participant retention the duration of follow-up was 12 months, consisting of four-monthly follow-up questionnaires, sent to the trial participants, along with a return envelope (Cockayne et al., 2021). If someone did not return their questionnaire within 21 days of the due date, they would also receive a reminder letter along with an extra copy of the questionnaire. All participants received £5 with their final follow-up questionnaire. 53 out of 430 participants (12.3%) in the intervention group, and 77 out of 901 participants (8.5%) in the control group were lost to follow-up; thus, the overall attrition rate was 9.8% (130 out of 1331 randomised participants). Excluding deaths and considering exclusively participants who decided to withdraw from the study, which is the relevant figure for the chapter, 41 out of 430 participants (9.53%) in the intervention group, and 63 out of 901 participants (7.00 %) in the control group were lost to follow-up; thus, the attrition rate, excluding deaths, is even lower, at 7.8% (104 out of 1,331 randomised participants). Nevertheless, there may be significant administration costs, as well as print and shipping costs of trial materials related to the recruitment and the follow-up of randomised participants lost to follow-up, acting as resource

misallocation to the trial commissioner's budget of £722,096.59 (NIHR Health Technology Assessment Programme 14/49/149). This chapter introduces a cost analysis for estimating such costs, the framework of which can be used for research teams to estimate the cost of participant loss to follow-up in their randomised trials.

3.2.3. Cost-utility analysis and missing data in the OTIS trial

The researchers concluded that a home environmental assessment and modification tool, administrated by OTs, was not found to reduce the incidence of falls for patients subject to a higher risk of falling, compared to usual care (Cockayne et al., 2021). In addition, following 10,000 bootstrap estimates of the incremental quality-adjusted life years (QALYs) and costs, the intervention was found to be more costly per patient (i.e. by £18.78, 95% confidence interval (CI): £16.33 to £21.24) and also less effective (i.e. by 0.0042 (95% CI 0.0041 to 0.0043) QALYs compared to usual care), thus indicating that the intervention was dominated by usual care under the base case analysis, a type of analysis using the most likely range of inputs and assumptions for producing economic evaluation results (Cockayne et al., 2021). In other words, usual care was found to be a dominant strategy, compared to the intervention, as it is both more effective and cheaper.

However, the sensitivity analysis challenged the cost-utility findings, as the complete case analysis, i.e. a type of analysis considering utility and costing data exclusively from participants whose responses did not have any missing data for producing economic evaluation results, concluded that the OT home visit intervention was less costly per patient but not significantly so (i.e. by £68.60, 95% CI: -£315.92 to £178.73) and also more effective but not significantly so (i.e. by 0.0076 (95% CI -0.0107 to 0.0259) QALYs) compared to usual care. The difference between base- and complete-case analysis is that the former applies multiple imputation (MI) methods to fill in missing data from participants who were either lost to follow-up or did not respond to follow-up questionnaires, whereas the latter considers data only from participants who provided complete data throughout the trial. In the case of the OTIS trial, observed data from 412 participants (121 in the intervention group and 291 in the usual care group) were used for the complete-case cost-utility analysis, whereas observed and multiply imputed data from 1331 participants (430 in the intervention group and 901 in the usual care group) were used for the base-case cost-utility analysis.

Thus, it is evident that missing data, for which participant loss to follow-up is partly responsible, creates uncertainty on trial health economists about which methods to use for handling missing data. To address this, the identification of the mechanism behind missing data should be identified. There are three mechanisms: data missing completely at random (MCAR), data missing at random (MAR) and data missing not at random (MNAR). MCAR "*means that missing values are randomly distributed over the data sample*" (Heymans and Twisk, 2022) , MAR "*means the probability of missing data is related to other variables*" (Heymans and Twisk, 2022) and MNAR means that "*the probability of missing data is dependent on the values of the variable itself*"(Heymans and Twisk, 2022). Unless data are MCAR, complete case analysis is not recommended. MI methods are recommended only when data are MAR. Neither method is valid when data are MNAR, which implies that the cost-utility or cost-effectiveness findings would be unreliable regardless of the approach adopted. Biasedness around the cost-effectiveness results could lead to significant costs for national healthcare systems as a result of having inconclusive or uncertain evidence surrounding the cost-effectiveness of treatments evaluated in a given randomised trial.

The authors demonstrated that the missing at random (MAR) assumption was likely to be reliable in the OTIS trial, following logistic regressions that assessed the correlation between missing data and control variables, and that between missing data and reported costs and utility scores (Cockayne et al., 2021). The only significant predictor behind participant loss to follow-up was the case where a participant had had at least one previous fall, which increased the likelihood of missing outcome data (odds ratio (OR): 5.84 (95% CI: 1.13, 30.21); p = 0.04). When this factor was used as a covariate in the primary analysis, the intervention effect estimate remained unchanged at the 5% significance level (adjusted incidence rate ratio (IRR): 0.17 (95% CI: 0.99,1.38); p = 0.07) (Cockayne et al., 2021). These results imply that the base-case analysis cost-utility findings were likely to be reliable, in contrast to those observed under the complete-case analysis.

3.2.4. Aim of Chapter 3

The primary aim of this chapter is to estimate the opportunity cost of participant losses to follow-up in the OTIS trial by estimating the aggregate and average protocol-driven costs related to participant loss to follow-up. This estimation will provide insights into the average cost of participant loss from the economic perspective of the trial team, and its impact on resource misallocation for trial teams and funders. Given the limited evidence on the

opportunity costs of participant attrition, the chapter intends to offer guidance to research teams on estimating these costs in their own randomised trials, facilitating informed budget decisions.

3.3. Methods

3.3.1. Types of participant loss to follow-up

There were multiple periods where follow-up occurred in the OTIS trial after randomisation, at 4, 8, and 12 months through follow-up questionnaires (Cockayne et al., 2021). Therefore, prior to estimating the average cost per participant lost to follow-up, it was recognised that attrition in the case study was differential by the following elements:

1) Time point at which a participant was lost to follow-up (i.e. (a) before receiving the 4-month follow-up questionnaire, due to withdrawal from trial; (b) after having received the 4-month follow-up questionnaire but before receiving the 8-month follow-up questionnaire, due to withdrawal from trial; (c) after having received the 4- and 8-month follow-up questionnaires but before receiving the 12-month follow-up questionnaire, due to withdrawal , and; (d) after having received all follow-up questionnaires, due to not responding to the 12-month follow-up questionnaire. Distinguishing the latest follow-up questionnaires that the participants, lost to follow-up received was necessary for estimating the actual indirect and direct costs the trial team incurred as a result of attrition.

2) The group they were allocated to (i.e. intervention group or usual care group).

3) Whether the participants lost to follow-up had responded to any of the follow-up questionnaires they received (i.e. excluding those lost to follow up before receiving the 4-month follow-up questionnaire). If someone did not respond to any follow-up questionnaire within 21 days of receipt, the trial team would incur an additional cost of sending a reminder letter plus an extra copy of the questionnaire (Cockayne et al., 2021). Please note that if a participant did not withdraw from the study or die, the trial team continued to send them future follow-up questionnaires (Cockayne et al., 2021).

By considering these differential elements, four different types of participant loss to follow-up were identified for the OTIS trial with regards to whether participants lost to follow-up had actually responded to any of the follow-up questionnaires. *Table 3.1* presents these different types of participant loss to follow-up in the OTIS trial plus the number of participants

associated with each of these types, based upon Table 5 of the published study (Cockayne et al., 2021) . *Supplemental Material 3.1* provides details about the allocation of participants lost to follow-up by type and subtype of attrition. The reported follow-up and response rates from the published study were considered to determine the sample size for each type and subtype of attrition (Cockayne et al., 2021). For ethical reasons, losses to follow-up generated by death events (12 in the intervention group and 14 in the control group) are not considered in the analysis.

3.3.2. Protocol-driven costs

Having defined attrition in the OTIS trial, the next step was to define the protocol-driven opportunity costs by following the economic perspective of the trial team.

Initially, the trial team incurred costs for recruiting patients with a higher risk of falling to the trial. Given the published trial report, the recruitment packs that all potentially eligible patients received consisted of the following case report forms (CRFs): invitation letters, participant information sheets (PIS), consent forms and screening forms (Cockayne et al., 2021). Evidently, the trial team incurred costs for administering, printing, and shipping recruitment packs to patients, as well as for processing and managing data collected from potentially eligible patients, who were eventually recruited but decided to withdraw from the trial.

Secondly, the trial team incurred costs related to the data management of recruited participants, as well as for administering, printing and shipping falls calendar packs (consisting of monthly falls calendars, the falls prevention leaflet and the baseline questionnaire), reminder letters, and the 4-, 8-, and 12-month follow-up questionnaires to randomised participants. Third, it incurred such costs for group newsletters, which were sent to recruited participants at three months after randomisation and two weeks before receiving their 12-month follow-up questionnaire (Cockayne et al., 2021). Fourth, the trial team incurred financial costs for paying £5 in cash to participants alongside the 12-month questionnaire.

Table 3.1: Types of participant loss to follow-up in the OTIS trial

Type of loss to follow-up	Intervention group	Usual care group	All participants
1. Attrition before receiving 4-month questionnaire due to withdrawal or death	**16**	**12**	**28**
2. Attrition before receiving 8-month questionnaire due to withdrawal or death	**9**	**15**	**24**
2a. Responded to 4-month questionnaire	0	0	0
2b. Not responded to 4-month questionnaire	9	15	24
3. Attrition before receiving 12-month questionnaire due to withdrawal or death	**4**	**12**	**16**
3a. Responded to both 4-month and 8-month questionnaires	0	0	0
3b. Not responded to neither 4-month nor 8-month questionnaires	2	10	12
3c. Responded to the 4-month, but not to the 8-month questionnaire	2	2	4
3d. Responded to the 8-month, but not to the 4-month questionnaire	0	0	0
4. Attrition due to not responding to 12-month questionnaire	**12**	**24**	**36**
4a. Responded to the 8-month questionnaire, but not to the 12-month questionnaire	11	16	27
4b. Responded to neither 8-month nor 12-month questionnaires	1	8	9
All types (total)*	**41**	**63**	**104**

* Note due to unequal randomisation (i.e. 2:1 in favour of the usual care group) it would be expected that twice as many patients be lost to follow-up from the control group if the percentage loss to follow-up was the same.

Fifth, shipping costs of monthly falls calendars sent by trial participants were incurred by the trial team; however, data regarding the frequency of participants lost to follow-up in responding to monthly falls calendars are not available. Information regarding the assumption I made to estimate the number of falls calendars returned by participants lost to follow-up to the trial team can be found in *Supplemental Material 3.2*. These costs are not differentiated by treatment group. Evidently, there seem to be three main costing components from the economic perspective of the trial team: administration, print, and shipping costs.

- Administration costs reflect the time required for trial staff to perform administrative tasks related to the recruitment of potentially eligible patients and retention of randomised participants. Relevant tasks include recording participant ID numbers, preparing and distributing trial materials, keeping records of questionnaire delivery and data entry for previous timepoints. Under this category fall both the administration costs related to trial materials and the costs associated with data management. The time dedicated to administering trial materials and data for participants lost to follow-up reflect an opportunity cost, in the sense it could have been allocated to other tasks within the trial if the marginal rates of productivity in other tasks were positive.
- Print costs are direct costs related to printing the relevant material any potentially eligible patient or recruited participant should receive before and after randomisation.
- Shipping costs are direct costs related to sending by post the relevant material any potentially eligible patient or recruited participant should receive before and after randomisation. In the OTIS trial, shipping costs consist of financial costs of pre-paid freepost envelopes, free-post licences, return freepost postages, and outgoing postages.

The sum of the administration, print and shipping costs represents the unit cost of a trial material (e.g.4-month follow-up questionnaire) from the economic perspective of the trial team.

3.3.3. Administration costs related to participants lost to follow-up

The administration team in the OTIS trial consisted of a part-time data clerk (annual salary cost of £10,228 based on 50% full-time equivalent (FTE)), a part-time Grade 8 data manager (annual salary cost of £10,795, based on 20% full-time equivalent (FTE)) and a full-time Grade 5 trial support officer (annual salary cost of £32,033). These salaries were reported, at 2015 price levels, in the grant application of the York Trials Unit (YTU) to the National Institute for

Health & Care Research (NIHR), including employer national insurance and workplace pension contributions. Access to the grant application, and thus to staff costs, has been given confidentially by the trial team.

The following formula estimates the aggregate administration cost, considering the data team's salary costs and proportion of time related to administrative tasks relevant for the analysis.

Administration costs = Data clerk's salary costs*υ + Data manager's salary costs*β + Trial support officer's salary costs*γ *(Equation 3.1)*

, where υ, β and γ correspond to the proportion of labour time devoted to administration tasks in relation to trial materials. Given the difficulty in estimating these parameters due to differing staff workloads, the following assumptions were made: υ=100% to reflect the primary focus of data clerks on data entry and processing, β=40% to account for the data management oversight by data managers, and γ=60% considering the dual role of trial support officers in providing administrative assistance for trial materials (plus keeping records of questionnaire delivery) and supporting trial researchers and investigators. Assuming υ=100%, a two-way sensitivity analysis was undertaken to explore the impact of varying β and γ jointly between 30% and 100% to evaluate the robustness of the findings given these assumptions. Moreover, since information on task allocation could not be obtained, the trial staff were assumed to have distributed equally their time in administration tasks across the trial materials.

After solving *Equation 3.1*, and considering the assumptions made, the total administration costs are estimated to be £33,765.80, with the administration costs per trial material being £4,823.69. To express the administration costs in terms of unit costs, the aforementioned figure is divided by the number of potentially eligible patients or recruited participants to whom the following trial materials were sent: the recruitment pack (n=19,308), the falls calendar pack (n=1,496), the 4-month follow-up questionnaire (n=1,301), the 8-month follow-up questionnaire (n=1,268), the 12-month follow-up questionnaire (n=1,237), the first group newsletter (n=1,301) and the second group newsletter (n=1,237) (Cockayne et al., 2021). Note that the administration costs of monthly falls calendar returns, and reminder letters are assumed to be included already in the unit administration costs of the follow-up questionnaires, as they formed part of the administrative follow-up efforts by the trial team.

3.3.4. Print and shipping costs of trial materials related to participant loss to follow-up

Following correspondence from YTU's member of staff, data about the unit print and shipping costs of the OTIS trial's CRFs (i.e. recruitment pack, falls calendar pack, follow-up questionnaires and group newsletters) were collected and reported at 2017 price levels. Hence, the administrative unit costs are later converted from 2015 price levels to 2017 price levels, using the Consumer price inflation time series data from the World Bank, for homogeneity purposes; the two-year UK inflation rate from 2015 to 2017 was 3.7% (World Bank, 2023).

The unit shipping costs were formed using the following costing figures: with regards to the recruitment pack, a non-window A4 envelope costed £0.03, on which a printed address label with a unit cost of £0.02 was added. With respect to the recruitment pack, the falls calendar pack and the follow-up questionnaires, the unit cost of a pre-paid A4 freepost envelope was £0.19, the unit cost of a return freepost postage for A4 letter was £0.66, and the unit cost of outgoing postage (A4 size) was £0.76 per envelope. With respect to the falls calendar and the follow-up questionnaires, the unit cost of window A4 envelope was £0.05. The unit shipping costs of the two group newsletters consisted of the outgoing postage (A5/C5 size: (£0.56)) and an envelope (£0.02). In relation to monthly falls calendar returns, the unit cost of a return freepost postage was £0.44.

The unit printing cost related to the recruitment pack includes the unit printing costs of the participant information sheet (£0.34), the invitation letter (£0.15), the consent form (£0.13), the screening form (£0.10), the contact form (£0.10), the baseline questionnaire (£0.32), the falls calendar (£1.23) and the falls prevention leaflet (£3.92). The unit printing cost related to the follow-up questionnaires includes the unit printing costs of the cover letter (£0.02) and the follow-up questionnaires (£0.36). Finally, the unit printing cost of a group newsletter is £0.25.

3.3.5. Unit costs related to participant loss to follow-up

By aggregating the unit administrative, print and shipping costs of each of the study's materials, it became possible to estimate their corresponding unit costs.

Unit cost= Unit administrative cost + Unit print cost + Unit shipping cost *(Equation 3.2)*

In addition to the unit costs of trial materials, the trial team incurred an incremental cost for delivering the intervention to participants who were randomised to the intervention group. This incremental cost consisted of the training the occupational therapists (OTs) received to

undertake home visits, the time spent at home visit and on follow-up telephone call, and the equipment/adaptations installed following the home visit. The figure, expressed in per participant terms, is provided in the published NIHR report related to the trial, i.e. £136.53 (Cockayne et al., 2021).

3.3.6. Average and aggregate costs related to participant loss to follow-up

The next step was to estimate the unit costs by type of attrition, from the economic perspective of the trial team. To do so, it was crucial to detect which materials participants lost to follow-up had received, so that the unit costs be estimated accordingly using the unit cost figures. A description of which trial materials participants lost to follow-up received, by type of attrition, is presented on *Table 3.2*.

It is assumed that participants who were lost to follow-up and had responded to a previous follow-up questionnaire did so without the need for receiving an additional copy of this questionnaire. Combining the information from *Table 3.2* and the unit costs of trial materials and the intervention, the unit costs by type of attrition, from the economic perspective of the trial team, were estimated.

The unit cost figures, by type of attrition and intervention group, were multiplied by the number of participants falling into each of these attrition types and intervention groups from *Table 3.1*, to estimate the aggregate monetary cost of participants lost to follow-up to the trial team. For simplicity purposes and due to absence of relevant information regarding intervention delivery in the study's report, it is assumed that all participants from the intervention group who were lost to follow-up had received an OT home visit; in the trial's original NIHR report, it is reported that 88.6% of the recruited participants in the intervention group received the home visit (Cockayne et al., 2021). The cost per participant lost to follow-up, from the economic perspective of the trial team, was also obtained.

Table 3.2: Trial materials received by participants lost to follow-up, by type of attrition

Type of loss to follow-up	**Intervention group**	**Usual care group**
1. Attrition before receiving 4-month questionnaire due to withdrawal or death	1x Recruitment pack 1x Falls calendar pack 1x OT home visit 1x First group newsletter 1.56x Falls calendar monthly return	1x Recruitment pack 1x Falls calendar pack 1x First group newsletter 1.56x Falls calendar monthly return
2. Attrition before receiving 8-month questionnaire due to withdrawal or death		
2a. Responded to 4-month questionnaire	1x Recruitment pack 1x Falls calendar pack 1x 4-month questionnaire 1x OT home visit 1x First group newsletter 6.32 x Falls calendar monthly return	1x Recruitment pack 1x Falls calendar pack 1x 4-month questionnaire 1x First group newsletter 6.32 x Falls calendar monthly return
2b. Not responded to 4-month questionnaire	1x Recruitment pack 1x Falls calendar pack 2x 4-month questionnaire 1x OT home visit 1x First group newsletter 6.32 x Falls calendar monthly return	1x Recruitment pack 1x Falls calendar pack 2x 4-month questionnaire 1x First group newsletter 6.32 x Falls calendar monthly return
3. Attrition before receiving 12-month questionnaire due to withdrawal or death		
3a. Responded to both 4-month and 8-month questionnaires	1x Recruitment pack 1x Falls calendar pack 1x 4-month questionnaire 1x 8-month questionnaire 1x OT home visit 1x First group newsletter 10.51 x Falls calendar monthly return	1x Recruitment pack 1x Falls calendar pack 1x 4-month questionnaire 1x 8-month questionnaire 1x First group newsletter 10.51 x Falls calendar monthly return
3b. Not responded to neither 4-month nor 8-month questionnaires	1x Recruitment pack 1x Falls calendar pack 2x 4-month questionnaire 2x 8-month questionnaire 1x OT home visit 1x First group newsletter 10.51 x Falls calendar monthly return	1x Recruitment pack 1x Falls calendar pack 2x 4-month questionnaire 2x 8-month questionnaire 1x First group newsletter 10.51 x Falls calendar monthly return
3c. Responded to the 4-month, but not to the 8-month questionnaire	1x Recruitment pack 1x Falls calendar pack 1x 4-month questionnaire 2x 8-month questionnaire 1x OT home visit 1x First group newsletter 10.51 x Falls calendar monthly return	1x Recruitment pack 1x Falls calendar pack 1x 4-month questionnaire 2x 8-month questionnaire 1x First group newsletter 10.51 x Falls calendar monthly return
3d. Responded to the 8-month, but not to the 4-month questionnaire	1x Recruitment pack 1x Falls calendar pack 2x 4-month questionnaire 1x 8-month questionnaire 1x OT home visit 1x First group newsletter 10.51 x Falls calendar monthly return	1x Recruitment pack 1x Falls calendar pack 2x 4-month questionnaire 1x 8-month questionnaire 1x First group newsletter 10.51 x Falls calendar monthly return
4. Attrition due to not responding to 12-month questionnaire		

4a. Responded to the 8-month questionnaire, but not to the 12-month questionnaire	1x Recruitment pack 1x Falls calendar pack 1x 4-month questionnaire 1x 8-month questionnaire 1x 12-month questionnaire 1x 12-month questionnaire + reminder letter 1x OT home visit 1x First group newsletter 1x Second group newsletter 1x £5 monetary reward 10.51 x Falls calendar monthly return	1x Recruitment pack 1x Falls calendar pack 1x 4-month questionnaire 1x 8-month questionnaire 1x 12-month questionnaire 1x 12-month questionnaire + reminder letter 1x First group newsletter 1x Second group newsletter 1x £5 monetary reward 10.51 x Falls calendar monthly return
4b. Responded to neither 8-month nor 12-month questionnaires	1x Recruitment pack 1x Falls calendar pack 1x 4-month questionnaire 1x 8-month questionnaire 1x 8-month questionnaire + reminder letter 1x 12-month questionnaire 1x 12-month questionnaire + reminder letter 1x OT home visit 1x First group newsletter 1x Second group newsletter 1x £5 monetary reward 10.51 x Falls calendar monthly return	1x Recruitment pack 1x Falls calendar pack 1x 4-month questionnaire 1x 8-month questionnaire 1x 8-month questionnaire + reminder letter 1x 12-month questionnaire 1x 12-month questionnaire + reminder letter 1x First group newsletter 1x Second group newsletter 1x £5 monetary reward 10.51 x Falls calendar monthly return

3.4. Results

3.4.1. Unit administration costs of trial materials related to participant loss to follow-up

At 2015 price levels, the unit administration cost of a recruitment pack is £0.25; the unit administration cost of a falls calendar pack, including the baseline questionnaire, is £3.22; the unit administration cost of a 4-month follow-up questionnaire is £3.70; the unit administration cost of an 8-month follow-up questionnaire is £3.80; the unit administration cost of a 12-month follow-up questionnaire is £3.90; the unit administration cost of the first group newsletter is £3.70; and the unit administration cost of the second group newsletter is £3.90. These are the unit administration costs reported at 2015 price levels.

Following the conversion to 2017 price levels (World Bank, 2023) , the unit administration cost of a recruitment pack is £0.26; the unit cost of a falls calendar pack, including the baseline questionnaire, is £3.34; the unit administration cost of a 4-month follow-up questionnaire is £3.84; the unit administration cost of an 8-month follow-up questionnaire is £3.94; the unit administration cost of a 12-month follow-up questionnaire is £4.04; the unit administration cost of the first group newsletter is £3.84 ; and the unit administration cost of the second group

newsletter is £4.04. The unit administration costs also consider the administration costs related to the monthly falls calendar returns, as they formed part of the follow-up efforts related to data collection from follow-up questionnaires by the trial team and the administration costs of the falls calendar pack.

3.4.2. Unit print costs of trial materials related to participant loss to follow-up

The unit print cost of a recruitment pack was £0.82, including the unit print costs of an invitation letter (£0.15), a PIS (£0.34), a consent form (£0.13), a screening form (£0.10), and a contact form (£0.10), all colour printed and sized in A4 papers. The unit print cost of a full falls calendar, consisting of 18 C5 cards, was £1.23. The unit print cost of a baseline questionnaire was £0.32. The unit print cost of the patient falls prevention leaflet was £3.92. The unit print cost of the falls calendar packs, including a copy of the baseline questionnaire, was £1.55. The unit print cost of a copy of the 4-, 8-, and 12-month follow-up questionnaires, each of which included an additional page with a cover letter, was £0.38. The unit print cost of a group newsletter (both first and second) was £0.25. In addition, the print costs of the reminder letter (1 page-coloured), i.e. £0.15, are considered for participants not responding to follow-up questionnaires. Note that these costs were reported at the 2017 price levels.

3.4.3. Unit shipping costs of trial materials related to participant loss to follow-up

The unit shipping cost of a recruitment pack was £1.66. The unit shipping costs of a falls calendar pack (including falls calendar plus baseline questionnaire) and a 4-, 8- and 12-month follow-up questionnaire were identical, at £1.66. The unit shipping cost of a group newsletter (both first and second), printed on A5 or C5 paper, was £0.58, which includes the unit cost of an envelope and outgoing postage. Note that these costs were reported at the 2017 price levels. Finally, the shipping costs of the falls calendar monthly returns are considered, with the return freepost postage for a C5 letter costing £0.44 to the trial team.

3.4.4. Unit costs related to participant loss to follow-up

By aggregating the administration, print and shipping unit costs, the unit cost of a recruitment pack is £2.74, the unit cost of a falls calendar pack, including a copy of the baseline questionnaire, falls calendars and the falls prevention leaflet, is £10.47, the unit cost of a 4-month follow-up questionnaire is £5.88, the unit cost of an 8-month follow-up questionnaire is £5.98, the unit cost of a 12-month follow-up questionnaire is £6.08, the unit cost of the first

group newsletter is £4.67, the unit cost of the second group newsletter is £4.87, the unit shipping cost of the monthly falls calendar return is £0.44, the unit print cost of reminder letters is £0.15 and the unit cost of the monetary reward is £5.The NIHR report of the study estimated the total unit cost of the OT home visit intervention to be £136.53 at 2017 price levels (Cockayne et al., 2021), which is also used in this chapter for estimating the costs of attrition to the OTIS trial. All unit costs from the economic perspective of the trial team are summarised in *Table 3.3*.

Table 3.3: Unit costs from the economic perspective of the trial team

Item	Unit administration cost (£, 2017)	Unit print cost (£, 2017)	Unit shipping cost (£, 2017)	Unit cost (£, 2017)
Recruitment pack	0.26	0.82	1.66	2.74
Falls calendar pack	3.34	5.47	1.66	10.47
4-month follow-up questionnaire	3.84	0.38	1.66	5.88
8-month follow-up questionnaire	3.94	0.38	1.66	5.98
12-month follow-up questionnaire	4.04	0.38	1.66	6.08
First group newsletter	3.84	0.25	0.58	4.67
Second group newsletter	4.04	0.25	0.58	4.87
OT home visit intervention	NA	NA	NA	136.53
Falls calendar monthly return	NA*	NA*	0.44	0.44
Reminder Letter	NA**	0.15	0.00**	0.15
£5 monetary reward in cash	NA	NA	NA	5.00

* Unit administration costs for falls calendar monthly returns already considered in the unit administration costs related to the falls calendar pack and the follow-up questionnaires. Unit print costs related to falls calendars already included in the unit print cost of the falls calendar pack.

** Unit administration costs for reminder letters already considered in the unit administration costs related to the follow-up questionnaires. Reminder letters were sent alongside an extra copy of a follow-up questionnaire; hence their additional unit shipping cost is zero.

3.4.5. Average and aggregate costs related to participant loss to follow-up

The unit costs by type of attrition, from the economic perspective of the trial team, are available in *Table 3.4*. The unit cost figures, by type of attrition and intervention group, were multiplied by the number of participants falling into each of these attrition types and intervention groups from *Table 3.1*, to generate an aggregate monetary cost of participants lost to follow-up to the trial team, expressed as the sum of the costing figures included in *Table 3.4*. The aggregate,

direct financial costs of participant loss to follow-up, from the economic perspective of the trial team, were estimated to be £10,234.90.

Therefore, the average cost per participant lost to follow-up, from the economic perspective of the trial team, was:

Average cost per participant lost to follow-up = £98.41 *(Equation 3.3)*

Given that; (1) the estimated cost of attrition is £10,234.90 and; (2) the NIHR awarded £722,096.59 to the trial team, it means that (£10,234.90/£722,096.59) *100% = 1.42 % of the funding has been lost because of participant loss to follow-up.

The two-way sensitivity analysis, which jointly varied the proportion of labour time dedicated by the data manager and the trial support officer for administration tasks related to our analysis between 30% and 100%, suggests that the findings remain robust despite the assumptions made on the parameters υ,β and γ, with the aggregate cost of attrition varying between £ 9,618.98 (when β= γ= 30%) and £11,346.50 (when β= γ= 100%).

Table 3.4: Unit and aggregate costs by type of attrition, from the economic perspective of the trial team

Type of loss to follow-up	**Intervention group (unit cost; £, 2017 level)**	**Usual care group (unit cost; £, 2017 level)**	**Intervention group (aggregate cost; £, 2017 level)**	**Usual care group (aggregate cost; £, 2017 level)**
1. Loss to follow-up before receiving 4-month questionnaire due to withdrawal or death	155.10	18.57	2481.66	222.89
2. Loss to follow-up before receiving 8-month questionnaire due to withdrawal or death				
2a. Responded to 4-month questionnaire	163.77	27.24	0	0
2b. Not responded to 4-month questionnaire	172.59	36.06	1553.27	540.83
3. Loss to follow-up before receiving 12-month questionnaire due to withdrawal or death				
3a. Responded to both 4-month and 8-month questionnaires	174.38	37.85	0	0
3b. Not responded to neither 4-month nor 8-month questionnaires	191.17	54.64	191.17	273.22

3c. Responded to the 4-month, but not to the 8-month questionnaire	189.76	53.23	569.29	372.63
3d. Responded to the 8-month, but not to the 4-month questionnaire	189.66	53.13	0	0
4. Loss to follow-up due to not responding to 12-month questionnaire				
4a. Responded to the 8-month questionnaire, but not to the 12-month questionnaire	201.19	64.66	2213.14	1034.64
4b. Did not respond to neither 8-month nor 12-month questionnaires	207.33	70.80	207.33	566.40

3.5. Discussion

3.5.1. Summary of findings

The OTIS trial had an overall attrition rate of 9.8%. Given that such a rate is low, the likelihood of attrition bias, the main source of threat to the statistical validity of a trial's findings, is low. Nevertheless, participant attrition still occurred, albeit in small numbers, and generated significant economic costs for trial teams and research commissioners.

The average cost per participant lost to follow-up is estimated to be £98.41. Such a figure includes the administration costs of data management, as well as administration, print, and shipping costs of trial materials, including the recruitment pack, the falls calendar pack, reminder letters, the follow-up questionnaires and newsletters, as well as the cost of the trial's intervention of OT home visit, related to participants lost to follow-up, plus the final £5 monetary reward. Given the timepoints and frequency of attrition, the aggregate direct cost is estimated to be £10,234.90. This figure also reflects the direct financial costs of attrition in the OTIS trial to the trial's funding source, the NIHR (Cockayne et al., 2021), as £10,234.90 from its allocated funding to the OTIS trial has been lost instead of invested in another clinical study/studies that could have generated additional health benefits for the population.

In 2016/17, the NIHR allocated £67.6 million for pilot or full-scale randomised trials (National Institute for Health & Care Research, 2024) (see *Supplemental Material 3.3* for more details). The OTIS trial, approved during that period, received £722,096.59 in funding (National Institute for Health & Care Research, 2024) . Assuming other RCTs experienced comparable losses to follow-up as a percentage of their funding, i.e. (£10,234.90/£722,096.59) *100% = 1.42 %), the estimated annual cost of attrition for the NIHR would be approximately 1.42%* £67,553,088 = £957,491. In other words, approximately £960,000 invested from the UK's

main health research funder may have not led to improvements in health outcomes for the UK population because of foregone resources originating from participant attrition in its funded studies. Considering that the average cost per funded feasibility or full-scale RCT was £1,535,297 in the 2016/17 period (National Institute for Health & Care Research, 2024) , these findings illustrate that had resource waste not existed as a result of attrition, the NIHR could have invested in 0.62 additional RCTs instead. In addition, a resource misallocation of £960,000 suggests that if the £20,000 cost-effectiveness threshold recommended by the National Institute for Health and Care Excellence (NICE) (Claxton et al., 2015b) , which corresponds to a maximum of £20,000 invested per additional Quality-Adjusted Life Year (QALY) gained, at least 48 QALYs could have been saved had attrition not occurred in NIHR-funded trials, and these economic resources were invested in cost-effective health technologies and treatments instead. This estimate, however, will understate the opportunity costs for those trials where follow up includes a clinical, face to face, appointment. Therefore, the estimate of the attrition cost to the NIHR HTA programme is probably very conservative.

3.5.2. Strengths and limitations of the study

Aside from the statistical implications of attrition, trial teams face opportunity costs because of administering data related to and providing trial materials for recruited participants lost to follow-up ; this chapter highlights such costs for a case study which, despite its low attrition rate, faced significant financial costs due to participant losses to follow-up. To estimate such costs, this chapter presents a detailed cost analysis which considers the unit protocol-driven costs of administering, printing and shipping each trial material participants lost to follow-up received. In addition, the methodology differentiates the nature of participants lost to follow-up, according to the time point at which they were lost to follow-up, the treatment group they were allocated to, and their history of responding to previous follow-up questionnaires. It is the first time a study attempts to estimate the direct financial costs of attrition associated with a specific randomised trial.

These findings highlight the importance for trial teams to find methods to reduce participant losses to follow-up, as the resources foregone could have been allocated more efficiently towards other trial-related activities. These activities may include the recruitment of additional participants to improve the statistical power of a trial, improving the delivery of interventions, extending the follow-up period to explore clinical or health outcomes in the longer term,

optimising data collection and statistical analysis methods, or implementing a more intensive monitoring of trial-related activities. This analysis could also be important for trial funders as attrition, which causes resource misallocation, can undermine their efforts to maximise population health outcomes (e.g. a net loss of 48 QALYs if the cost of attrition accounts for 1.42% of the funding allocated by the NIHR to randomised trials) and their return on investment (ROI) within their constrained budget. Finally, trial teams could consider the methodology presented to estimate the economic costs of attrition in their randomised trials, even before commencing the trials, in order to optimise their budget by considering alternative retention strategies that could reduce loss to follow-up. Given the opportunity costs of attrition, such a course of action could enhance the research impact of their studies given the budget constraints.

However, the generalisability of the costing methodology may be limited, as the follow-up data collection methods are not solely limited to postal follow-up questionnaires. If a trial involves clinical appointments for follow-up data collection, costs such as staff costs related to follow-up activities, frequency of clinic visits for follow-up or reimbursable participant transport costs (e.g. parking fees) to the clinic may also need to be considered in the estimation of the opportunity costs of attrition. Alternatively, if a trial involves telephone interviews, duration of calls, call charges and trial staff costs related to the conduct of telephone interviews could be relevant opportunity costs. On the other hand, if a trial involves electronic and web-based questionnaires, costs related to electronic data collection and management may be relevant rather than print and shipping costs of trial materials delivered to participants by post. If a trial involves postal follow-up methods, the framework presented can be more easily adapted in such a trial. Moreover, if a trial involves alternative interventions, there can be variations in terms of resource utilisation that may require an adjustment to the estimation of the opportunity costs of attrition. For instance, trials of nursing interventions may involve more intensive ongoing training, administrative and monitoring activities, whose costs are incurred by the trial team. On the other hand, trials of medical interventions may involve medical technologies for intervention delivery such as diagnostic tests and more intensive monitoring of adverse events. Thus, trial teams need to adjust the relevant resource items that should be considered for the estimation of the opportunity costs of attrition in their trials according to the follow-up data collection methods and intervention design. Nevertheless, the framework presented offers a detailed methodology as to how the opportunity costs of attrition were estimated in a case

study, the rationale and the approach of which could be applicable to trials of different follow-up methods or intervention designs.

In addition, several assumptions were made for the estimation of the opportunity cost of attrition in the OTIS trial because of the unavailability of relevant data. For instance, assumptions have been made with regards to the sample size of participants lost to follow-up corresponding to each type and subtype of attrition in the OTIS trial (S*upplemental Material 3.1*), as well as with respect to the number of monthly falls calendars that were returned by post to the trial, by type and subtype of attrition (*Supplemental Material 3.2*). It was also difficult to assign precise proportions of labour time devoted to trial-related administrative activities, something which led to assumptions regarding such proportions; however, the two-way sensitivity analysis has confirmed the robustness of the findings despite the assumptions made.

3.5.3. Direction for future research

The analysis presented could work as a guidance for research teams to estimate the average and aggregate costs of participant attrition in their randomised trials, even prior to their recruitment phase. In addition, estimates of the aggregate and average cost of attrition could be stated in grant applications to trial commissioners, such as the NIHR, as participant losses to follow-up affect indirectly these institutions due to resource waste as discussed already.

Given the costs of participant attrition, it seems imperative that trial teams and decision makers identify strategies that could improve participant retention in randomised trials. A study design which could identify such strategies is Studies Within A Trial (SWATs) (Treweek et al., 2018a). Research on SWATs of retention strategies is already in place (Gillies et al., 2021). Improving retention in randomised trials could reduce the direct financial costs of attrition and ensure statistical accuracy. So far, no retention strategy has been found to be effective with high certainty of evidence (Gillies et al., 2021). In addition to SWATs, individualised approaches and removing barriers to participation for improving participant retention are mentioned as solutions in the literature although further evidence would be necessary to explore the impact of such approaches (Brueton, 2022). Finally, I hope the findings of this chapter will encourage trial teams to understand the wider implications of participant losses to follow-up in their studies; according to a recent scoping review, only 36.8% of RCT protocols have made reference to plans for promoting participant retention (Murphy et al., 2022a).

In addition, whereas MI methods managed to address missing data because they were MAR, they would not have been able to do so had data been MNAR. Given this scenario, it would be interesting for health economists to explore the costs of missing data, and thus attrition, from the economic perspective of a healthcare system as established statistical methods for dealing with missing data would be invalid under MNAR data.

3.5.4. Concluding remarks

Despite the low attrition rate of the OTIS trial, which is lower than the median attrition rate usually observed in randomised trials, loss to follow-up has still generated considerable financial costs for the trial team. It is strongly recommended that decision makers focus on identifying strategies which could improve participant retention in randomised trials. The conduct of SWATs for novel and existing retention strategies could positively contribute to the aim of improving trial efficiency, with benefits for patients, trial teams and funders.

Chapter 4: Existing evidence on SWATs for improving recruitment and retention in RCTs

Given the remarkable costs poor recruitment and attrition can generate to trial teams and national healthcare systems, as evidenced by *Chapter 2* and *Chapter 3*, it is strongly suggested that strategies addressing such challenges be identified. The following chapter aims to gather and critically appraise the evidence on the effectiveness and the cost-effectiveness of existing recruitment and retention strategies that have been evaluated from SWATs.

With regards to the *effectiveness* of recruitment and retention strategies, two systematic reviews have been published in Cochrane (Treweek et al., 2018b) (Gillies et al., 2021). The chapter summarises the methodology and the findings of the two reviews, before proceeding to the first systematic review that has ever been undertaken, with respect to the *cost-effectiveness* of recruitment and retention strategies.

4.1. Overview of Cochrane reviews on strategies to improve recruitment and retention in RCTs

4.1.1. Cochrane systematic review of recruitment strategies (Treweek et al., 2018b)

The latest Cochrane review focusing on recruitment aimed to "*quantify the effects of strategies for improving recruitment of participants to randomised trials*" (Treweek et al., 2018b).

4.1.1.1. Methods

The study selection process focused on randomised and quasi-randomised trials, i.e. trials that allocate participants to treatments using non-random techniques such as treatment allocation according to age or baseline clinical status, (including SWATs and other types of RCTs) related to healthcare and non-healthcare settings, in order to appraise the existing evidence on the effectiveness of strategies for improving recruitment. In addition to randomised trials recruiting real-life participants making real decisions, the authors included studies of hypothetical trials, although they later suggested that future associated reviews avoid hypothetical trials due to a potentially high risk of bias. The considered population consisted of potential trial participants, trial collaborators and research ethics committees. The interventions were any strategies designed to improve recruitment of participants to RCTs. The primary outcome was "*the proportion of eligible participants or centres recruited*" (Treweek et al., 2018b). Given the

context of this review, no comparators were pre-specified. Studies related to retention interventions and qualitative studies related to improving questionnaire response were excluded from this review.

To identify the relevant studies, the reviewers applied a search strategy to electronic databases, such as the Cochrane Methodology Review Group Specialised Register (CMR), MEDLINE, Embase, Science Citation Index & Social Science Citation Index (ISI) and ERIC, with records searched up to 11 February 2015. The data extraction generated records on methods evaluated, country, population, study setting, nature of the study to be recruited into, randomisation or quasi-randomisation, and numbers and proportions of participants in the intervention and control groups of the study that compared recruitment interventions. The Cochrane Risk of Bias tool was used to assess the risk of bias of the included studies. A GRADE assessment was also undertaken for all interventions, regardless of whether they were associated with a single or more studies. Hypothetical and high-risk-of-bias studies were considered in the data analysis. Heterogeneity of the included trials was captured through the i-squared (I^2) test, whilst the factors the authors believed to have determined substantial heterogeneity across the affected studies could be the following: type of design and allocation concealment, setting of the study recruiting participants, disease area, design of the study recruiting participants, target group and recruitment to hypothetical versus real trials. The data synthesis categorised trials based on the Online Resource for Recruitment research in Clinical triAls (ORRCA) project: design; pre-trial planning; trial conduct changes; modifications to the consent process; modification to the information given to potential participants about the trial; interventions aimed at the recruiter or recruitment site; incentives (Kearney et al., 2018). All results were reported in 95% confidence intervals (CI) associated with fixed-effect and random-effect risk differences (RD).

4.1.1.2. Results

In total, 25,432 titles and abstracts were screened, with 377 deemed eligible for full text based on the inclusion and exclusion criteria. In addition, the full texts of six studies were obtained. Studies totalling 68 were included, and thus were eligible for data extraction. Out of the 68 included studies, 63 were designed for trial participants, whereas the remaining five were designed for professionals recruiting trial participants. The included studies were undertaken in 12 countries, but most frequently in the USA and the UK. The Cochrane risk of bias was assessed to be low for 22 studies, high for 32 studies and unclear for 14 studies.

With respect to the design, an open design, i.e. a design where the participants and/or trial researchers are aware of participants' treatment allocation, compared to a placebo-controlled design, i.e. a design where the participants and/or trial researchers are not aware of participants' treatment allocation, was shown to increase recruitment by RD=10% (95% CI: 7% to13%), with high GRADE certainty of evidence; this figure was obtained following a meta-analysis of two studies (Avenell et al., 2004, Hemminki et al., 2004). In addition, the review found that patient preference designs, i.e. designs allowing participants to express their preference on the treatment they would receive and whose preferences are eventually considered for treatment allocation, may not improve the recruitment rate; internet-based data collection may reduce recruitment compared to paper-based data collection; cluster-randomised design, i.e a design where the unit of randomisation is not individual patients but groups of patients such as clinical practices. may improve recruitment, compared to a Zelen's design, i.e. an alternative design under which patients are allocated to treatment groups before giving informed consent.; and up-front randomisation may be preferable to delayed randomisation for recruitment purposes.

With respect to trial conduct changes, a telephone reminder to contact participants who did not respond to postal trial invitations was found to improve recruitment by RD= 6% (95% CI: 3% to 9%) with high GRADE certainty of evidence; this figure was obtained following a meta-analysis of two studies (Nystuen and Hagen, 2004, Wong et al., 2013). Nevertheless, the RD figure has low uncertainty when the baseline recruitment rate, i.e. proportion of eligible patients randomised, is less than 10%; otherwise, if the baseline recruitment rate is above 10%, the GRADE of evidence is downgraded to moderate. According to single studies, and with moderate GRADE of evidence, several trial conduct modifications may enhance recruitment, such as quotes from previous participants in SMS messages and highlighting shortage of trial places in SMS messages.

With respect to modifications to the consent process, five out of eight included related studies were disregarded from reporting results as they were assessed to have a high risk of bias. The main finding in this category is that opt-out consent forms are likely to increase recruitment by RD=19% (95% CI: 3% to 35%) (Trevena et al., 2006).

With respect to modifications to the information given to potential participants about the trial, a bespoke process that incorporates formal user-testing for improving the participant information leaflet (PIL) makes a small or no improvement in recruitment rates [RD = 1% (95% CI = -1% to 3%)], with high GRADE certainty of evidence (Cockayne et al., 2017, Man

et al., 2015). Results were reported for further 17 information modification-related recruitment interventions. Findings from the review indicated that, with a moderate GRADE of evidence and a low risk of bias in the corresponding included study, the use of a brief patient information leaflet (PIL) compared to a lengthier standard PIL may have mixed effects on recruitment rates. Additionally, including a questionnaire that focuses on trial issues alongside the invitation could potentially enhance recruitment. Furthermore, the inclusion of user feedback in PIL forms may slightly or may not improve recruitment rates. The use of a primer letter may or may not lead to improved recruitment. On the other hand, the provision of trial information via telephone does not appear to significantly improve recruitment. The results of the remaining interventions were not reported by the review authors, as the reviewers assessed the GRADE of evidence to be either low or very low and the risk of bias of the related studies to be unclear. With respect to interventions aimed at the recruiter or the recruitment site(s), using a postcard teaser campaign might slightly or might not improve recruitment, with moderate GRADE certainty of evidence. Finally, financial incentives appear to improve recruitment by RD = 4% (95% CI = -1% to 8%), with moderate GRADE certainty of evidence due to heterogeneity among the included studies (Free et al., 2010, Jennings et al., 2015).

4.1.1.3. Conclusions of the review

In summary, the reviewers concluded that an open design, compared to a placebo-controlled design, and using telephone reminders to non-responders are effective interventions to improve recruitment to RCTs; these findings are judged to be generalisable. In addition, a bespoke process incorporating formal user-testing for improving the participant information leaflet (PIL) was found to make a small or no improvement in recruitment rates, with high GRADE certainty of evidence. Whereas the review includes results for many potential recruitment interventions, the uncertainty regarding their generalisability in improving trial efficiency originates from the conduct of single studies exploring different potential recruitment interventions. Thus, the reviewers suggested that *"the research community...establish a process for prioritising which recruitment interventions are most in need of evaluation*" (Treweek et al., 2018b) , and further recruitment trials, related to interventions associated with a moderate GRADE of evidence (e.g. brief PIL vs lengthier PIL, postcard teaser campaign, etc.), be undertaken (Treweek et al., 2018b). Most importantly, as the results of the recruitment interventions with high GRADE of evidence were based upon SWATs, the authors recommend that future recruitment interventions be assessed with SWAT designs and that Trial Forge be informed of such recruitment trials for optimising coordination and research efficiency. The

priority interventions are the following: 1) telephone reminders for host trials having a reported recruitment of more than 10%; 2) moderate financial incentives (i.e. significantly less than £100 per participant recruited (Free et al., 2010)) to reduce heterogeneity among studies; 3) text messaging interventions which had a moderate GRADE certainty of evidence. Finally, the conduct and development of recruitment interventions for paediatric trials in the future is highly encouraged, as the difficulties of recruitment to paediatric trials are differential in nature to other trials because of the additional ethical requirements.

4.1.2. Cochrane systematic review of retention strategies (Gillies et al., 2021)

The latest Cochrane review focusing on retention aimed to *"quantify the effect of strategies to improve retention of participants in randomised trials"* (Gillies et al., 2021).

4.1.2.1. Methods

The study selection process focused on participant randomised and quasi-randomised trials involving patients, members of the public, healthcare professionals, etc., in order to appraise the existing evidence on the effectiveness of strategies for improving retention in RCTs. There were no restrictions with respect to the age, gender, ethnicity, language, and geographic group of the participants of host trials. The trial setting was not constrained to healthcare settings only since education and social sciences topics were additionally considered. However, for inclusion purposes, the outcomes of the host trials were health related or clinical. In contrast to the Cochrane recruitment review, hypothetical trials were excluded from data extraction, as such trials were assessed to have a high Cochrane risk of bias (Treweek et al., 2018b). The interventions were any strategies designed to improve retention of participants in RCTs. Related retention trials could include randomised comparisons of two or more retention strategies, or randomised comparisons of one or more retention strategies with a usual procedure. Retention interventions could be related to data collection, participant strategies, sites and staff, central study management, and study design. The primary outcome was the proportion of participants retained at the primary analysis point; if unspecified, the first time point of the statistical analysis of the retention trial was considered, or, if multiple points were reported, the data for the time point closest to the retention intervention were collected.

To screen the relevant studies, the reviewers identified trial retention studies published up to the end of December 2017 from ORRCA. This is a helpful database including a wide range of electronic databases, in order to capture studies evaluating strategies to improve recruitment

and retention to RCTs. For studies published from 2018 and onwards, the reviewers applied a search strategy to electronic databases, such as Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE (OVID), CINAHL, PsycINFO, SCOPUS, and Web of Science Core Collection, with records searched up to January 2020. In addition, they searched the SWAT repository to capture retention SWATs that were either ongoing or unpublished. For the host trial, the data extraction form provided information about the design, location, setting, population, intervention, and comparator. For the retention SWAT, the data extraction provided information about the type of design (i.e. randomised, or quasi-randomised trial), design, aim, definition of retention used, retention period, the source of the retention trial sample, and participant characteristics. For the retention strategy, the data extraction provided information about the type, whether it is theoretically based, the frequency, mode of delivery, co-interventions, economic information, resource requirements, and the numbers and proportions of participants in the intervention and control groups of the retention SWAT. The Cochrane risk of bias tool was applied to assess the quality of the included studies. For any retention strategy, the GRADE tool was applied to assess the quality of its cumulated evidence, in a similar fashion to the Cochrane recruitment review. Studies reported to have a high risk of bias were not excluded from the data analysis. The RD and 95% CIs were reported for retention strategies. In the case of cluster RCTs, the standard error was inflated to derive an adjusted standard error. Heterogeneity across the included trials relevant to a specific retention strategy was assessed through the Chi^2 statistic, the I^2 statistic and subgroup analyses. The latter considered the type of design used, the host trial setting, the disease area, the follow-up duration, and the amount of the financial incentive. The data synthesis distinguished trials into the following categories, in line with the ORRCA project: data collection, participants, sites and site staff, central study management, and study design. All results were reported in 95% confidence intervals associated with fixed-effect and random-effect RD. Results were not pooled if heterogeneity was substantial. Overall, the methodology of this review tends to be similar to the methodology of the most recent Cochrane recruitment review (Treweek et al., 2018b); the key methodological differences are that the retention review excluded hypothetical trials, and that it did not pool results when heterogeneity was unexplainable.

4.1.2.2. Results

In total, 18,756 titles and abstracts (18,655 from electronic databases, 76 from ORRCA database, and 25 from SWAT repository) were screened, of which 150 were eligible for full text, given the inclusion criteria. Eventually, 72 studies were included; these were undertaken

in eight countries, but most frequently in the USA and the UK. 70 studies were designed for trial participants, whereas the remaining two studies were associated with individuals involved in trial retention. The trial settings of these studies were broad and mainly clinical. The Cochrane risk of bias was assessed to be low for 14 studies, high for eight studies and unclear for 50 studies. The reviewers obtained 52 comparisons that included no study with a high risk of bias.

With respect to data collection, relevant findings were obtained from 14 low or uncertain risk of bias studies with 9 comparisons and n=35,215 participants. No findings for any associated retention intervention were yielded with high GRADE certainty of evidence. The review found that using a diary versus no diary for follow-up may reduce retention by -3% (95% CI: -4% to -2%), with moderate GRADE certainty of evidence (Griffin et al., 2019, Marques et al., 2013). Another finding, with moderate GRADE certainty of evidence, is that self-sampling kits (sent directly by post or proposed via invitation to order) may increase retention by 9% (95% CI: 4% to 13%) (Tranberg et al., 2018); however, given the limited applicability of such a strategy to most trials, the reviewers did not consider it a priority retention strategy Several results were reported for alternative retention strategies; however, their GRADE certainty of evidence was either low or very low due to single associated studies or multiple studies with unclear or high risk of bias.

With respect to retention strategies targeting recruited trial participants, the GRADE certainty of evidence for 35 interventions was either low or very low; hence no generalisable conclusions can be made from the findings. Reminders was one sub-category of such interventions, with electronic reminders having a highly uncertain RD of 1 % (95% CI: -4% to 6%) (Starr et al., 2015, Ashby et al., 2011, Keding et al., 2016), personalised reminders having a highly uncertain RD of -1% (95% CI: -11% to 8%) (Nakash, 2007) and telephone reminders also having a highly uncertain RD of -1% (95% CI: -18% to 15%) (Severi et al., 2011). Prompts was another sub-category of such interventions, with electronic prompts versus no prompts having a very low GRADE certainty of evidence and RD of 2% (95% CI: -1% to 6%) (Clark et al., 2015, Bradshaw et al., 2020, Keding et al., 2016, Man et al., 2011, Starr et al., 2015), telephone prompts having RD of 1% (95% CI: -10% to 12%) (Edwards et al., 2016, MacLennan et al., 2014) and prenotification cards having RD of 3% (95% CI: -3% to 10%) (Treweek et al., 2021). Similar uncertainty persists on monetary incentives, non-monetary incentives, participant engagement and behavioural interventions. However, the results from

the introduction of a monetary reward to both host trial intervention and control groups alongside with a cover letter are more certain. The RD is 9% (95% CI: 3% to 15%), with a moderate GRADE certainty of evidence (Hardy et al., 2016).

With respect to retention interventions targeting sites and site staff or central study management, only three studies explored comparisons, with prompts being uncertain with a RD of -3% (95% CI: -13% to 7%) (Land et al., 2007), monitoring visits being uncertain with a RD of -5% (95% CI: -20% to 10%) (Liénard et al., 2006) and a peer led follow-up intervention having a positive RD of 22% (95% CI: 14% to 30%), but a low GRADE certainty of evidence (Fouad et al., 2014). Finally, five studies considering two comparisons explored potential retention interventions targeting study design. Whilst the uncertainty of the evidence persists, giving a pen during recruitment may improve retention rates, with a moderate GRADE certainty of evidence and a positive RD of 20% (95% CI: 7% to 32%) (Whiteside et al., 2019).

4.1.2.3. Conclusions of the review

In summary, there is limited evidence to recommend an effective strategy improving retention in RCTs. With moderate GRADE certainty of evidence, adding a pen to recruitment, adding self-sampling kits to follow-up, and introducing a monetary reward to both host trial intervention and control groups, alongside a cover letter, can increase retention in RCTs. In addition, adding a diary to follow-up may reduce retention with moderate GRADE certainty of evidence. Similar to the recruitment review, the uncertainty of the evidence primarily originates from the assessment of many potential retention interventions from single studies, with no existing replications for 33 out of 51 comparisons. Nevertheless, nearly 50% of the included studies were published recently, i.e. from 2016 to 2020, thus signalling the key role of SWATs in gathering evidence to identify effective retention strategies. Similar to the Cochrane recruitment review, the reviewers summarised several retention strategies that could be prioritised for replication. These are electronic reminders versus usual follow-up; personalised prompts versus no prompt; adding a pen versus no pen; adding monetary rewards to all trial arms; adding a social incentive to a cover letter versus a standard cover letter; personalising a card versus no card; using a telephone reminder versus usual follow-up; comparing an electronic reminder with an electronic prompts; introducing a societal benefit message versus normal follow-up; adding a pen during recruitment versus no pen; and using self-sampling kits versus no self-sampling kits. It is believed that for retention strategies with a moderate GRADE certainty of evidence, a single replication might improve their overall GRADE certainty of

evidence. The authors also encourage further SWATs associated with patient and public involvement (PPI) interventions in case a trial team believes the effects of such interventions can outweigh their costs, as PPI consists one of the key unanswered research questions about trial retention (Brunsdon et al., 2019) Moreover, the reviewers viewed the reporting of many included trials to be poor, due to lapses of CONSORT diagrams, unclear primary outcomes, unclear sample sizes and invariable participant sociodemographic characteristics (e.g. findings came mainly from high income countries and white participants). Thus, considering these reporting aspects is important for deciding whether further trials of a particular retention intervention need to be undertaken in different contexts. Finally, the reviewers assert that *"future evaluations (of any retention intervention) should also consider economic evaluation"* (Gillies et al., 2021) .

4.2. A systematic review of economic evaluations alongside Studies Within a Trial (SWATs) for improving recruitment and retention in RCTs

Note: This study is published in a peer-reviewed journal (Gkekas et al., 2023). The work presented is identical to the published study, including the supplemental material.

4.2.1. Abstract

Aim: To review the cost-effectiveness of strategies to improve participant recruitment and retention in randomised controlled trials.

Methods: All included studies from the latest Cochrane recruitment and retention reviews were considered (Treweek et al., 2018b, Gillies et al., 2021). To identify articles published since the Cochrane reviews, electronic databases were searched until March 2021. Hand searching of conference databases and journals was also undertaken. The inclusion criteria included Studies within a Trial (SWATs). The main outcome was the incremental cost-effectiveness ratio (ICER). Quality assessment of papers used the Cochrane risk of bias 1 tool. The CRD guidance was used to assess the quality of economic evaluation. Random-effect meta-analyses were undertaken. The GRADE certainty of evidence was applied for each strategy, and Trial Forge Guidance 2 was used for strategies included in meta-analyses to evaluate the uncertainty of the findings. Cost-effectiveness ranks summarises the cost-effectiveness of all strategies.

Results: We identified 6569 records and included 29 SWATs (earliest conducted in 1999 and latest in 2021) including more than 35800 participants. There is no strategy which we would recommend trial teams and researchers adopt with complete statistical certainty. Recruitment strategies which could be cost-effective include financial incentives, trial-branded pens, telephone reminders and pre-notification leaflets. Retention strategies which could be cost-effective include vouchers and trial-branded pens.

Conclusion: Future SWATs should replicate existing recruitment and retention strategies, rather than evaluate novel ones. We recommend that economic evaluations be carried out alongside all future SWATs, costs and benefits be recorded transparently, and the cost-effectiveness of existing recruitment or retention strategies be evaluated.

4.2.2. Introduction

Recruitment of participants into randomised controlled trials (RCTs) is usually poor (Fletcher et al., 2012). Under-recruited, and hence under-powered, trials result in research waste (Treweek et al., 2018b). Another main challenge with RCTs is attrition, when recruited participants fail to complete follow-up assessments. A systematic review of 151 trials associated with the National Institute for Health & Care Research (NIHR) Health Technology Assessment (HTA) Programme has found the median retention rates to be 89% (Walters et al., 2017). Poor retention not only diminishes the power of the trial but also can introduce attrition bias, thus threatening the statistical analysis of RCTs (Torgerson and Torgerson, 2008).

Studies Within A Trial (SWATs) are a study design for identifying strategies to improve recruitment and retention in RCTs (Bower et al., 2014, Treweek et al., 2018a). SWATs' primary objective is to improve trial methodology and efficiency Treweek et al. (2018a). Two systematic reviews have appraised the evidence on the effectiveness of strategies for improving recruitment and retention in RCTs (Treweek et al., 2018b, Gillies et al., 2021). Both reviews have implied poor progress on identifying effective recruitment/retention strategies.

Moreover, no appraisal of the evidence on the *cost-effectiveness* of recruitment and retention strategies has been undertaken so far. Given the anticipated high direct and indirect costs of poor recruitment and retention rates (McDonald et al., 2006, Kitterman et al., 2011), economic evaluations of recruitment and retention strategies alongside SWATs are useful. More broadly, an "*economic evaluation offers an organised consideration of the range of possible alternative courses of action and the evidence of the possible effects of each. This is more likely to lead to better decisions that improve overall social value*" (Drummond et al., 2015). There is an urgent need to strengthen the evidence arising from SWATs towards identifying strategies that could improve recruitment and retention in RCTs. There is a further need to develop a framework for the economic analysis of SWATs to enable research organisations to make more informed decisions.

This review accumulates and critically appraises the existing evidence on economic evaluations alongside SWATs for improving recruitment and retention in SWATs. The primary aim is to improve trial efficiency by increasing the evidence available for making trial process

decisions. The secondary aim is to make recommendations for improvement of future economic evaluations alongside recruitment and retention SWATs.

4.2.3. Methods

A protocol is registered on PROSPERO (CRD record code: 42021236824), in line with the 2020 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) guidelines (Page et al., 2021).

The studies eligible for inclusion were quasi- or fully randomised SWATs. The corresponding host trials of SWATs had to be quasi- or fully randomised, and within the context of healthcare or any field applicable to healthcare settings. Hypothetical studies (i.e. studies that ask potential patients whether they would participate in a trial that will not take place in reality) were excluded as these were assessed to have a high Cochrane risk of bias (Treweek et al., 2018b), and hence were also excluded from the retention review (Gillies et al., 2021).

Any strategies designed to improve recruitment and/or retention of participants in RCTs were eligible for inclusion in the study. The target population was any potentially eligible trial participants. For SWATs associated with improving recruitment, the strategies were aimed at potential trial participants who could be recruited to a host trial. For SWATs associated with improving retention, the strategies were aimed at already randomised trial participants who were asked to provide follow-up data. In contrast to the Cochrane recruitment and retention reviews, strategies aimed at collaborators or research ethics committees were not considered (Treweek et al., 2018b, Gillies et al., 2021). There were no restrictions regarding comparators. There are several potential types of economic evaluation alongside SWATs, including cost-effectiveness analysis, cost-benefit analysis, cost-consequence analysis, and cost-utility analysis. Therefore, the primary outcome was reported in terms of the incremental cost-effectiveness ratio (i.e. the incremental cost per additional patient recruited or per additional participant retained), the (monetary) net benefit of a given strategy or the willingness to pay (WTP) for a given strategy. The secondary outcomes were any costs and health utilities (benefits) of recruitment/retention strategies. The measures of effect could be reported as incremental/unit/total costs, or incremental utilities/effects/benefits expressed in recruitment or retention rates. If the primary outcome in a study was unavailable but its secondary outcomes were reported appropriately, such a study would not be excluded on these grounds.

As this review focused on SWATs of recruitment or retention strategies that included economic evaluations in their analyses, the SWATs from the most recent Cochrane recruitment and retention reviews were considered in the study selection process (Treweek et al., 2018b, Gillies et al., 2021). Further potential SWATs were identified after the final dates of the study searches in these Cochrane reviews, i.e. on and after 12 February 2015 until 3 March 2021 for recruitment strategies (Treweek et al., 2018b), and on and after 1 March 2020 until 3 March 2021 for retention strategies (Gillies et al., 2021). Thus, we developed a search strategy for the identification of more recent SWATs on recruitment and/or retention strategies that involved economic evaluation. The search strategy is available in *Supplemental Material 4.2*.

We searched the following electronic databases:
• MEDLINE (OVID)
• Embase (OVID)
• CINAHL
• Cochrane Methodology Review Group Specialised Register (CMR) in the Cochrane Library
• Science Citation Index and Social Citation Index
• ERIC (EBSCO)
• PsycINFO (OVID)
• Scopus

Hand searching of conference abstracts associated with SWATs was also undertaken. Journals were also hand searched, including ClinicalTrials.gov, OpenTrials, EU Clinical Trials Register, Cochrane Central Register of Controlled Trials (CENTRAL) and the Online Resource for Recruitment research in Clinical triAls (ORRCA). The search dates were the same as those for the electronic databases.

The titles, abstracts, and full texts of identified records were independently screened by two authors (AG and AE). We independently extracted the data through a standardised data extraction form on Microsoft Word, which included information on both the host trial and the SWAT. The data extraction form provided information about the host trial name, design, location, clinical setting, population, intervention(s), and comparator(s) (*Supplemental Material 4.1*). With respect to SWATs, the data extraction form provided information about the design, the strategy(-ies), the comparator(s), study objective, time horizon, frequency and

timing of strategy, measure(s) of benefit and costs, type of economic evaluation, numbers, and proportions of participants in the intervention and control groups, results of health economic outcomes in the intervention and control groups, and perspective adopted in the economic evaluation (*Supplemental Material 4.1*). No automation tools were applied in the data collection process. We did not need to contact any study investigators to obtain further data or to ask for clarification of published data.

In line with the Cochrane recruitment and retention reviews' methodology, the Version 1 Cochrane risk of bias tool was used to appraise the quality of the included studies (Higgins et al., 2011) .For included studies that were obtained from the recruitment and retention reviews, the risk of bias presented in these reviews was assumed to be valid, and hence their quality appraisal was adapted from the corresponding reviews to avoid duplication of effort (Treweek et al., 2018b, Gillies et al., 2021). In addition, a descriptive quality assessment was independently undertaken to assess the quality of the included SWATs (including those from the Cochrane reviews) with respect to their economic evaluation. Such assessment followed explicitly the University of York's Centre for Reviews and Dissemination (CRD) guidance on systematic reviews of economic evaluations by considering the following: methods of deriving the effectiveness data, cost analysis, valuation and measurement of health benefit, methods of synthesising the costs and effects, and, if applicable, analysis of uncertainty (Tacconelli, 2010). To allow for broader inclusion of studies associated with economic evaluations alongside SWATs, studies with a high risk of bias, or studies with a low quality of economic evaluation, or studies that were not peer-reviewed, were still included if they met the inclusion criteria.

Random-effect meta-analyses, through the Cochran-Menthel-Haenszel weighting method, were undertaken for the primary outcome. Since all included SWATs were associated with cost-effectiveness analysis, this was carried out by initially obtaining the odds ratios (ORs) of the recruitment or retention rates for each strategy. Then, the ORs were converted to effect sizes, i.e. incremental recruitment or retention rates, by dividing the natural logarithm of the OR with 1.81 (Chinn, 2000). This conversion is assumed for both continuous and dichotomous outcomes (Chinn, 2000), and hence we applied such a conversion in our study.

(i.e. effect size= incremental recruitment rate/retention rate= ln(OR)/1.81).

The Cochran-Menthel-Haenszel method was applied for weighting the incremental costs of each strategy from each included study, to obtain the aggregate figure for the incremental cost of each strategy. The final step was to calculate the incremental cost per patient recruited or participant retained for each strategy by dividing the incremental cost with the incremental recruitment or retention rate. 95% confidence intervals are presented for the primary outcome, the OR, and the incremental recruitment or retention rate. RevMan was the software used for meta-analysis (Nordic Cochrane Centre, 2014). The figures for the primary outcome were adjusted to 2019 USD Purchasing Power Parity (PPP) rates. The use of PPP, defined by the International Monetary Fund (IMF) as "*the rate at which the currency of one country would have to be converted into that of another country to buy the same amount of goods and services in each country*" (Callen, 2012) can reflect more accurately any cost variations among countries. We anticipated the included studies to be potentially subject to between-group (study) heterogeneity; hence the I^2 statistic, ranging from 0% to 100%, was computed for all strategies, whose primary outcome was obtained through the inclusion of multiple SWATs. The greater a reported value of I^2, the greater the extent of between-study heterogeneity. More details about the meta-analyses of each recruitment and retention strategy can be found in *Supplemental Material 4.3* and *Supplemental Material 4.4* respectively.

Following meta-analysis, the GRADE approach was applied to the effect measure (i.e. the OR) and consequently the primary outcome (i.e. the ICER), to assess the certainty of the evidence for each recruitment and retention strategy; this tool explores the extents of risk of bias, imprecision, inconsistency, indirectness, and publication bias in the included studies. Such an assessment was undertaken by the two reviewers (AG and AE), and details about the GRADE assessment of each recruitment and retention strategy can be found in *Supplemental Material 4.5*. Furthermore, the Trial Forge Guidance 2 was explicitly used to qualitatively assess whether more SWATs should be conducted for recruitment and/or retention strategies included in the study's meta-analyses (Treweek et al., 2020). The assessment using Trial Forge Guidance 2 comprises of five criteria: risk of bias, imprecision, inconsistency, balance of benefit and disadvantage to participants, and balance of benefit and disadvantage to the host trial (Treweek et al., 2020). Such an assessment was undertaken by the two reviewers (AG and AE), and details about the application of Trial Forge Guidance 2 to each recruitment and retention strategy comprising of at least two SWATs can be found in *Supplemental Material 4.6*.

Cost-effectiveness ranks of strategies for improving recruitment and retention in RCTs are presented. The rank was based on the mean ICER of each recruitment or retention strategy, whether the lower 95% odds ratio confidence intervals indicate the associated strategy is significantly effective, and the GRADE certainty of evidence.

4.2.4. Results

4.2.4.1. Searching of records

The full texts of 68 studies from the recruitment review (Treweek et al., 2018b) and 71 studies from the retention review (Gillies et al., 2021) were assessed for inclusion to our study. Following the searches, 8113 records were retrieved from the electronic databases overall. 28 additional studies were identified from manually searched registers. Nine studies were identified from hand searching reference lists of two studies that were retrieved from the electronic databases. After deduplication, 6569 records were screened, and 267 full texts were assessed for eligibility. 22 studies were included in this review. A PRISMA flow diagram is shown in *Figure 4.1*.

4.2.4.2. Characteristics of the included studies

Three studies had more than one SWAT (Jennings et al., 2015, Free et al., 2010, Khadjesari et al., 2011), and hence 29 SWATs were included in this review. Nine studies (15 SWATs) assessed recruitment strategies, whereas 15 studies (16 SWATs) evaluated retention strategies. Two studies included SWATs with strategies targeting both recruitment and retention (Whiteside et al., 2019, Jolly et al., 2019). The characteristics of the included studies, i.e. study (author, date, country), host trial design, participants, SWAT intervention(s) (and comparator(s)), and SWAT outcome(s), are presented in *Table 4.1*.

All SWATs of retention strategies were already included in the retention review (Gillies et al., 2021). However, two of these SWATs had data that were publicly inaccessible at the time of the publication of the retention review (Cook et al., 2021, Dorling et al., 2020). These SWATs eventually became publicly accessible in journals. As there was an uncertainty regarding which data the reviewers from the retention review had accessed, the papers' risk of bias was re-assessed.

For recruitment strategies, four included studies (four SWATs) were retrieved from the electronic databases(Jolly et al., 2019, Bracken et al., 2019, Arundel et al., 2017, Rogers et al., 2019), one study (one SWAT) was retrieved from manual searching (Hancocks et al., 2019), four studies (10 SWATs) were already included in the recruitment review (Treweek et al., 2018b), and one study (one SWAT) was already included in the retention review (Khadjesari et al., 2011).

Most SWATs had individually randomised designs; however, two studies (two SWATs) were quasi-randomised (Miller et al., 1999, Gates et al., 2009), and two studies (two SWATs) were cluster randomised (Jolly et al., 2019) (Marsh and Kendrick, 1999).

Primary outcomes were available in seven out of 22 studies (11 out of 29 SWATs) and reported in terms of the incremental cost per additional patient recruited/ incremental cost per additional participant retained, respectively. Accordingly, the incremental cost-effectiveness ratio (ICER), for recruitment strategies, was defined as:

ICER= Incremental cost per additional patient recruited= $$\frac{Cost\ of\ recruitment\ strategy - cost\ of\ baseline\ strategy}{Recruitment\ rate\ of\ recruitment\ strategy - Recruitment\ rate\ of\ baseline\ strategy}$$

(Equation 4.1)

For retention strategies, the ICER was defined as:

ICER= Incremental cost per additional participant retained= $$\frac{Cost\ of\ retention\ strategy - cost\ of\ baseline\ strategy}{Retention\ rate\ of\ retention\ strategy - Retention\ rate\ of\ baseline\ strategy}$$

(Equation 4.2)

Therefore, cost-effectiveness analysis was the sole method of economic evaluation available alongside the included SWATs. The primary outcome was manually computed by the reviewers in the remaining 15 studies (18 SWATs), using the incremental costs and the incremental recruitment and/or retention rates of a given strategy to obtain the ICER. The perspective adopted by all economic evaluations related to the trial teams, i.e. the reported effects and costs of recruitment or retention strategies were direct and associated with the trial teams' budget. In total, 35864 participants from 29 SWATs were involved.

The Cochrane risk of bias was assessed to be high for four included studies, unclear for nine studies, and low for nine studies (*Table 4.2*). In terms of the quality of the economic evaluation, this was assessed to be low for seven studies, moderate for six studies, unclear for one study,

and high for eight studies (*Table 4.3*). One included study affected the quality appraisal detrimentally, as a full text was unavailable (Hancocks et al., 2019).

Figure 4.1: PRISMA Flow Diagram for the systematic review

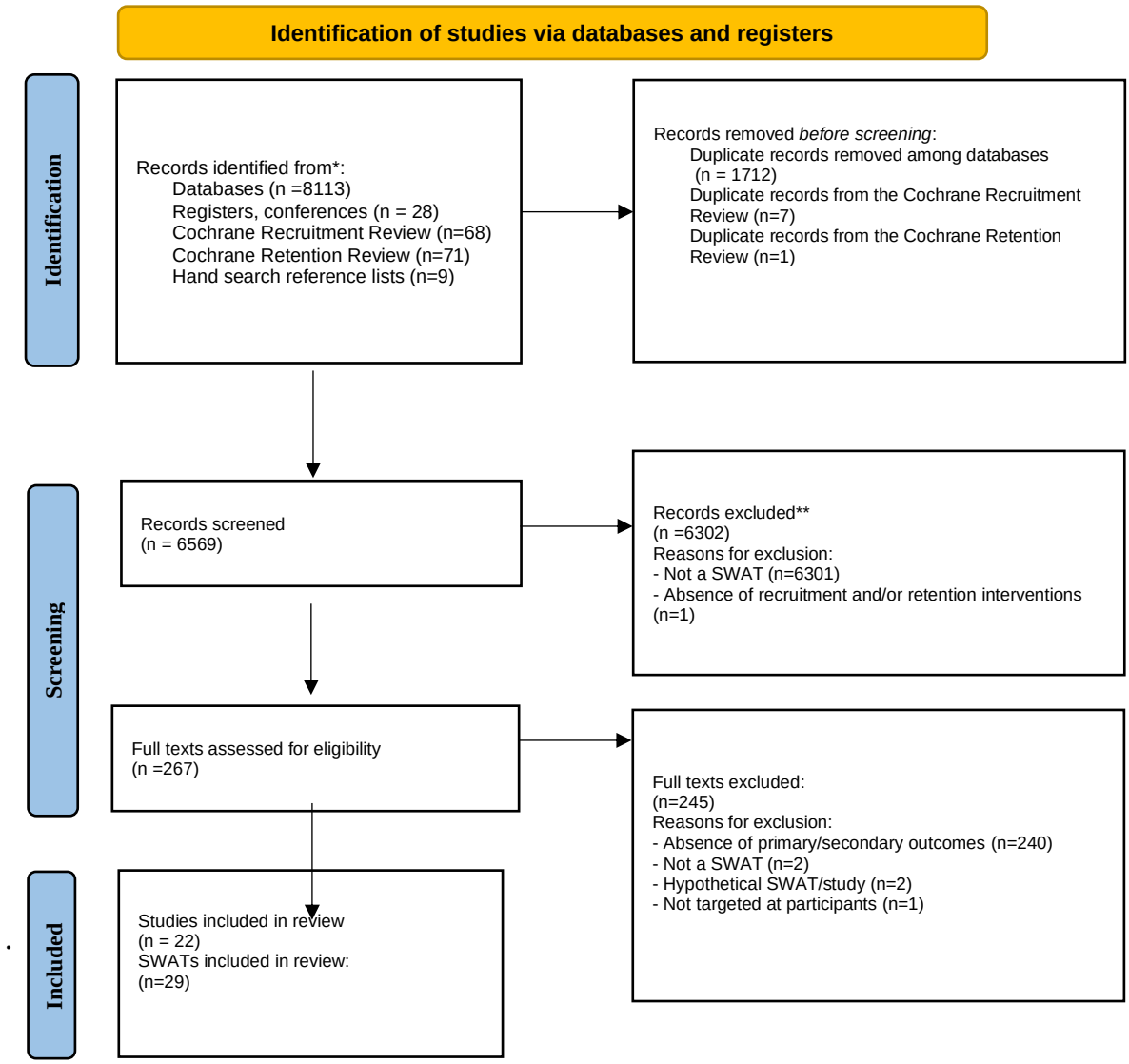


Table 4.1: Characteristics of the included studies

Study	Host trial design	Participants	SWAT Interventions	SWAT Outcome(s)
Jennings et al. (2015a) Country: United Kingdom	Prospective Randomised Open Blinded End point (PROBE) design.	People aged 60 or over taking long-term NSAIDS for arthritis	Intervention: Invitation letter with a fixed payment of £100. Comparator: An invitation letter with no fixed payment incentive.	Increase in consented patients with incentive.
Jennings et al. (2015b) Country: United Kingdom	Prospective Randomised Open Blinded End point (PROBE) design.	People aged 60 or over with chronic hyperuricaemia in conditions where	Intervention: Invitation letter with a fixed payment of £100. Comparator: An invitation letter with no fixed payment incentive.	Increase in consented patients with incentive.

		urate deposition has already occurred.		
Jennings et al. (2015c) Country: United Kingdom	Randomised Controlled Trial (RCT); open, parallel, double-blind	People aged 18-79 with newly diagnosed hypertension	Intervention: Invitation letter with a fixed payment of £100. Comparator: An invitation letter with no fixed payment incentive.	Increase in consented patients with incentive.
Jennings et al. (2015d) Country: United Kingdom	Randomised Controlled Trial (RCT); open, parallel, double-blind	People aged 18-79 with uncontrolled blood pressure.	Intervention: Invitation letter with a fixed payment of £100. Comparator: An invitation letter with no fixed payment incentive.	Increase in consented patients with incentive.
Jennings et al. (2015e) Country: United Kingdom	Randomised Controlled Trial (RCT); open, parallel, double-blind	People aged 18-80 with at least one component of the metabolic syndrome.	Intervention: Invitation letter with a fixed payment of £100. Comparator: An invitation letter with no fixed payment incentive.	Increase in consented patients with incentive.
Free et al. (2010a) Country: United Kingdom	Randomised Controlled Trial (RCT); pilot, single-blind	People aged 16 and above who are smokers and willing to stop smoking in next month.	Intervention: Research staff sending a text message regarding the newly available online registration facility. Comparator: Research staff calling the participants' mobile numbers to register them for the trial (no text message).	Consent to be randomised into the Txt2stop trial (i.e. host trial) within 2 weeks.
Free et al. (2010b) Country: United Kingdom	Randomised Controlled Trial (RCT); pilot, single-blind	People aged 16 and above who are smokers and willing to stop smoking in next month.	Intervention: Letter containing study and consent information and a £5 note. Comparator: Participants received the normal trial procedures.	Consent to be randomised into the Txt2stop trial (i.e. host trial) within 2 weeks.
Free et al. (2010c) Country: United Kingdom	Randomised Controlled Trial (RCT); pilot, single-blind	People aged 16 and above who are smokers and willing to stop smoking in next month.	Intervention: Four text messages over one week containing quotes from existing participants Comparator: No text messages.	Consent to be randomised into the Txt2stop trial (i.e. host trial) within 2 weeks.
Miller et al. (1999) Country: United States	Two Randomised Controlled Trials (RCTs). No further information about their designs.	Participants aged 18-75 with DSM-IV dysthymic disorder, double depression (major depression superimposed on antecedent dysthymia), or chronic major depression.	Intervention: Phone screening by a Senior Investigator. Comparator: Phone screening by a trained Research Assistant.	Proportion of participants recruited to the two host trials.
Bell et al. (2016) Country: United Kingdom	Randomised Controlled Trial (RCT); pragmatic, unblinded, two-arm, parallel	Females aged 70-85 who are not currently on prescription medication to prevent osteoporotic fractures before randomisation.	Intervention: Trial-branded pen with the 60-month follow-up questionnaire. Comparator: 60-month follow-up questionnaire alone	Questionnaire return rate.
Cunningham-Burley et al. (2020) Country: United Kingdom	Randomised Controlled Trial (RCT); two-arm, 1:1 randomisation	NHS staff who are subject to a Trust dress code	Intervention: Trial-branded pen with the 14-week follow-up questionnaire. Comparator: 14-week follow-up questionnaire alone.	Questionnaire return rate.
Clark et al. (2015) Country: United Kingdom	Randomised Controlled Trial (RCT); two-arm, 1:1 block randomisation	Smokers who are aged 35 or more, who are invited to undertake a series of case-finding tools,	Intervention: Electronic prompt (i.e. SMS or e-mail) to return the study questionnaire. Comparator: No prompt.	Questionnaire return rate.

		which comprise lung function tests and several symptom based case-finding questionnaires, for the potential identification of COPD.		
Cochrane et al. (2020) Country: United Kingdom	Cohort, pragmatic, two-arm, open Randomised Controlled Trial (RCT)	People aged 65 or older who either have experienced a fall in the last year or feel worried about falling in the future.	Intervention: Personalised text message, as a retention strategy. Comparator: Generalised text message	Questionnaire return rate.
Hardy et al. (2016) Country: United Kingdom	Multicentre Randomised Controlled Trial (RCT).	Adult women who are nulliparous, have a single cephalic presentation, greater than or equal to 37 weeks' gestation, intend spontaneous vaginal birth, are in second stage of labour and with an effective mobile epidural in situ.	Interventions: 1) an incentive cover letter with a promise of a £10 gift voucher on the return of a completed questionnaire. The covering letter thanked participants for their time and effort. All reminder letters included a sentence about the incentive. Comparator: no incentive mentioned in the cover letter. If the questionnaire was not returned, all reminder letters included the promise of a £10 gift voucher on the return of a completed questionnaire.	Questionnaire return rate.
Gates et al. (2009) Country: United Kingdom	Cluster Randomised Controlled Trial (RCT)	Participants who attended Emergency Departments (EDs) with an acute whiplash injury of whiplash-associated disorder grades I-III were eligible for Step 1. People who attended EDs with whiplash injuries and had persistent symptoms 3 weeks after ED attendance were eligible for Step 2.	Intervention: £5 gift voucher, redeemable at a range of shops with their questionnaire, and a covering letter thanking participants for their time and effort. Comparator: No gift voucher, and a standard covering letter.	Questionnaire return rate.
James et al. (2020) Country: United Kingdom	Cohort, pragmatic, two-arm, open Randomised Controlled Trial (RCT)	People aged 65 or older who either have experienced a fall in the last year or feel worried about falling in the future.	Interventions: 1) a branded pen and a standard cover letter. 2) a branded pen and a social incentive cover letter. 3) no pen and a social incentive cover letter. Comparator: no pen and a standard cover letter.	Questionnaire return rate.
Kenyon et al. (2005) Country: United Kingdom	Cohort, pragmatic, two-arm, open Randomised Controlled Trial (RCT)	Children whose mothers joined the MRC ORACLE Trial. Their mothers	Intervention: monetary incentive (£5 voucher redeemable at high street stores) together with the reminder questionnaire.	Questionnaire return rate.

		have had preterm, prelabour rupture of the fetal membranes (pROM). The parents of the survived children are the participants in the SWAT.	Comparator: No monetary incentive. The same reminder questionnaire was sent.	
Khadjesari et al. (2011a) Country: United Kingdom	Randomised Controlled Trial (RCT); 2-arm, randomisation stratified by age and gender.	People who visited DownYourDrink while browsing the web, and who had an AUDIT-C score greater than 5.	Intervention: Participants who did not complete the first follow-up questionnaire within 1 week received either a £5 Amazon.co.uk voucher, £5 donation to Cancer Research UK, or entry in a £250 prize draw in the second prompt for completion of questionnaires. Comparator: No incentive. The participants received another prompt for completion of questionnaires.	Questionnaire return rate.
Khadjesari et al. (2011b) Country: United Kingdom	Randomised Controlled Trial (RCT); 2-arm, randomisation stratified by age and gender.	People who visited DownYourDrink while browsing the web, and who had an AUDIT-C score greater than 5.	Intervention: Participants received a £10 Amazon.co.uk voucher in the first prompt for completion of questionnaires. Comparator: No incentive. The participants received another prompt for completion of questionnaires.	Questionnaire return rate.
Marsh et al. (1999) Country: United Kingdom	Cluster Randomised Controlled Trial (RCT)	Parents of children aged 3-12 months registered with 36 participating practices in Nottingham.	Intervention: 1) postal administration with financial incentive (£2 voucher) once the completed diary had been received or postal group without financial incentive. 2) telephone administration with financial incentive (£2 voucher) once a completed diary had been received or telephone group without financial incentive. Comparator: no postal or telephone administration; either with or without financial incentive.	Diary return rate.
Treweek et al. (2021) Country: United Kingdom	Randomised Controlled Trial (RCT); four-centre, 1:1 parallel group	Women aged 50-70 who are overweight and attending routine breast screening in four Scottish breast screening service centres.	Intervention: pre-notification card Comparator: no pre-notification card	Proportion of participants attending the host trial primary outcome measurement visit (i.e. retention).
Whiteside et al. (2019) Country: United Kingdom	Cohort, pragmatic, two-arm, open Randomised Controlled Trial (RCT)	People aged 65 or older who either have experienced a fall in the last year or feel worried about falling in the future.	Intervention; branded pen with trial invitation pack Comparator: no pen, but with trial invitation pack	1) Randomisation rate 2) Proportion of participants who remained in the trial at 3 months post randomisation,i.e. retention.

Jolly et al. (2019) Country: United Kingdom	Pragmatic, multicentre Randomised Controlled Trial (RCT)	People aged 18 or older who are on the practice COPD register and have mild dyspnoea.	Interventions: The practices recruiting participants for the host trial accessed standard printed patient information materials, as well as a multimedia information resource, developed by patient and public involvement (PPI) contributors and researchers. Comparator: The practices accessed standard printed patient information materials, with no extra multimedia information resource.	1) Recruitment rate 2) Retention rate (after 6 months) 3) Retention rate (after 12 months)
Bracken et al. (2019) Country: Australia	Randomised Controlled Trial (RCT): multicentre, double-blind	Men aged 50-74 years, obese or overweight, with prediabetes or newly diagnosed type 2 diabetes, and a low serum testosterone.	Intervention: SMS reminder text which provided key enrolment information as well as including a peripheral cue based on the concept of social proof. Comparator: Telephone reminder, with a reminder call script used by staff members.	Attendance rate.
Arundel et al. (2017) Country: United Kingdom	Cohort Randomised Controlled Trial (cRCT); two-arm, pragmatic, open, multicentre	Patients aged 65 years and over who have attended routine podiatry services within the past 6 months, have had one fall in the past 12 months; or one fall in the past 24 months requiring hospital attention; or report a fear of falling on their baseline questionnaire in the past 4 weeks.	Intervention: A pre-notification leaflet, 2–3 weeks before the trial recruitment pack. Comparator: Trial recruitment pack only	Randomisation rate.
Rogers et al. (2019) Country: United Kingdom	Randomised Controlled Trial (RCT): prospective, open-label, blinded	People aged 60 or over, taking allopurinol for chronic gout, and with additional cardiovascular risk factors.	Intervention: DVD presentation containing an audio-visual presentation explaining the background and operation of FAST, and a standard invitation pack. Comparator: Standard invitation pack only.	Randomisation rate.
Hancocks et al. (2019) Country: United Kingdom	Randomised Controlled Trial (RCT): pragmatic, multicentred, parallel, two group	People aged 18 or over, who are smokers and smoke at least 10 cigarettes per day (for at least one year).	Interventions: 1) Full invitation pack from a GP. 2) Single-page invitation from a GP. Comparator: Text message invitation	Recruitment rate.
Cook et al. (2021) 15 European countries	Randomised Controlled Trial (RCT): pragmatic, multicountry, adaptive, two group, phase IV	People aged 1 or over, who have sudden onset of self-reported fever, with at least one respiratory symptom (cough, sore throat, running or congested nose)	Intervention: Unconditional monetary incentive of £20 given to participants at recruitment, as an intervention to boost retention in the host trial. Comparator: Conditional monetary incentive of £20 given to participants only once a questionnaire had been returned.	Diary return rate.

		and one systemic symptom (headache, muscle ache, sweats or chills or tiredness) during a period of increased influenza activity.		
Dorling et al. (2019) Countries: Ireland, United Kingdom	Randomised Controlled Trial (RCT): parallel, multicentre, two group	Infants born at <32 weeks' gestation or a weight of <1500g, who were receiving <30 ml/kg/day of milk at trial enrolment.	Intervention: The first paper follow-up letter to parents would include a promise of an incentive (£15 (€15 for Irish residents) gift voucher redeemable at high-street shops) after receipt of a completed form. Comparator: The first paper letter to parents would enclose the incentive (£15 (€15 for Irish residents) gift voucher redeemable at high-street shops) before the receipt of a completed form.	Questionnaire return rate.

Table 4.2: Cochrane risk of bias in the included studies

Study	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other bias	Overall risk of bias
Jennings (2015a) *	+	+	+	+	+	+	+	+
Jennings (2015b) *	+	+	+	+	+	+	+	+
Jennings (2015c) *	+	+	+	+	+	+	+	+
Jennings (2015d) *	+	+	+	+	+	+	+	+
Jennings (2015e) *	+	+	+	+	+	+	+	+
Free (2010a) *	+	+	+	+	+	+	+	+
Free (2010b) *	+	+	+	+	+	+	+	+
Free (2010c *)	+	+	+	+	+	+	+	+
Miller (1999) *	-	-	?	+	+	+	+	-
Bell (2016) **	?	+	+	+	+	+	+	?
Cunningham-Burley (2020) **	+	?	+	+	+	+	+	?
Clark (2015) **	?	+	+	+	+	+	+	?
Cochrane (2020) **	+	+	+	+	+	+	+	+
Gates (2009) **	-	-	+	+	-	?	?	-
Hardy (2016) **	+	+	+	+	+	+	+	+
James (2020) **	+	+	+	+	+	+	+	+
Kenyon (2005) **	+	?	+	+	+	+	+	?
Khadjesari (2011a) **	+	?	?	?	?	+	+	?
Khadjesari (2011b) **	+	?	?	?	?	+	+	?
Marsh (1999) **	-	?	+	+	?	?	+	-
Treweek (2021) **	+	+	+	+	+	+	+	+
Whiteside (2019) **	+	+	+	+	+	+	+	+
Jolly (2019)	+	+	?	?	?	+	+	?
Bracken (2019)	+	+	+	+	+	+	+	+
Arundel (2017)	+	+	+	+	+	+	+	+
Rogers (2019)	+	?	-	-	?	+	+	-
Hancocks (2019)	+	?	?	?	?	?	?	?
Cook (2021)	+	?	+	+	+	+	+	?
Dorling (2020)	+	?	+	+	+	+	+	?

Table 4.3: Quality of economic evaluation in the included studies

Study	Reliable derivation of effectiveness data	Reliable cost analysis	Reliable valuation and measurement of benefits	Reliable cost and benefit synthesis	Reliable analysis of uncertainty	Ec. Evaluation Quality
Jennings (2015a)	+	?	+	?	N/A	Moderate
Jennings (2015b)	+	?	+	?	N/A	Moderate
Jennings (2015c)	+	?	+	?	N/A	Moderate
Jennings (2015d)	+	?	+	?	N/A	Moderate
Jennings (2015e)	+	?	+	?	N/A	Moderate
Free (2010a)	+	+	+	?	N/A	Moderate
Free (2010b)	+	?	+	?	N/A	Moderate
Free (2010c)	+	+	+	?	N/A	Moderate
Miller (1999)	+	+	+	?	N/A	Moderate
Bell (2016)	+	+	+	+	N/A	High
Cunningham-Burley (2020)	+	+	+	+	N/A	High
Clark (2015)	+	+	+	-	N/A	Low
Cochrane (2020)	+	+	+	?	N/A	Moderate
Gates (2009)	+	+	+	+	N/A	High
Hardy (2016)	+	-	+	?	N/A	Low
James (2020)	+	+	+	+	N/A	High
Kenyon (2005)	+	+	-	-	N/A	Low
Khadjesari (2011a)	+	+	+	-	N/A	Low
Khadjesari (2011b)	+	+	+	-	N/A	Low
Marsh (1999)	+	+	+	+	N/A	High
Treweek (2021)	+	+	+	?	N/A	Moderate
Whiteside (2019)	+	+	+	+	N/A	High
Jolly (2019)	+	-	+	-	N/A	Low
Bracken (2019)	+	+	+	+	N/A	High
Arundel (2017)	+	+	+	+	N/A	High
Rogers (2019)	+	+	+	?	N/A	Moderate
Hancocks (2019)	?	?	?	?	N/A	Unclear
Cook (2021)	+	+	+	+	N/A	High
Dorling (2020)	+	-	+	+	N/A	Low

4.2.4.3. Recruitment strategies

Financial incentives

The ICER of a financial incentive, against no financial incentive, was estimated from two studies (six SWATs) (Jennings et al., 2015, Free et al., 2010). With an odds ratio of 1.65 (95% CI: 0.86, 3.18) and an incremental cost of $133.44, it costs $476.57 (95% CI: from $ 208.50 to N/A[3]) to recruit an additional patient (see *Table 4.S1* and *Figure 4.S1* in *Supplemental Material 4.3* for more details). All SWATs have a low Cochrane risk of bias, but moderate quality of economic evaluation. The I^2 statistic is 49%, signalling evidence of substantial between-study heterogeneity. There are three potential sources of between-group heterogeneity; 1) variations in healthcare settings across the host trials of the SWATs; 2) variations in the monetary incentives among SWATs and 3) variations in the populations across SWATs. The GRADE certainty of evidence is moderate for this recruitment strategy, due to inconsistency. In line

[3] N/A implies that the intervention is not effective at the lower bound of the 95% confidence interval. By the definition of ICER, since the correlation between the ICER and incremental recruitment rate is inverse, the higher bound (lower bound) effect size, i.e. incremental recruitment rate, is associated with the lower bound (higher bound) ICER. Therefore, if the effect size is negative, the corresponding ICER is undefined.

with Trial Forge Guidance 2 (Treweek et al., 2020), the GRADE criterion is met, the cumulative meta-analysis criterion is met, the PICOT criterion[4] is partially met, the balance of benefit and disadvantage to participants criterion is not met, and the balance of benefit and disadvantage to host trial criterion is not met. We suggest further studies including different monetary incentives be conducted in the future so that a figure of additional patients recruited by a $1 increase in monetary incentive be obtained.

Nudge interventions

The ICER of nudge interventions against usual recruitment procedures was estimated from three studies (three SWATs) (Free et al., 2010, Jolly et al., 2019, Rogers et al., 2019). Nudge interventions related to recruitment included: quotes from existing participants over text messages, a multimedia information resource that was developed through patient and public involvement (PPI) contributors and researchers, and a DVD presentation containing an audio-visual presentation explaining the host trial. With an odds ratio of 1.13 (95% CI: 0.72, 1.77) and an incremental cost of $22.00, it costs $314.29 (95% CI: from $68.75 to N/A) to recruit an additional patient (see *Table 4.S2* and *Figure 4.S2* in *Supplemental Material 4.3* for more details). The risk of bias is unclear for two studies (Jolly et al., 2019) (Rogers et al., 2019), and low for one study (Free et al., 2010) . In addition, the quality of economic evaluation is moderate for two studies (Free et al., 2010) (Rogers et al., 2019), and low for one study (Jolly et al., 2019). The I^2 statistic is 74%, signalling evidence of substantial between-study heterogeneity. There are four potential sources of such heterogeneity; 1) variations in healthcare settings across the host trials of the SWATs; 2) variations in the "nudge interventions"; 3) variations in the populations across SWATs and 4) variations in the designs of the included SWATs. The GRADE certainty of evidence is very low for this recruitment strategy, due to risk of bias, inconsistency, and indirectness of the included studies. In line with Trial Forge Guidance 2 (Treweek et al., 2020), the GRADE criterion is met, the cumulative meta-analysis criterion is met, the PICOT criterion is partially met, the balance of benefit and disadvantage to participants criterion is met, and the balance of benefit and disadvantage to host trial criterion is partially met. Therefore, further replications of SWATs associated with nudge interventions are encouraged.

[4] PICOT stands for: population, intervention, comparator, outcome, time frame

Screening of a trial by a senior investigator

The cost-effectiveness of this strategy was estimated according to a single SWAT (Miller et al., 1999). Screening for the host trial undertaken by a senior investigator, versus screening for the host trial undertaken by a research assistant, is not cost-effective, since the odds ratio is 0.19 (95% CI: 0.11, 0.32), with an incremental cost of $37.05 (see *Table 4.S3* in *Supplemental Material 4.3* for more details). Given the low sample size of the included study (Miller et al., 1999), and its high Cochrane risk of bias, the GRADE certainty of evidence is low due to imprecision and risk of bias. The included study has a moderate quality of economic evaluation.

Text messages versus telephone calls

The cost-effectiveness of this strategy was estimated according to a single SWAT (Free et al., 2010). With an odds ratio of 3.47 (95% CI: 1.27, 9.48) and an incremental cost of $22.00, it costs only $4.41 (95% CI: from $2.45 to $23.38) to recruit an additional patient (see *Table 4.S3* in *Supplemental Material 4.3* for more details). Given the sample size of the included study (Free et al., 2010), and its low Cochrane risk of bias, the GRADE certainty of evidence is moderate due to imprecision. The included study has a moderate quality of economic evaluation.

Pre-notification leaflet

The cost-effectiveness of this strategy was estimated according to a single SWAT (Arundel et al., 2017). With an odds ratio of 1.17 (95% CI: 0.87, 1.57) and an incremental cost of $2.25, it costs $25.97 (95% CI: from $9.00 to N/A) to recruit an additional patient (see *Table 4.S3* in *Supplemental Material 4.3* for more details). Given the sample size of the included study (Arundel et al., 2017), and its low Cochrane risk of bias, the GRADE certainty of evidence is moderate due to imprecision. The included study has a high quality of economic evaluation.

Telephone reminders versus text reminders

The cost-effectiveness of this strategy was estimated according to a single SWAT (Bracken et al., 2019). With an odds ratio of 1.37 (95% CI: 0.95, 1.98) and an incremental cost of $3.98, it costs $23.37 (95% CI: from $10.47 to N/A) to recruit an additional patient (see *Table 4.S3* in *Supplemental Material 4.3* for more details). Given the sample size of the included study (Bracken et al., 2019), and its low Cochrane risk of bias, the GRADE certainty of evidence is moderate due to imprecision. The included study has a high quality of economic evaluation.

Invitation packs by GP

The cost-effectiveness of this strategy was estimated according to a single SWAT (Hancocks et al., 2019). With an odds ratio of 7.75 (95% CI: 1.04, 57.97) and an incremental cost of $1.13, it costs $1.00 (95% CI: from $0.50 to $57.47) to recruit an additional patient (see *Table 4.S3* in *Supplemental Material 4.3* for more details). However, since these figures were obtained from an abstract (Hancocks et al., 2019), the GRADE certainty of evidence is very low due to risk of bias, imprecision, indirectness and publication bias. The included study has an unclear quality of economic evaluation.

Trial-branded pen

The cost-effectiveness of this strategy was estimated according to a single SWAT (Whiteside et al., 2019). With an odds ratio of 1.04 (95% CI: 0.65, 1.66) and an incremental cost of $0.47, it costs $21.41 (95% CI: from $1.68 to N/A) to recruit an additional patient (see *Table 4.S3* in *Supplemental Material 4.3* for more details). Given the sample size of the included study (Whiteside et al., 2019), and its low Cochrane risk of bias, the GRADE certainty of evidence is moderate due to imprecision. The included study has a high quality of economic evaluation.

In line with Trial Forge Guidance 2 (Treweek et al., 2020), we encourage all the recruitment strategies to be replicated in future SWATs, but the comparison of pre-notification leaflet against no leaflet.

Ranking recruitment strategies

The cost-effectiveness rank of the eight recruitment strategies is available in *Table 4.4* and has been determined according to the following descending order: GRADE evidence, statistical certainty, ICER. Providing financial incentives might be an effective recruitment strategy, but its ICER is relatively high, at $476.57, thus questioning its cost-effectiveness; more SWATs of financial incentives with moderate amounts (i.e. significantly less than £100) are needed to estimate the ICER. The following may be considered cost-effective strategies: providing a telephone reminder versus a SMS reminder, or a branded pen versus no pen, or a pre-notification leaflet versus no leaflet. However, whereas their corresponding ICERs are relatively low, their OR lower bounds signal they may not actually be effective recruitment strategies. Providing primary text message, versus primary call and no text message, might be a very cost-effective strategy; however, more SWATs of this strategy are needed since its

GRADE certainty of evidence is low. It remains uncertain whether the remaining recruitment strategies are cost- effective since their GRADE certainty of evidence is very low.

Overall, there is no complete certainty up to date on which recruitment strategies would be cost-effective for trial teams to use for recruiting eligible patients to their trials. Nevertheless, strategies such as financial incentives, trial-branded pens, telephone reminders and pre-notification leaflets could possibly provide recruitment benefits to future trials in a cost-effective manner. More evidence is needed to determine the cost-effectiveness of such strategies.

Table 4.4: Cost-effectiveness rank of different recruitment strategies

Cost-effectiveness rank of different recruitment strategies					
Rank	**Strategy**	**Number of SWATs**	**Sample size**	**GRADE certainty of evidence**	**ICER**
1	Financial incentive vs no financial incentive	6	1506	Moderate	$476.57 ($ 208.50, N/A)
2	Branded pen with trial invitation pack vs no pen	1	1862	Moderate	$21.41 ($1.68,N/A)
3	Telephone reminder vs SMS reminder	1	709	Moderate	$23.37 ($10.47,N/A)
4	Pre-notification leaflet vs no leaflet	1	4314	Moderate	$25.97 ($9.00,N/A)
5	Primary text message vs primary call and no text message	1	937	Moderate	$4.41 ($2.45, $23.38)
6	Invitation pack from a surgeon vs text message	1	1267	Very Low	$1.00 ($0.50,$57.47)
7	Nudge intervention vs usual recruitment	3	6054	Very Low	$314.29 ($68.75 , N/A)
8	Screening for the host trial undertaken by a senior investigator vs screening undertaken by a research assistant	1	347	Low	N/A (ineffective)

4.2.4.4. Retention strategies

Trial-branded pen

The ICER of providing a trial-branded pen versus no pen was estimated from three studies (three SWATs) (Cunningham-Burley et al., 2020, Whiteside et al., 2019, Bell et al., 2016). With an odds ratio of 1.14 (95% CI: 1.00, 1.30) and an incremental cost of $0.52, it costs $6.98 (95% CI: from $3.63 to N/A) for an additional participant to be retained in a host trial (see *Table 4.S4* and *Figure 4.S3* in *Supplemental Material 4.4* for more details). One included study has a low Cochrane risk of bias, (Whiteside et al., 2019) whereas the remaining studies have an unclear risk of bias (Cunningham-Burley et al., 2020, Bell et al., 2016). All studies have a

high quality of economic evaluation. The I^2 statistic is negligible at 0%, signalling no evidence of substantial between-study heterogeneity. However, there are four potential sources of such heterogeneity; 1) variations in healthcare settings across the included SWATs; 2) variations in retention periods among SWATs; 3) variations in the populations across the SWATs and 4) variations in the SWATs' designs. The GRADE certainty of evidence for this retention strategy is moderate, due to inconsistency. In line with Trial Forge Guidance 2 (Treweek et al., 2020), the GRADE criterion is met, the cumulated evidence criterion is not met, the PICOT criterion is partially met, the balance of benefit and disadvantage to participants criterion is met and the balance of benefit and disadvantage to host trial criterion is not met. Therefore, we argue further SWATs associated with trial-branded pens as a retention strategy to be undertaken.

Financial incentives

The ICER of providing financial incentives versus no incentives was estimated from three studies (three SWATs) (Khadjesari et al., 2011, Gates et al., 2009, Kenyon et al., 2005). With an odds ratio of 1.33 (95% CI: 1.15, 1.53) and an incremental cost of $8.20, it costs $15.89 (95% CI: from $10.65 to $32.42) for an additional participant to be retained in a host trial (see *Table 4.S5* and *Figure 4.S4* in *Supplemental Material 4.4* for more details). Once the quasi-randomised SWAT included in the three SWATs of financial incentives is removed, the result is not significantly affected: the OR slightly falls to 1.32 (95%: 1.01, 1.73), the incremental cost increases to $11.08 and the ICER slightly increases to $22.05 (95% CI: from $11.17 to $615.21). The Cochrane risk of bias is low in one study (Kenyon et al., 2005), unclear in one study (Khadjesari et al., 2011), and high in one study (Gates et al., 2009). Furthermore, two studies have a low quality of economic evaluation (Khadjesari et al., 2011) (Kenyon et al., 2005), whereas one study has a high quality of economic evaluation (Gates et al., 2009). The I^2 statistic is 37%, signalling low evidence of substantial between-study heterogeneity. There are two sources of such heterogeneity: 1) variations in retention periods among SWATs and 2) variations in the monetary incentives among SWATs. The GRADE certainty of evidence for this retention strategy is moderate due to risk of bias. In line with Trial Forge Guidance 2 (Treweek et al., 2020), the GRADE criterion is met, the cumulative meta-analysis criterion is not met, the PICOT criterion is partially met, the balance of benefit and disadvantage to participants criterion is partially met, and the balance of benefit and disadvantage to host trial criterion is not met. We encourage replications of further SWATs associated with financial incentives as a strategy for improving participant retention in RCTs. Three further studies (or three SWATs) were not included in this meta-analysis (Khadjesari et al., 2011, Marsh and

Kendrick, 1999, James et al., 2021). The results of these studies, and the reasons for which they were not included are available in *Supplemental Material 4.7*.

Nudge interventions

The ICER of a nudge intervention versus usual retention was estimated from three studies (three SWATs) (Jolly et al., 2019, Hardy et al., 2016, Cochrane et al., 2020). Nudge interventions related to retention included: a personalised text message instead of a generalised one, a multimedia information resource that was developed through patient and public involvement (PPI) contributors and researchers, and a social incentive cover letter instead of a standard one. With an odds ratio of 1.14 (95% CI: 0.94, 1.39) and an incremental cost of $0.84, it costs $11.55 (95% CI: from $4.61 to N/A) for an additional participant to be retained in a host trial (see *Table 4.S6* and *Figure 4.S5* in *Supplemental Material 4.4* for more details).The Cochrane risk of bias is low in two studies (Hardy et al., 2016, Cochrane et al., 2020), and unclear in one study (Jolly et al., 2019). The quality of economic evaluation is high in one study (Hardy et al., 2016), moderate in one study (Cochrane et al., 2020), and low in one study (Jolly et al., 2019). The I^2 statistic is negligible at 0%, signalling low evidence of substantial between-study heterogeneity. However, there are still three sources of such heterogeneity; 1) variations in "nudge" interventions among SWATs; 2) variations in retention periods among SWATs and 3) variations in the SWATs' designs. The GRADE certainty of evidence for this retention strategy is moderate due to inconsistency. In line with Trial Forge Guidance 2 (Treweek et al., 2020), the GRADE criterion is met, the cumulative meta-analysis criterion is met, the PICOT criterion is partially met, the balance of benefit and disadvantage to participants criterion is not met, and the balance of benefit and disadvantage to host trial criterion is partially met. We encourage replications of further SWATs associated with nudge interventions for improving participant retention in RCTs.

Unconditional monetary incentive versus conditional monetary incentive

The ICER of an unconditional monetary incentive, versus a conditional one, was estimated from two studies (two SWATs) (Cook et al., 2021, Dorling et al., 2020). With an odds ratio of 0.90 (95% CI: 0.31, 2.64) and an incremental cost of $18.61, such a strategy is not cost-effective, since its estimated odds ratio is less than 1 (see *Table 4.S7* and *Figure 4.S6* in *Supplemental Material 4.4* for more details). The Cochrane risk of bias is unclear for both studies, whereas the quality of economic evaluation is high in one study (Cook et al., 2021)and

moderate in the other study (Dorling et al., 2020). The I^2 statistic is 93%, demonstrating high evidence of substantial between-study heterogeneity. There are five potential sources of such heterogeneity; 1) variations in healthcare settings between the host trials of the included SWATs; 2) variations in the populations between SWATs; 3) differences in the interventions between SWATs; 4) variations in retention periods between SWATs and 5) variations in the SWATs' designs. The GRADE certainty of evidence is low, due to risk of bias and inconsistency. In line with Trial Forge Guidance 2 (Treweek et al., 2020), the GRADE criterion is met, the cumulative meta-analysis criterion is met, the PICOT criterion is partially met, the balance of benefit and disadvantage to participants criterion is met, and the balance of benefit and disadvantage to host trial criterion is met. We highly encourage replications of further SWATs comparing unconditional with conditional monetary incentives for improving participant retention in RCTs.

Pre-notification card

The cost-effectiveness of this strategy was estimated according to a single SWAT (Treweek et al., 2021). With an odds ratio of 1.26 (95% CI: 0.99, 2.19) and an incremental cost of $1.02, it costs $4.86 (95% CI: from $2.76 to $N/A) to retain an additional participant in a host trial (see *Table 4.S8* in *Supplemental Material 4.4* for more details). Given the low sample size of the included study and its low Cochrane risk of bias, (Treweek et al., 2021), the strategy's GRADE certainty of evidence is moderate, due to imprecision. The included study has a moderate quality of economic evaluation.

Electronic prompts

The cost-effectiveness of this strategy was estimated according to a single SWAT (Clark et al., 2015). With an odds ratio of 1.48 (95% CI: 0.81, 1.96) and an incremental cost of $0.12, it costs $0.55 (95% CI: from $0.28 to $N/A) to retain an additional participant in a host trial (see *Table 4.S8* in *Supplemental Material 4.4* for more details). Given the low sample size of the included study and its unclear Cochrane risk of bias (Clark et al., 2015), the strategy's GRADE certainty of evidence is low, due to risk of bias and imprecision. The included study has a low quality of economic evaluation.

Trial-branded pen (before recruitment)

The cost-effectiveness of this strategy was estimated according to a single SWAT (Whiteside et al., 2019). With an odds ratio of 8.27 (95% CI: 1.04, 66.00) and an incremental cost of $0.47,

it costs $0.40 (95% CI: from $0.20 to $23.50) to retain an additional participant in a host trial (see *Table 4.S8* in *Supplemental Material 4.4* for more details). Given the sample size of the included study and its low Cochrane risk of bias (Whiteside et al., 2019), the strategy's GRADE certainty of evidence is moderate, due to imprecision. The included study has a high quality of economic evaluation.

In line with Trial Forge Guidance 2 (Treweek et al., 2020), we highly encourage the aforementioned retention strategies to be replicated in future SWATs.

Ranking retention strategies

A summary of the cost-effectiveness rank of the seven retention strategies is provided on *Table 4.5* and has been determined according to the following descending order: GRADE evidence, statistical certainty, ICER. Providing pens before patient recruitment to an RCT is potentially a very cost-effective strategy, with an ICER of $0.40. However, as this finding is derived from a single study with a low sample size, more evidence is needed to confirm the figure. A retention strategy, which also seems to be cost-effective with moderate GRADE certainty of evidence, is the provision of £5 up to £10 vouchers; the ICER is relatively low at $15.89. Providing a trial-branded pen is potentially another cost-effective retention strategy, with its ICER being very low, at $6.98. However, its lower bound OR=1, meaning there is still a chance such a strategy is not (cost-) effective. Due to either low GRADE certainty of evidence or wide confidence intervals of the remaining retention strategies, it is inconclusive whether these strategies are cost-effective or not.

We encourage trial researchers to consider financial incentives of up to £10 and/or trial-branded pens as retention strategies, while we recommend more SWATs of these strategies be undertaken. Despite the reported lower bound OR, we still encourage pens as a retention strategy due to its low reported ICER and low incremental costs, especially for trials involving postal questionnaires.

Table 4.5: Cost-effectiveness rank of different retention strategies

Cost-effectiveness rank of different retention interventions					
Rank	**Strategy**	**Number of SWATs**	**Sample size**	**GRADE certainty of evidence**	**ICER**
1	Trial-branded pen versus no trial-branded pen (before recruitment)	1	92	Moderate	$0.40 ($0.20, $23.50)
2	Financial incentive versus no financial incentive	3	5753	Moderate	$15.89 ($10.65,$32.42)
3	Trial-branded pen versus no trial-branded pen	4	9790	Moderate	$6.98 ($3.63, N/A)
4	Nudge intervention versus usual recruitment procedure	3	5276	Moderate	$11.55 ($4.61,N/A)
5	Pre-notification card versus no pre-notification card	1	558	Moderate	$4.86 ($2.76,N/A)
6	Electronic prompts versus no electronic prompts	1	437	Low	$0.55 ($0.28, N/A)
7	Unconditional monetary incentive versus conditional monetary incentive	2	1268	Low	$465.25 ($97.95,N/A)

4.2.5. Discussion

4.2.5.1. Summary of findings

Whereas Cochrane reviews have explored the *effectiveness* of strategies for improving recruitment and retention in RCTs (Treweek et al., 2018b, Gillies et al., 2021), this review additionally appraises the *cost-effectiveness* of recruitment and retention strategies. The findings demonstrate an uncertainty regarding which strategies are cost-effective for improving participant recruitment and/or retention in RCTs. For both recruitment and retention strategies, the uncertainty of the evidence primarily originates from the evaluation of several potential strategies from single studies, but without any replications. The corresponding Cochrane reviews on recruitment and retention in RCTs suggested that replications of SWATs with strategies having a moderate GRADE certainty of evidence be undertaken (Treweek et al., 2018b, Gillies et al., 2021), a recommendation we also make for bolstering the evidence on the cost-effectiveness of strategies for improving recruitment or retention in SWATs.

Overall, there is no retention strategy which we would recommend trial teams and researchers adopt with complete statistical certainty. Providing vouchers of up to £10 during follow-up

could be a cost-effective retention strategy with an estimated ICER of $15.89; it costs only $15.89 for an additional participant to be retained in a host trial. Providing a trial-branded pen may also be a cost-effective strategy, with an ICER of $6.98, yet not statistically significant since its lower bound OR=1 (hence its lower bound effectiveness is zero). Also, providing a trial-branded pen before recruitment, may be a cost-effective strategy, with an ICER of $0.40 which is also statistically significant. However, the GRADE certainty of evidence for both strategies is moderate, meaning that additional SWATs of these strategies would be beneficial for making more certain inferences about their cost-effectiveness. ICERs were derived for further retention strategies; however, it remains inconclusive whether these are cost-effective due to their low or very low GRADE certainty of evidence. Whereas the retention review found the inclusion of self-kits or a diary to be effective strategies (Gillies et al., 2021), no data for evaluating their cost-effectiveness were available. Therefore, we highly encourage the undertaking of future SWATs of these strategies and the inclusion of economic evaluations alongside such SWATs. Similarly to the retention review (Gillies et al., 2021), we also encourage the undertaking of further SWATs associated with the cost-effectiveness of patient and public involvement (PPI) interventions, since PPI is a key unanswered question about trial retention (Brunsdon et al., 2019). Overall, due to the low ICER and incremental costs we recommend trial teams use trial-branded pens as a retention strategy, especially in trials involving postal questionnaires for follow-up. Providing vouchers of up to £10 could be another beneficial retention strategy for trial-teams.

Also, there is no recruitment strategy which we would recommend trial teams and researchers adopt with complete statistical certainty. Including a branded pen with a trial invitation pack, or a telephone reminder versus an SMS reminder, could be cost-effective strategies, with their ICERs being low, at $21.41 and $23.37 respectively. However, as their lower bound ORs are less than 1, their cost-effectiveness is not statistically significant. In addition, their GRADE certainty of evidence is moderate, implying these strategies would benefit from further SWATs to determine their cost-effectiveness with less uncertainty. Providing financial incentives may be an effective yet a costly strategy, with $476.57 required to recruit an additional patient. However, there is substantial heterogeneity among the associated SWATs, since very different monetary incentives were present (i.e. from £5 up to £100). Therefore, we encourage the cost-effectiveness of moderate financial incentives (i.e. less than £100 per participant recruited (Treweek et al., 2020)) to be evaluated in future SWATs. Unfortunately, we could not estimate the cost-effectiveness of using an open design, compared to a placebo-controlled design, as the

associated SWATs did not undertake any relevant economic evaluations or provide costs related to such strategies. Since this strategy appears to be effective at improving recruitment (Treweek et al., 2018b), economic evaluations of such a strategy alongside future SWATs are welcome. We also encourage the estimation of the cost-effectiveness of incorporating user-testing for improving the participant information leaflet (PIL) in future SWATs, as a recruitment strategy.

4.2.5.2. Recommendations for future economic evaluations alongside SWATs

To minimise the uncertainty regarding the findings from SWATs on the cost-effectiveness of recruitment and/or retention strategies, we highly recommend the application of Value of Information (VoI) analyses. Such an analysis can inform decision makers on whether more trials are needed to minimise the uncertainty of the cost-effectiveness of a strategy. A VoI analysis could be used in line with the Trial Forge Guidance 2 (Treweek et al., 2020) to confirm whether a further SWAT associated with a recruitment/retention strategy should be undertaken. For instance, since we concluded trial-branded pens to be a potentially cost-effective retention strategy with moderate GRADE certainty of evidence, and the Cochrane review concluded pens to be a potentially effective strategy with low GRADE certainty of evidence, the GRADE criterion is met in Trial Forge Guidance 2 and hence further SWATs on pens are recommended. However, it seems that such a strategy could potentially be a very cost-effective one for participant retention, and hence it may not be necessary to undertake another SWATs, which would require the financing of resource costs. To determine whether more SWATs are needed for determining its effectiveness, a VoI analysis for trial-branded pens could be undertaken. A framework of VoI analysis related to SWATs which trial researchers could follow is available in the literature, and applicable after a standard meta-analysis of a recruitment or retention strategy (Claxton et al., 2015a).

A concern was that although 139 studies were originally included in the recruitment and retention reviews (Treweek et al., 2018b, Gillies et al., 2021), only 17 of these studies were included in our review. Therefore, economic evaluations were not undertaken alongside the majority of SWATs. Whereas capturing the effectiveness of different recruitment or retention strategies is useful, cost considerations are equally important due to limited availability of financial resources. We highly encourage trialists and researchers to undertake economic

evaluations alongside *all* SWATs in the future. In addition, the costs of obtaining outcome data were not provided in six studies (10 SWATs), meaning that the cost-effectiveness of some strategies may have been overestimated. Therefore, the reporting of costs should be transparent, expressed in unit terms, and stratified into different types of direct and indirect costs, including the costs of obtaining outcome data. Finally, in all cost-effectiveness analyses there is a defined cost-effectiveness threshold to determine whether a given intervention, which is both more effective compared to existing interventions and costlier, is cost-effective. As long as the ICER is less (more) than the threshold, then a strategy is (not) cost-effective. We recommend trial researchers to define such a threshold for determining which recruitment or retention strategies should be considered as cost-effective. In our review, we did not set out a cost-effectiveness threshold, as there has not been any research in this area; instead we presented cost-effectiveness ranks of recruitment and retention strategies for comparisons with respect to their cost-effectiveness to be made.

The perspective all SWATs followed was related to the trial teams. However, poor recruitment into RCTs may also lead to indirect costs through the generation of foregone health benefits to an affected population not experiencing the clinical benefits of a potentially effective intervention. For instance, a study modelled the impact on human lives lost due to poor recruitment in the COVID-19 RECOVERY trial, which showed that over 2,800 lives could have been saved in the UK (Knowlson and Torgerson, 2020). Similarly, the financial costs of poor attrition can be significant, with the time costs of researchers dealing with follow- up being dominant (Peterson et al., 2012). When the follow up to a funded RCT is poor, this may generate huge costs for RCT funders, as they could have instead provided funding to trials with better follow-up rates and hence with more statistical accuracy in their results. Therefore, in future SWATs it is recommended that researchers adopt a broader perspective where possible when conducting economic evaluations alongside SWATs, such as the perspective of a national healthcare system or the societal perspective (i.e. through cost-benefit analysis instead of cost-effectiveness analysis).

4.2.5.3. Strengths and limitations of the review

The major limitations in our study were the differential definitions and computations of cost-effectiveness outcomes among the included studies. These were partially captured through manual conversions of ICERs, or any other secondary economic outcome, into unit incremental costs, by stringently following the definition of ICER (*Equation 4.1*, *Equation 4.2*) and the

reported recruitment or retention rates. This approach enabled us to obtain cost-effectiveness figures from 20 out of 22 studies, or from 26 out of 29 included SWATs in a homogeneous manner. Another limitation could be our flexible approach towards including studies with high Cochrane risk of bias or low quality of economic evaluation, or studies that have not been peer reviewed or published yet. However, as an appraisal of the cost-effectiveness of strategies for improving participant recruitment and retention in RCTs was not explored in the past, we encouraged this flexible approach during the screening of records and inclusion of studies. Finally, there were differences in the definitions of "recruitment rate" or "retention rate", especially in terms of the recruitment and retention periods, across the included studies. However, we encouraged flexibility in the definitions of such terms by the same means.

Overall, the review benefits from such flexibility so that the evidence on the cost-effectiveness or recruitment and/or retention strategies is fully captured. In addition, all studies were subject to extensive quality appraisals, including the Cochrane risk of bias and quality of economic evaluation. Moreover, the certainty of the evidence for each recruitment and/or retention strategy was extensively assessed through the GRADE approach and Trial Forge Guidance 2. We believe the use of multiple tools strengthens the reliability of our findings. Finally, we believe our review motivates the research community to undertake economic evaluations alongside all future SWATs; we have also made recommendations on how such economic evaluations could be optimally undertaken.

4.2.6. Conclusion

There is no recruitment or retention strategy which we would recommend trial teams and researchers adopt with full certainty. Improving recruitment and retention in RCTs is a priority for trial teams, reflected through the emergence of SWATs as a study design to improve trial efficiency. It is of paramount importance for future SWATs to replicate existing recruitment and/or retention strategies, rather than focus on novel strategies. We also recommend that economic evaluations be carried out alongside *all* future SWATs, costs and benefits be reported clearly and transparently, the cost-effectiveness of existing recruitment or retention strategies be repeatedly evaluated, and broader perspectives be adopted in future SWATs if applicable. Finally, we encourage researchers to undertake VoI analyses for each recruitment and retention strategy, in combination with Trial Forge Guidance 2, to minimise the uncertainty of the evidence.

Chapter 5: A Value of Information Analysis framework for SWATs of recruitment and retention strategies

5.1. Abstract

Background: Trial Forge Guidance 2 is a guidance for researchers to determine whether a further SWAT, related to a specific recruitment or retention strategy, is recommended so as to reduce the underlying uncertainty (Treweek et al., 2020). Due to its stringent qualitative criteria, however, this guidance is subject to an increased likelihood of mistakenly encouraging an investment into further SWATs of a given strategy given budget constraints.

Methods: This chapter introduces an already established framework of Value of Information (VoI) analysis to address the optimisation of SWAT-related research given the financial constraints (Claxton et al., 2015a). Two case studies of telephone reminders as a recruitment strategy and pens as a retention strategy, were used, for which Trial Forge Guidance 2 recommends that further SWATs be undertaken according to the uncertainty of the evidence. The modified VoI analysis framework was gradually adopted across the two and five SWATs related to telephone reminders and pens respectively, enabling the estimation of figures related to the value of implementation and the value of additional research.

Results: In the case of telephone reminders, the value of additional research is 0 and the value of implementation is 67,765. In other words, 740 additional SWATs need to be funded for other recruitment strategies to observe similar expected incremental recruitment rates, to those obtained under the existing evidence for telephone reminders, and no further SWAT would be required, so as to reduce any uncertainty related to telephone reminders. The VoI analysis suggests that no further SWATs of telephone reminders be undertaken, and that telephone reminders be instantly adopted as a recruitment strategy in all trials. In the case of pens the value of additional research is 0 and the value of implementation is 6,227. In other words, 72 additional SWATs need to be funded for other retention strategies to observe similar expected incremental retention rates, to those obtained under the existing evidence for pens, and no further SWAT would be sufficient, so as to reduce any uncertainty related to pens. The VoI analysis suggests that no further SWATs of pens be undertaken, and that pens be instantly adopted as a retention strategy in all trials involving returns of questionnaires during follow-up.

Conclusion: The VoI analysis framework for making investment decisions on SWATs is applicable for all recruitment and retention strategies. It is strongly encouraged that VoI analysis be combined with Trial Forge Guidance 2 to prioritise SWAT-related research into finding a balance between qualitative quality and financial constraints when making investment decisions on SWATs of recruitment and retention strategies and deciding which of these strategies should be adopted in recruitment and retention processes for future trials.

5.2. Introduction

5.2.1. Challenges of poor recruitment and retention in RCTs

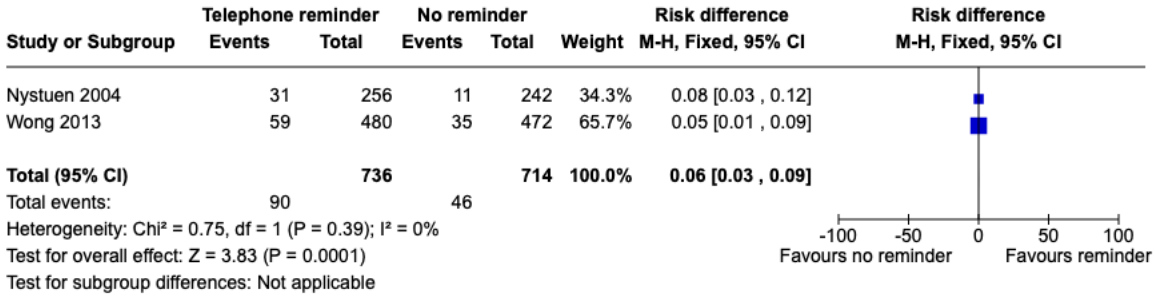
Poor recruitment of patients to RCTs leads to underpowered trials that are subject to a high risk of Type II errors, causing: research waste, the rejection of effective healthcare interventions, delays in meta-analyses ascertaining the effectiveness of rejected yet effective interventions, ethical issues emerging from exposing participants to uncertainty during and after the trial and, most importantly, extension of the length of a given trial, which puts a huge strain on its existing budget (Treweek et al., 2018b). Poor retention of already recruited participants in RCTs may threaten the internal and the external validity of such research designs, as it not only diminishes the power of a trial but can also introduce selection bias, i.e. randomisation of trial arms is not achieved eventually and hence a trial's results become more uncertain and/or upwardly or downwardly biased (Torgerson and Torgerson, 2008). As a result, the research community wants to address these threats to trial efficiency, which may also have a remarkable impact on healthcare systems, trial teams' finances and funders, as *Chapter 2* and *Chapter 3* have demonstrated. Studies Within A Trial (SWATs) are highly encouraged for identifying effective (and cost-effective) recruitment and retention strategies (Treweek et al., 2018a). Two Cochrane reviews have critically appraised the evidence surrounding the effectiveness of recruitment and strategies (Gillies et al., 2021, Treweek et al., 2018b), whereas the review of *Chapter 4* has critically appraised the evidence surrounding their cost-effectiveness (Gkekas et al., 2023).

5.2.2. Evidence on telephone reminders as a recruitment strategy

One of the recruitment strategies identified by the Cochrane reviewers was the use of telephone reminders to invited patients not responding to postal invitations, to participate in randomised trials (Treweek et al., 2018b). Using the proportion of participants recruited to each of the included studies as the primary outcome, and the risk difference (RD) as the unit of statistical analysis, telephone reminders generated a noticeable, statistically significant improvement in recruitment rates by RD=6% (95% confidence interval (CI): 3% to 9%). The corresponding Cochran-Mantel-Haenszel meta-analysis was undertaken using two SWATs, with a sample size of 1450 participants (Nystuen and Hagen, 2004, Wong et al., 2013). There is low evidence of statistical uncertainty, with the cumulative and study-specific lower bounds of the 95% CIs exceeding zero. The I^2 statistic, which is a test of heterogeneity between the included studies in a meta-analysis, is equal to 0%, thus demonstrating insufficient evidence to reject the

hypothesis that the included studies are homogeneous. The results from the meta-analysis are shown on *Figure 5.1* below, produced with Reviewer Manager (RevMan).

Figure 5.1: Meta-analysis of telephone reminders for non-responders versus no telephone reminders, for recruitment to randomised trials



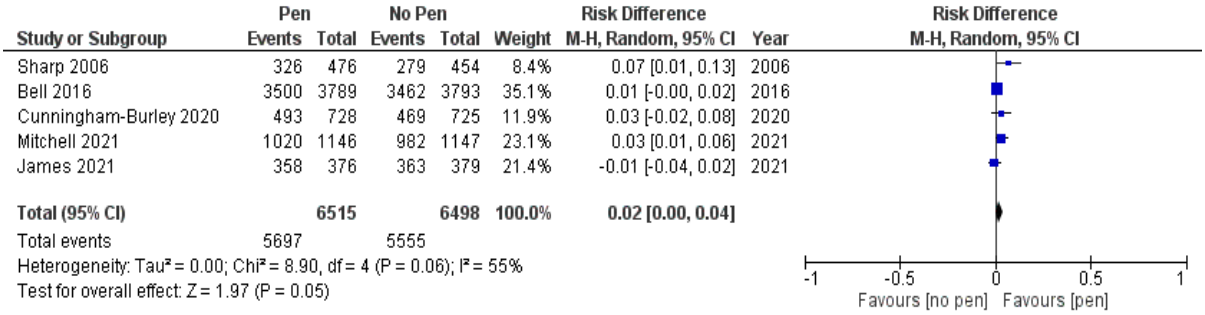
According to *Figure 5.1,* the reviewers concluded that the GRADE certainty of evidence for this recruitment strategy is high, meaning they have high confidence that the RD estimate of using telephone reminders for patients who are invited to a randomised trial but do not respond to postal invitations, is similar to the true effect. Both included SWATs were assessed to have a low Cochrane risk of bias (Nystuen and Hagen, 2004, Wong et al., 2013). In addition, there was no evidence for substantial unexplained heterogeneity or imprecision. The reviewers thus assessed this strategy as having a high GRADE certainty of evidence. Nevertheless, as recruitment rates were below 10% in both of the host trials of the included studies, the indirectness criterion is not met for trials with recruitment rates exceeding 10%, and therefore the reviewers downgraded the GRADE certainty of evidence to moderate in this case (Treweek et al., 2018b).

5.2.3. Evidence on pens as a retention strategy

One of the retention strategies that was identified by the reviewers was the additional provision of pens to participants being followed up for data collection from host RCTs (Gillies et al., 2021). Using the proportion of participants retained in each of the included studies as the primary outcome, and the risk difference (RD) as the unit of statistical analysis, adding a pen to questionnaire letters versus no pen generated a slight, borderline statistically significant improvement in retention rates by RD=2% (95% CI: 0% to 4%). The corresponding inverse variance weighted meta-analysis was undertaken using five SWATs, with a sample size of

13013 participants (Sharp et al., 2006) (Bell et al., 2016) (Cunningham-Burley et al., 2020) (James et al., 2021) (Mitchell et al., 2021). The results from the meta-analysis are shown on *Figure 5.2* below, produced with Reviewer Manager (RevMan).

Figure 5.2: Meta-analysis of adding pens to questionnaires for follow-up versus no pen for follow up



According to *Figure 5.2*, there is uncertainty in the aforementioned RD estimate since three SWATs have a CI lower bound RD that is slightly below or equal to zero, thus indicating poor evidence of a statistically significant improvement in retention rates because of the addition of pen for follow-up (versus no pen), whereas two SWATs have a CI lower bound RD that is above zero, thus demonstrating a statistically significant improvement in retention rates because of the addition of pen for follow-up. The I^2 statistic is equal to 55%, a figure exceeding the cut-off value of 50%, thus demonstrating evidence of moderate heterogeneity among the included SWATs. The main factor for potential heterogeneity could be clinical diversity, as each of the five SWATs' corresponding host trials was undertaken in different clinical settings, e.g. one host RCT was related to knee replacement (Mitchell et al., 2021) whereas another host RCT was related to falls prevention (James et al., 2021) , which unavoidably implies variability in trial participants. Nevertheless, the reported p-value (=0.06) of the chi-squared test for heterogeneity, which is different from the I^2 statistic in the sense that it depends upon the number of studies included for meta-analysis, is below the 10% confidence level, thus implying insufficient evidence for rejecting the null hypothesis of homogeneity among the included SWATs. Overall, there seems to be no substantial heterogeneity in the meta-analysis, and the reasons for moderate heterogeneity can be largely explained by the different clinical settings of the host RCTs.

According to these findings, the reviewers concluded that the GRADE certainty of evidence is low, meaning that their confidence in the RD estimate of adding pens for follow-up, versus no pen, is limited and hence "*the true effect may be substantially different from the estimate of the effect*" (Gillies et al., 2021). Two of the included SWATs were assessed to have a low Cochrane risk of bias (James et al., 2021, Mitchell et al., 2021), whereas the remaining SWATs were assessed to have an uncertain Cochrane risk of bias (Sharp et al., 2006, Bell et al., 2016, Cunningham-Burley et al., 2020). In addition, there was no evidence for substantial unexplained heterogeneity. Therefore, the main reason for the reviewers having downgraded the rating of evidence seems to be imprecision, as it is related to the reported 95% CIs of RD associated with the addition of pens for follow-up.

5.2.4. Trial Forge Guidance 2

Apart from the GRADE rating of evidence, Trial Forge, a collaborative group aiming to improve and disseminate rigorous evidence for improving the conduct of RCTs, has produced guidance for researchers to determine whether a further SWAT, related to a specific retention (or recruitment) strategy, is recommended (Treweek et al., 2020). This guidance, called Trial Forge Guidance 2, is based on the following five criteria: GRADE, cumulated evidence, PICOT, balance of benefit and disadvantage to participants, and balance of benefit and disadvantage to the host trial (Treweek et al., 2020). By applying the Trial Forge Guidance 2 criteria in the cases of telephone reminders as a recruitment strategy and of using pens as a retention strategy, it is recommended that further SWATs of telephone reminders and pens be undertaken (see *Table 5.1* from Treweek et al. (2020) and *Table 5.2* using my own judgment respectively). Nevertheless, whilst this guidance has produced criteria that ensure the soundness and the qualitative quality of SWAT-related research, the authors themselves recognise that "*it is currently unlikely that applying the five criteria to any body of evidence will lead to a decision not to start another evaluation*" (Treweek et al., 2020) and that *"the most efficient way of approaching the limited time and money available for evidence generation about trial processes may be to focus on whether something clears a threshold that makes it worth doing, rather than having a precise estimate of its effect. There would be little to gain from pursuing perfection if it will not change decisions*" (Treweek et al., 2020). The present chapter attempts to introduce such a threshold, by using a well-established health economic framework for valuing the anticipated benefits and costs of reducing the uncertainty about the effectiveness of recruitment and retention strategies through additional SWATs; this framework is called the value of information (VoI) analysis.

Table 5.1: Trial Forge Guidance 2: Using telephone reminders as a recruitment strategy, adapted from Treweek et al. (2020)

Trial Forge Guidance 2 (Using telephone reminders as a recruitment strategy)		
Criterion	**Comments**	**Criterion Met (Yes or No or Partially)**
GRADE rating of evidence	Data are available for recruitment only (two trials, n = 1450). The GRADE certainty in the evidence for the two trials in the review is high but is considered moderate for trials that do not have low (< 10%) underlying recruitment.	Partially
Cumulated evidence	Data are available for recruitment only. There are only two trials and it seems too early to claim the cumulative meta- analysis has converged.	Yes
PICOT	P: One study was done in Norway in 2002–2003 and involved people aged 16–66 years who were sick-listed for > 7 weeks due to non-severe psychological problems or musculoskeletal pain (Nystuen and Hagen, 2004). The second study was carried out in Canada in 2010 and involved people aged 50–70 years from family practice lists who were eligible for colorectal cancer screening (Wong et al., 2013). I: The host trial intervention in the Norwegian study was solution-focused sessions led by psychologists; these were one-on-one or in groups and aimed to help people get back to work (Nystuen and Hagen, 2004). The host trial interventions in the Canadian study were of virtual colonoscopy, optical colonoscopy or faecal occult blood testing (Wong et al., 2013) C: The host trial comparator in the Norwegian study was usual care: written information from the social security office (Nystuen and Hagen, 2004). The Canadian host trial effected a head-to-head evaluation of three screening methods, so the three interventions mentioned above were also the comparators (Wong et al., 2013) O: Both studies measured recruitment to the host trial. Both host trials had low underlying recruitment. T: Mobile telephones have replaced home-based, landline phones for many people and neither study explicitly includes mobile telephones.	Partially
Balance of benefit and disadvantage to participants	There is little or no direct benefit to participants, although some may like being reminded about the trial. One potential disadvantage is that some participants may be irritated by the reminder call, but what proportion would be irritated is unclear.	Yes
Balance of benefit and disadvantage to the host trial	The benefit to the host trial is a small increase in recruitment if underlying recruitment is low but it is unclear what the benefit would be if underlying recruitment was higher. There is a potential disadvantage to the host trial of over-burdening trial staff with making the reminder telephone calls, but the size of this disadvantage is unclear.	Yes

Table 5.2: Trial Forge Guidance 2: Using pens as a retention strategy

Trial Forge Guidance 2 (Using pens as a retention strategy)		
Criterion	**Comments**	**Criterion Met (Yes or No or Partially)**
GRADE rating of evidence	Found to be "low" (Gillies et al., 2021). Since the rating is lower than "high", this criterion is met.	Yes
Cumulated evidence	Five SWATs related to this strategy were undertaken. According to *Figure 5.1*, the RD has converged.	No
PICOT	P: No SWAT achieved sufficient sample size from young adult men. (Sharp et al., 2006) - UK, secondary care setting, mean age: 34 (age range: 20-59), females only with low-grade abnormal cervical smear living in Tayside, Grampian or Nottingham (Bell et al., 2016) - UK, primary care setting, age range: 70-85, females only at risk of fracture (Cunningham-Burley et al., 2020)- UK, secondary care setting (NHS staff), mean age: 43 (age range: ≥18), males and females irrespective of any health conditions (James et al., 2021) - UK, community setting, mean age: 79.7 (age range: ≥65), able to walk 10 feet, males and females irrespective of any health conditions (Mitchell et al., 2021) - UK, secondary care setting, mean age: 66.8 (age range: 70-85, females only irrespective of any health conditions I: The interventions are health-based, but some of them are treatment-related while others are prevention-related. Also, the health conditions are differential among the host trials. (Sharp et al., 2006) - Colposcopy as a diagnostic tool. (Bell et al., 2016) - 10-year fracture risk assessment using a WHO algorithm to reduce the risk of fracture (Cunningham-Burley et al., 2020) - slip-resistant footwear to prevent slips (James et al., 2021) - home environmental assessment and modification to reduce the risk of falling (Mitchell et al., 2021) - Community-based screening programme of fracture risk to reduce hip fractures C: There is homogeneity in the comparators of the five host trials. (Sharp et al., 2006) - A six-monthly cervical smear. (Bell et al., 2016) - Usual care from healthcare professional (Cunningham-Burley et al., 2020) - Usual footwear (James et al., 2021) - Usual care from healthcare professional (Mitchell et al., 2021) - Usual care O: The questionnaire response rate was reported as an outcome in all included SWATs and is relevant for the RD in retention rates between pen and no pen. T: Sending questionnaires via post remains a popular practice during an RCT.	Partially
Balance of benefit and disadvantage to participants	Providing a pen may be perceived as useful to some participants, but not useful to some others. The mechanisms according to which a participant would be more likely to feel benefited by a pen are unknown.	Yes
Balance of benefit and disadvantage to the host trial	There is a clear benefit to the host trials, as adding pens for follow-up may improve the retention rate at a minimal incremental cost (see Chapter 2.2 for more information).	No

5.2.5. Fundamentals of Value of Information (VoI) analysis

Before introducing VoI analysis for estimating the benefits of additional SWATs to reduce the uncertainty on the effectiveness of the two recruitment and retention strategies, the rationale behind VoI analysis in a health economics framework should be explained. In health technology assessment (HTA), VoI analysis is commonly undertaken after primary cost-utility analyses of technologies in order to estimate the value, related to health outcomes and costs, of conducting further research on important input parameters likely to reduce the existing uncertainty on the primary results of such evaluations (Briggs et al., 2006). Such input parameters could be health utilities, costs, prior probabilities and treatment effects associated with a health condition of interest and/or a proposed health technology under evaluation. An original VoI analysis requires a pre-planned decision model, followed by probabilistic sensitivity analysis (PSA). *Decision models* synthesise the available evidence and consider all possible scenarios, related to a specific health condition of interest, that could arise from a set of alternative health technologies that are evaluated through cost-utility analysis (Briggs et al., 2006). A *probabilistic sensitivity analysis (PSA)* can capture the uncertainty in the results of a decision model through the estimation of the joint uncertainty across all input parameters.

Using the prior input parameters of the model, the values of which are usually different with respect to the technology under evaluation, it is feasible to estimate the probabilities of each scenario occurring. Each scenario also has its own costs and benefits, typically expressed in terms of a health outcome. Therefore, for each technology under evaluation, their expected benefits and costs are calculated through the summation of all benefits and costs from each scenario, weighted by the probability of each scenario occurring (Briggs et al., 2006). Given the estimation of the expected benefits and costs associated with each health technology, the cost-effectiveness of each technology can be subsequently evaluated so that the decision maker may choose which of the recommended technologies maximises the population's expected net benefits under a pre-specified cost-effectiveness threshold, where the latter sets an upper limit of how much a decision maker is willing to pay to gain an additional unit of health effect or quality-adjusted life year (QALY). A QALY captures the changes in the quality of life and life expectancy as a result of adopting a proposed technology and its scale lies between 0 (i.e. death) and 1 (i.e. perfect health) in a given year.

The primary outcome related to a cost-utility analysis, along with the decision model, is the incremental cost-effectiveness ratio (ICER). This can be expressed in terms of incremental cost per additional QALY gained. For a technology to be cost-effective, the ICER needs to be lower than the cost-effectiveness threshold. An outcome which combines the ICER with the cost-effectiveness threshold is the incremental net benefit (INB), which for a proposed technology A versus a baseline technology B is equal to:

$INB_A = incremental\ effect_{A,B} * cost\text{-}effectiveness\ threshold - incremental\ cost_{A,B}$ *(Equation 5.1)*

where $incremental\ effect_{A,B}$ is the increment in the expected health effects associated with technology A versus technology B and $incremental\ cost_{A,B}$ is the increment in the expected costs associated with technology A versus technology B. If $INB_A > 0$, technology A is cost-effective relative to technology B. If $INB_A \leq 0$, technology A is equally or less cost-effective relative to technology B.

Alternatively, the INB for technology A versus technology B (INB_A) can be expressed as:

$INB_A = incremental\ net\ benefit_A - incremental\ net\ benefit_B = INB_A - INB_B$ *(Equation 5.2)*

Under PSA, "*probability distributions are applied to the specified ranges for the key parameters and samples drawn at random from these distributions*" (Drummond et al., 2015). Such distributions can represent the uncertainty around the mean estimate of variables, i.e. the typical outcomes associated with cost-utility analysis such as incremental cost-effectiveness ratio (ICER) or incremental net benefit (INB). The types of distributions used for random sampling, e.g. Bernoulli or Dirichlet distribution, depend upon the nature of the input parameters in a specific decision model. To estimate the overall uncertainty, values from such distributions are randomly selected for each parameter input using Monte Carlo simulations. PSA is thus an essential tool for capturing the volume of uncertainty in the original findings as well as the underlying sources of such uncertainty. PSA has been previously applied in *Chapter 2* to assess the robustness of the cost-effectiveness findings with respect to improving patient recruitment to the RECOVERY trial.

To capture parameter uncertainty, it is recommended that a cost-effectiveness acceptability curve (CEAC) be generated; such a curve demonstrates the likelihood that a technology is cost-effective when the threshold changes within a predetermined range (Briggs et al., 2006). Under uncertainty, e.g. due to uncertain input parameters related to the effectiveness of treatments A and B, it is possible that technology A has a higher INB, but not statistically significant, than that of technology B (and hence $INB_A > 0$, since $INB_A > INB_B$ from *Equation 5.2*), but still has

a significant error probability, according to the CEAC, that could prevent its approval in a given threshold. Such disapproval could generate opportunity costs on patients who could benefit from the foregone INBs following the rejection of technology A. As a result, there is an expected cost of uncertainty, the estimate of which is based upon the probability of error from the CEAC and the opportunity costs of decision error that include the foregone health benefits and resources due to the rejection of technology A. Such an expected cost is defined as *the expected value of perfect information (EVPI)* related to the comparison of technology A to technology B, since having access to perfect information will remove the probability of any underlying uncertainties causing a wrong decision to be made (Briggs et al., 2006). Perfect information could arise from additional investment in research that could provide clear evidence on input parameters responsible for the imperfect information under the existing evidence. Such additional research could take place in the form of any research design, such as an RCT or a systematic review.

There are three factors that influence the estimation of EVPI: 1) the INB of technology A (INB_A) and the INB of technology B (INB_B) according to the existing evidence; 2) the distribution of INB_A according to the existing evidence, which could take the form of a linear (e.g. Normal) or a non-linear distribution, and; 3) a loss function which considers the opportunity costs of decision error according to the existing evidence (Briggs et al., 2006). An illustrative example of EVPI is provided in Briggs et al. (2006), based upon five iterations of Technology A and Technology B, and is presented on *Table 5.3.*

Table 5.3: An example of expected value of perfect information (EVPI), adapted from Briggs et al. (2006)

	Technology A	**Technology B**	**Optimal choice**	**Maximum net benefit**	**Opportunity loss**
Iteration 1	9	12	B	12	0
Iteration 2	12	10	A	12	2
Iteration 3	14	20	B	20	0
Iteration 4	11	10	A	11	1
Iteration 5	14	13	A	14	1
INB	*12*	*13*	*B*	*13.8*	*0.8*

According to the above figures, originating from five iterations and a pre-specified cost-effectiveness threshold, the optimal decision, under the existing evidence, is to choose Technology B over Technology A, as its expected net benefit is greater than that of Technology A ($INB_B=13>INB_A=12$). However, if perfect information was available, the decision maker would optimally choose Technology A over Technology B in Iterations 2, 4 and 5, as its observed net benefits are greater than those of technology B (i.e. $INB_{Iteration2,B}=10<ONB_{Iteration2,A}=12$, $INB_{Iteration4,B}=10<INB_{Iteration4,A}=11$, $INB_{Iteration5,B}=13<INB_{Iteration5,A}=14$), and Technology B over Technology A in Iterations 1 and 3 (i.e. $INB_{Iteration1,B}=12> INB_{Iteration1,A}=9$, $INB_{Iteration3,B}=20> INB_{Iteration3,A}=14$).

By adopting Technology B over Technology A across all iterations, the net benefits in Iterations 1 and 5 are maximised but the net benefits in the remaining iterations are not, as they generate opportunity losses due to the rejection of a more cost-effective technology. Therefore, by choosing Technology B in Iterations 1 and 3, and Technology A in Iterations 2, 4 and 5, the expected net benefit under perfect information would have been equal to 13.8, a figure greater than that of the expected net benefit under imperfect information based on choosing Technology B across all iterations (=13). The difference between the expected net benefit under perfect information and the expected net benefit under the existing information is called the expected opportunity cost, in other words the *expected value of perfect information (EVPI)*. It is also worth noting that whereas technology B has a higher overall expected net benefit than Technology A, the likelihood of it being cost-effective is only 0.4 under the given threshold.
As the EVPI is initially expressed in individual terms, it is important: to aggregate this figure across the relevant population; to decide the period over which information about the decision of which treatment to approve will be useful; and to find estimates of the incidence of the disease associated with Treatments A and B so as to generate a population-wide EVPI. The population-level EVPI should also be discounted at a pre-specified annual rate to consider the overall EVPI for both current and future patients.

If it costs £x overall to acquire additional information and hence a maximum expected net benefit for the population, a value of perfect information curve, with Population EVPI on the vertical axis and cost-effectiveness threshold on the horizontal axis, could be drawn. In this way, it can be shown under which range of cost-effectiveness thresholds it would have been cost-effective to invest £x to acquire perfect information. As long as the EVPI is greater than the cost £x of acquiring perfect information, it would be cost-effective to seek perfect

information on uncertain input parameters, as specified according to a prior PSA. With proper modelling, it is feasible to vary the net benefits presented in *Table 5.2* with different levels of cost-effectiveness threshold so as to generate an appropriate value of perfect information curve.

5.2.6. Value of Information (VoI) analysis in the framework of meta-analysis

The VoI analysis in *Section 5.2.4* is applicable to situations where: 1) the benefits of proposed technologies are associated with maximising the QALYs gained for a population; 2) the source of uncertainty originates from input parameters of an appropriate decision model. Such benefits are evaluated, along with a cost-effectiveness threshold, to determine which technology is more cost-effective according to the obtained INB or ICER estimates. Under this framework, decision modelling, followed by subsequent PSA, is necessary to estimate the EVPI for eliminating parameter uncertainty. Following the aggregation of EVPI to generate the population-level EVPI, it is compared against the fixed cost of acquiring perfect information of input parameters at a pre-specified cost-effectiveness threshold, to determine whether acquiring perfect information would be a cost-effective strategy.

In the framework of using telephone reminders as a recruitment strategy, or pens as a retention strategy, for patients in RCTs, however, the observed benefits from previous SWATs are not QALY-dependent, as they focus explicitly on the efficacy of telephone reminders or pens in improving the recruitment or retention rates. Therefore, outcomes such as INB or EVPI are not applicable for VoI analysis of SWATs, as the components of the former are directly associated with health-related outcomes, such as QALYs, by definition (see *Section 5.2.4*). In addition, the source of uncertainty, in the framework of telephone reminders and pens, would not be input parameters since the evidence does not originate from decision models but from original primary SWATs. Therefore, the source of uncertainty would instead be the 95% CIs around the RD acquired from the existing two swats (five) SWATs of using telephone reminders (pens) as a recruitment (retention) strategy, i.e. *effect uncertainty*. Therefore, an alternative methodology for VoI analysis needs to be adopted for exploring the benefits of additional research on SWATs of recruitment and retention strategies. A study by Claxton et al. (2015a) suggests that "*a simple extension of standard meta-analysis can provide quantitative estimates of the potential health benefits of further research and of implementing the findings of existing research, which can help inform research prioritisation and efforts to change clinical practice*" (Claxton et al., 2015a).

The results of a meta-analysis for all strategies under comparison are typically expressed in terms of an absolute measure, such as follow-up (i.e. questionnaire return) rate in retention SWATs. For such measures, 95% CIs expressing a range of statistically reasonable values of the effect are also reported. The distribution of such values lying within the 95% CI can reflect the uncertainty with respect to the relative effect size of a given proposed intervention or strategy (Claxton et al., 2015a). If such an effect uncertainty is unacceptable, it is recommended that further research occur since the costs of implementing it are justified given the underlying uncertainty (Claxton et al., 2015a). Instead, if such an effect uncertainty is acceptable, it is not recommended that further research occur since the costs of implementing it are not justified given the underlying uncertainty (Claxton et al., 2015a). To observe the importance of effect uncertainty and its broader consequences, it should be aggregated in active RCT participant population figures in the framework of retention SWATs (Claxton et al., 2015a). Following population aggregation, by using a relative measure of effect, such as risk ratio (RR), risk difference (RD) or odds ratio (OR), it becomes feasible to predict the estimated loss in a desirable outcome due to effect uncertainty, such as recruitment rate in SWATs of recruitment strategies or follow-up rate in SWATs of retention strategies. Such a loss describes the expected consequences of uncertainty. In the present framework, such consequences are not expressed in terms of foregone QALYs, but rather as foregone improvements in recruitment or follow-up rates for the active RCT population, due to potentially approving a less effective strategy or intervention (Claxton et al., 2015a). In contrast to the classical framework of VoI analysis, further research through SWATs is not anticipated to eliminate the underlying effect uncertainty, but could still reduce it enough such that more certain decisions about which retention strategy to adopt could be made, and therefore conclude whether investing £x for additional research on this strategy would or would not be a cost-effective thing to do (Claxton et al., 2015a).

5.2.7. Value of Information (VoI) analysis in the framework of SWATs

Available in the literature are examples of VoI analyses following the extension of meta-analyses of clinical RCTs, whose methodology could feasibly be applied in meta-analyses of SWATs, since a SWAT itself presents similar -if not identical- features of randomisation and statistical inference to a standard RCT (Claxton et al., 2016, McKenna et al., 2016, McKenna et al., 2015).

For instance, by placing the five existing SWATs related to pens in chronological order (Sharp et al., 2006, Bell et al., 2016, Cunningham-Burley et al., 2020, James et al., 2021, Mitchell et al., 2021), consecutive cumulative random effect meta-analyses of using pens versus not using pens for participant retention could be gradually undertaken to observe the evolvement of relative measures of effect (i.e. RD) over time. The 95% CIs and the p-values could be simultaneously estimated and used for predicting the extent of effect uncertainty throughout the consecutive meta-analyses. RevMan software was used for these meta-analyses. For each meta-analysis, an uncertain distribution of 5000 combinations of RD values (i.e. in line with the reported 95% CIs) and baseline retention rates (e.g. from 80% to 97%), could be generated through random sampling, via Microsoft Excel, to capture the uncertainty with respect to the retention rates of pens versus no pens. Such uncertainty was aggregated in the most recent English trial community figures, that is 1,390,483 participants involved in clinical research across England from April 2020 to March 2021 (National Institute for Health & Care Research, 2021) However, since many trials are expected not to include postal questionnaires as part of their data collection, as happens in the case of pens as a retention strategy, it is assumed that 25% of all trials include questionnaires during follow-up. Therefore, by assumption, the annual population is 0.25*1,390,483= 347,621 trial participants. The costs of additional research through SWATs could be estimated via the funding that the National Institute for Health & Care Research (NIHR) currently provides, i.e. up to £30,000 for an additional SWAT into every Health Technology Assessment (HTA) trial (National Institute for Health & Care Research, 2023). Nevertheless, four out of five of the considered SWATs were funded through the PROMETHEUS programme (currently expired), which provided a funding of up to £5,000 for each SWAT (Clark et al., 2022). Therefore, it is assumed that a new SWAT of pens would cost up to £5,000, in line with the funding that the already undertaken SWATs of pens had received. Also, since pens are typically included in pre-paid envelopes that contain questionnaires, the reported cost per additional participant retained was found to be £57.90 given the existing evidence (Murphy et al., 2022b). This figure is based upon a hypothetical randomised trial undertaken in the UK, consisting of 500 participants, 10 trial sites, 1-year participant follow-up and triple trial visits (Murphy et al., 2022b).

For each consecutive meta-analysis, to determine whether investing in an additional SWAT of pens as a retention strategy is a cost-effective procedure, the following factors must be taken into consideration: the expected number of trial participants withdrawing from trials due to uncertainty; the expected incremental number of trial participants remaining in trials as a result

of receiving pens during follow-up; the cost per additional participant retained using pre-paid envelopes (questionnaire) in a future trial (£57.90); and the available funding for an additional SWAT (£5,000). If the value of implementation (i.e. the gains in retention rates when establishing pens as a retention strategy across all host trials, given the existing evidence) is sufficiently larger than the value of additional research (i.e. the expected number of trial participants withdrawing from trials due to uncertainty), it becomes highly likely that investing in an additional SWAT on pens may not be cost-effective, given the existing budget (McKenna et al., 2016, Claxton et al., 2015a).

For illustrative purposes, let us assume a scenario of meta-analysis, similar to McKenna et al. (2016), of two SWATs comparing pen arms versus no-pen arms where:

- The cumulative retention rate in the pen arm is 0.79 (=316/400)
- The cumulative retention rate in the no pen arm is 0.773 (=402/520)

The RD is 0.02, with the lower bound of the 95% CI being -0.04 and the higher bound of the 95% CI being 0.07. The reported 95% CI suggests that using pens could be up to 7 % effective in improving retention rates in RCTs, against no pen, but it also means that using pens could be up to 4% less effective in improving retention rates, against no pen. The baseline retention rate is assumed to be the median retention rate observed in NIHR-funded trials, which equals 88%, with its interquartile range (IQR) being 80%-97% (Jacques et al., 2022).

A process of random sampling, i.e. bootstrapping, from uncertain distributions of RD figures lying in the above 95% CI and baseline retention rates (from 80% to 97%), could be undertaken, with 5000 (or more) simulations, in order to generate a relevant non-linear distribution which would capture the effect uncertainty of pens as a retention strategy (McKenna et al., 2016). The random sampling process presents a 69.43% chance of pens being a more effective retention strategy than usual follow-up, and a 30.57% chance of excess attrition rates because of using pens as a retention strategy. Furthermore, given 5,000 random combinations of baseline retention rates and RDs, it can be assumed that the average of their product generates 6,330, when multiplied by the trial population figure of 347,621 and then subtracted by the average product of this population figure and a random baseline retention rate: this is the expected number of additional trial participants remaining across all trials because of receiving a pen during follow-up. In addition, given a 30.57% chance of increased attrition because of adding pens during follow-up, the average of the product of the trial population and the negative difference between the baseline retention rate and the product of

the baseline retention and RD reflects the expected number of trial participants lost to follow-up as a result of effect uncertainty, a figure equal to 1,380 in this example. Therefore, *the value of additional research* is 1,380 and that *the value of implementation* is 6,330. Assuming the incremental cost per additional participant retained across trials using envelopes is £57.90, the NIHR would need to fund (£57.90* 6,330=) £366,507/£5,000=73 SWATs, in order to generate similar incremental benefits in retention rates through research with SWATs of alternative retention strategies. In addition, however, 1,380/ (£5000/£57.90) =16 further SWATs of pens would be needed to reduce effect uncertainty associated with 1,380 participants lost to follow-up. Each of these further SWATs would be expected to reduce uncertainty by £5000, i.e. the maximum funding available allocated, divided by £57.90, i.e. the incremental cost of retaining an additional participant via postal questionnaires, which is equal to approximately 86 participants, which means that 16 further SWATs would be needed to reduce uncertainty significantly. Evidently, by comparing the value of additional research with the value of implementation, it is uncertain whether more SWATs of pens as a retention strategy need to be funded by the NIHR in this example.

5.2.8. Aims of Chapter 5

The direct aim of this chapter is to observe whether, given the existing evidence, investing in an additional SWAT or telephone reminders as a recruitment strategy and pens as a retention strategy would be cost-effective research decisions for reducing the uncertainty of the cumulative RD reported in *Figure 5.1* and *Figure 5.2*. By considering the expected benefits and costs of additional research, the chapter, through the VoI analysis framework in *Section 5.2.6*, will determine whether funding additional SWATs of telephone reminders as a recruitment strategy and pens as a retention strategy would be efficient investments. Depending upon the findings, the VoI analysis will be compared with the findings encouraging further SWATs for pens as a retention strategy, according to Trial Forge Guidance 2, as shown in *Table 5.1*. The indirect aim of this chapter is to introduce a framework for reducing research waste by considering the cost-effectiveness of undertaking further SWATs on recruitment and retention strategies, through the application of VoI analysis techniques discussed in this chapter.

5.3. Methods

5.3.1. Meta-analysis in the framework of SWATs of telephone reminders as a recruitment strategy

The characteristics of the two SWATs associated with telephone reminders to non-responders as a recruitment strategy, introduced in *Figure 5.1*, are presented in *Table 5.1* (Nystuen and Hagen, 2004, Wong et al., 2013). The primary outcome across all studies is the recruitment rate in the control and intervention groups. Such an outcome is expressed in terms of proportion of patients recruited to a trial (see *Table 5.4*). To compare the difference in recruitment rates between the trial arms, i.e. the effectiveness of the provision of telephone reminders to non-responders as a recruitment strategy, the Cochrane recruitment review used the risk difference (RD) as a measure of effect, since the recruitment rate is a dichotomous variable (Gillies et al., 2021). To ensure methodological homogeneity, the RD is also the statistical measure of effect in this chapter. If RD>0, using telephone reminders to non-responders is a more effective recruitment strategy than no reminders. If RD=0, using telephone reminders is neither a more nor a less preferable retention strategy than no reminders. If RD<0, using pens is a less effective retention strategy than no reminders.

Table 5.4: Characteristics of the two SWATs related to using telephone reminders to non-respondents as a recruitment strategy, adapted from Treweek et al. (2018b)

Study	**Methods**	**Data**	**Comparisons**	**Outcomes**	**Control recruitment rate**
Nystuen et al. (2004)	Randomised controlled trial	Norway, community setting.	Effects of different telephone reminders. Written invitation to participate in a community-based trial followed by a telephone reminder if no response within 2 weeks; guide used for discussion. This was compared to written invitation to participate in a community-based trial followed by no reminder if no response within two weeks.	Proportion recruited to trial	4.55 %

Wong et al. (2013)	Randomised controlled trial	Canada, primary care setting.	Investigated use of telephone reminders to non-responders Intervention: up to three telephone reminders to those not responding to initial posted invitation Comparison: no telephone reminders (but did get a second invitation)	Proportion recruited to trial	7.42 %

The two types of model commonly used for meta-analysis are the fixed-effect and the random-effect model. In the former it is assumed that any effect differences between the included studies are due to random error (i.e. sampling variability (Riley et al., 2011)), whereas in the latter the variability of the effect between studies is due to variations in the magnitude of the effect across studies, in addition to random error (Riley et al., 2011). A crucial element in determining whether a random-effect (fixed-effect) should be preferred to a fixed-effect (random-effect) model is to observe the characteristics of the studies included and assess whether they do (not) present substantial heterogeneity regardless of the findings from common heterogeneity tests, such as the I^2 statistic. In the presence of substantial heterogeneity, i.e. variation in terms of design and methodology of the studies, it is highly likely that a variation in the size of the effect across studies actually exists, and hence the assumption related to a fixed-effect model does not hold and a random-effect model should be adopted instead.

Whereas the Cochrane review on recruitment strategies applied fixed-effect models (Treweek et al., 2018b), this chapter applies random-effect models to undertake the corresponding meta-analysis, labelled as *Meta-Analysis 1*, so as to account for potential heterogeneity. Given the description of the SWATs in *Table 5.4*, there is some evidence for substantial heterogeneity among the included studies, due to different clinical settings (primary care vs community setting) and location (Canada vs Norway). The reviewers of the Cochrane recruitment review do not explain why they used fixed-effect models (Treweek et al., 2018b). For this recruitment strategy, an applied random-effects model using the Mantel-Haenszel method of weighting the contribution of each included SWAT to the overall RD estimate is applied. The meta-analysis is undertaken on Reviewer Manager (RevMan) software, using random-effect models with the Mantel-Haenszel estimation method.

5.3.2. Cumulative meta-analyses in the framework of SWATs of pens as a retention strategy

The five SWATs associated with pens as a retention strategy, as introduced in *Figure 5.2*, were placed in chronological order (by date of publication) to observe how the evidence on using pens as a retention strategy has evolved, over time, through cumulative meta-analyses (Claxton et al., 2015a). This order is the following:

- Sharp et al. (2006)
- Bell et al. (2016)
- Cunningham-Burley et al. (2020)
- James et al. (2021)
- Mitchell et al. (2021)

The characteristics of the five studies, including methods, data, comparisons and outcomes, were collected from the Cochrane review (Gillies et al., 2021) and are shown in *Table 5.4*.The primary outcome across all studies was the retention rate in the control group, i.e. usual follow-up without pens, and in the intervention group, i.e. pens in addition to usual follow-up. Such an outcome could be expressed in terms of questionnaire response rate or questionnaire return rate in the included SWATs (see *Table 5.4.*). To compare the difference in retention rates between the trial arms, i.e. the effectiveness of pens as a retention strategy, the Cochrane retention review used the risk difference (RD) as a measure of effect (Gillies et al., 2021). To ensure methodological homogeneity, the RD is also the statistical measure of effect used in the VoI analysis of pens. as a retention strategy.

In line with the Cochrane review (Gillies et al., 2021), this chapter also applies random-effect models to undertake the corresponding meta-analyses. Given the description of the SWATs in *Table 5.5*, however, there is insufficient evidence for substantial heterogeneity among the included studies. Also, the reviewers of the Cochrane retention review do not explain why they used random-effect models (Gillies et al., 2021). Nevertheless, for purposes of homogeneity with the Cochrane retention review's methodology, this chapter applies a random-effects model, using the Mantel-Haenszel method of weighting the contribution of each included SWAT to the overall RD estimate.

The following four cumulative meta-analyses are undertaken to estimate the RD of pens plus usual follow-up as a retention strategy, compared to no pens and usual follow-up (Claxton et al., 2015a):

- *Meta-Analysis 2.1:* Sharp et al. (2006) and Bell et al. (2016)
- *Meta-Analysis 2.2:* Sharp et al. (2006), Bell et al. (2016) and Cunningham-Burley et al. (2020)
- *Meta-Analysis 2.3:* Sharp et al. (2006), Bell et al. (2016), Cunningham-Burley et al. (2020) and James et al. (2021)
- *Meta-Analysis 2.4:* Sharp et al. (2006), Bell et al. (2016), Cunningham-Burley et al. (2020), James et al. (2021) and Mitchell et al. (2021)

Table 5.5: Characteristics of the five SWATs related to using pens as a retention strategy, adapted from Gillies et al. (2021)

Study	**Methods**	**Data**	**Comparisons**	**Outcomes**	**Control retention rate**
Sharp et al. (2006)	2x2x2 factorial parallel, individually-randomised trial	UK, secondary care setting.	Three trials evaluated. 1) TAMBOLA-branded pen versus no pen 2) Questionnaire dispatched by first class post versus questionnaire dispatched by second class post 3) Pre-addressed return envelope with second-class postage stamp versus freepost business-reply envelope.	Questionnaire response rate. Retention period: 12, 18, 24 and 30 months	64.24%
Bell et al. (2016)	Parallel, individually-randomised trial	UK, primary care setting.	Intervention group received trial-branded pen with the 60-month follow-up questionnaire. Control group 60-month follow-up questionnaire alone.	Questionnaire return rate Retention period: 60 months	91.27%
Cunningham-Burley et al. (2020)	Parallel, individually-randomised trial	UK, secondary care setting.	Intervention group received a branded pen with their questionnaire. Control group did not receive a pen.	Proportion of participants who return the questionnaire. Retention period: 14 weeks	64.69%

James et al. (2021)	2x2 factorial parallel, individually-randomised trial	UK, community setting.	Intervention group 1 received a branded pen and a standard cover letter. Intervention group 2 received a branded pen and a social incentive cover letter. Intervention group 3 received no pen and a social incentive cover letter. Control group received no pen, standard cover letter. All participants received an unconditional £5 note with the questionnaire.	Questionnaire return rate. Retention period: 12 months	95.78%
Mitchell et al. (2021)	Parallel, individually-randomised trial	UK, secondary care setting.	Intervention group received a pen alongside the 12-month questionnaire. Control group did not receive any pen alongside the 12-month questionnaire.	Questionnaire return rate. Retention period: 12 months	85.61%

All meta-analyses are undertaken on RevMan software, using random-effect models with the Mantel-Haenszel estimation method.

5.3.3. Value of Information (VoI) analysis in the framework of SWATs of telephone reminders as a recruitment strategy

Following the introduction of the framework for undertaking VoI related to SWATs in *Section 5.2.6*, the following components were considered for VoI in the framework of SWATs of telephone reminders as a recruitment strategy:

- For *Meta-Analysis 1*, the recruitment rates in the telephone reminder group and the no reminder group were reported and expressed as follows:

$$\text{RecruitmentRate}_{\text{REMINDER}} = n_{\text{REMINDER}}/N_{\text{REMINDER}} \quad \textit{(Equation 5.3)}$$

where n_{REMINDER} is the number of participants recruited to a host trial from the telephone reminder group, and N_{REMINDER} is the total number of participants allocated to the telephone reminder group. The recruitment rate is defined as the percentage of potentially eligible patients eventually recruited to a randomised trial.

$RecruitmentRate_{NOREMINDER} = n_{NOREMINDER}/N_{NOREMINDER}$ *(Equation 5.4)*

where $n_{NOREMINDER}$ is the number of participants recruited to a host trial from the no reminder group, and $N_{NOREMINDER}$ is the total number of participants allocated to the no reminder group.

- For *Meta-Analysis 1*, the RD in recruitment rate between the telephone reminder arm and the no reminder arm were reported and labelled as $RD_{REMINDER}$. The corresponding 95% CI lower and upper bounds were also reported.
- For *Meta-Analysis 1*, the baseline recruitment rate, i.e. proportion of eligible participants randomised, was assumed to be equal to the most accurate median recruitment rate of 43% (95% CI: 37.2% to 48.7%), as estimated for cancer trials (Reynolds et al., 2023).

Therefore,

$BaselineRecruitment = 0.43$ *(Equation 5.5)*

- In line with the methodology of estimating the additional benefits of health research through meta-analysis (McKenna et al., 2016), an estimate of the population affected by trial methodology-related research is needed. The most relevant estimate to date could be the number of people estimated to have a long-term condition (LTC) across the UK, who would be more likely to be invited to participate in a randomised trial. According to NHS estimates, approximately 26 million people in the UK live with at least one LTC (NHS, 2018). Assuming that 10% of such patients would be invited and then would consent to participate in a clinical trial in a given year, the population estimate is 2,600,000.
- Estimates of the population recruited in randomised trials were obtained. Such estimates were differentiated by intervention arm (i.e. telephone reminders vs no reminder). The equation representing the expected population of patients invited and recruited to trials if they received a telephone reminder during the recruitment stage is:

$$\begin{aligned} PopulationRecruitment_{REMINDER} &= \text{Population of invited patients} * JointRecruitment_{REMINDER} \\ &= \text{Population of invited patients} * BaselineRecruitment \times (RD_{REMINDER} + 1) \\ &= 2{,}600{,}000 * BaselineRecruitment * (RD_{REMINDER} + 1) \end{aligned}$$ *(Equation 5.6)*

where $JointRecruitment_{REMINDER} = BaselineRecruitment * (RD_{REMINDER} + 1)$ *(Equation 5.7)*

$JointRecruitment_{REMINDER}$ is the product of the baseline recruitment rate and the risk difference (RD) in recruitment rates between the telephone reminder and the no reminder groups plus 1. The use of "plus 1" is necessary, since the RD in this case reflects an increase in exposure to a positive outcome, i.e. recruitment to a randomised trial. 5000 random combinations of

BaselineRecruitment and $RD_{REMINDER}$ were generated through random sampling in Microsoft Excel in order to estimate the expected value of additional research (i.e. $VOI_{REMINDER}$).

The corresponding equation representing the expected population of trial participants recruited to trials if they did not receive any telephone reminder is:

$PopulationRecruitment_{NOREMINDER}$= Population of invited patients * BaselineRecruitment

= 2,600,000 * BaselineRecruitment *(Equation 5.8)*

- Thus, the effectiveness of telephone reminder as a recruitment strategy, in terms of the population of invited patients, is expressed as the expected difference between $PopulationRecruitment_{REMINDER}$ and $PopulationRecruitment_{NOREMINDER}$

$Effectiveness_{REMINDER}$ = (5.6) – (5.8)

= $PopulationRecruitment_{REMINDER}$ - $PopulationRecruitment_{NOREMINDER}$

(Equation 5.9)

In line with McKenna et al. (2016), random sampling with respect to the uncertain distribution of $JointRecruitment_{REMINDER}$ (=BaselineRecruitment * ($RD_{REMINDER}$+1)) was undertaken in Microsoft Excel. A joint uncertain distribution with 5000 iterations was generated. $RD_{REMINDER}$ was randomly varied between its lower bound and its upper bound from the estimated 95%CI from *Meta-Analysis 1*, whereas BaselineRecruitment was randomly varied between 37.2% and 48.7%, to generate 5000 values of $JointRecruitment_k$, where k=1,..., 5000 is the labelled number of iteration .

Therefore, for each k,

$JointRecruitment_k$= $BaselineRecruitment_{,k}$ * ($RD_{REMINDERk}$ +1) *(Equation 5.10)*

where $BaselineRecruitment_{,k}$ = $Range_{BaselineRecruitment}$= (0.372, 0.487) (*Equation 5.11)*

and $RD_{REMINDERk}$ is the random value of $RD_{REMINDER}$ generated in iteration k.

For each k simulation, $JointRecruitment_{REMINDER,k}$ was multiplied by the population figure of 2,600,000 to obtain $PopulationRecruitment_{REMINDER,k.}$

$PopulationRecruitment_{REMINDER,k}$=Population of invited patients * $JointRecruitment_{REMINDER,k}$

=Population of invited patients * $BaselineRecruitment_k$ * ($RD_{REMINDER,k}$ +1)

= 2,600,000*$BaselineRecruitment_k$ * ($RD_{REMINDER,k}$ +1) *(Equation 5.12)*

Simultaneously, for each k simulation, $PopulationRecruitment_{NOREMINDERk}$ was estimated through a given generated value of $BaselineRecruitment_{k.}$

$PopulationRecruitment_{NOREMINDERk}$ = Population of invited patients * $BaselineRecruitment_k$

$$= 2{,}600{,}000 * \text{BaselineRecruitment}_k \qquad \textit{(Equation 5.13)}$$

Finally, for each k simulation, the effectiveness of telephone reminders as a recruitment strategy, in terms of the population of trial participants, is expressed as the expected difference between $\text{PopulationRecruitment}_{\text{REMINDER}k}$ and $\text{PopulationRecruitment}_{\text{NOREMINDER}k}$:

$$\text{Effectiveness}_{\text{REMINDER},k} = (5.12) - (5.13)$$

$$= \text{PopulationRecruitment}_{\text{REMINDER}k} - \text{PopulationRecruitment}_{\text{NOREMINDER}k} \qquad \textit{(Equation 5.14)}$$

where $\text{Effectiveness}_{\text{REMINDER},k} < 0$ implies that telephone reminders is not an effective recruitment strategy , compared to no reminders , $\text{Effectiveness}_{\text{REMINDER},k} = 0$ implies that telephone reminders is an equally effective recruitment strategy with no reminders, and $\text{Effectiveness}_{\text{REMINDER},k} > 0$ implies that telephone reminders is an effective recruitment strategy compared to no reminders.

Given that K=5000 iterations related to *Meta-Analysis 1*, it becomes feasible to estimate the probability of telephone reminders being an effective recruitment strategy (i.e. $P(\text{ReminderEffective}_k > 0)$) and the probability of telephone reminders not being an effective recruitment strategy (i.e. $P(\text{ReminderEffective}_k \leq 0)$). Such probabilities are frequentist and were estimated by measuring the frequency in which $\text{Effectiveness}_{\text{REMINDER},k} > 0$, $\text{Effectiveness}_{\text{REMINDER},k} = 0$ and $\text{Effectiveness}_{\text{REMINDER},k} < 0$ were generated following 5000 iterations for each meta-analysis. The algebraic expressions are provided below:

$$P(\text{ReminderEffective}_k > 0) = \frac{\sum_{k=1}^{5000}(Xk)}{k} \qquad \textit{(Equation 5.15)}$$

where for each iteration $X_k = 1$ when $\text{Effectiveness}_{\text{REMINDER},k} > 0$ and $X_k = 0$ when $\text{Effectiveness}_{\text{REMINDER},k} \leq 0$. Correspondingly,

$$P(\text{ReminderEffective}_k \leq 0) = 1 - \frac{\sum_{k=1}^{5000}(Xk)}{k} \qquad \textit{(Equation 5.16)}$$

By the law of total probability, it is always the case that:

$$P(\text{ReminderEffective}_k > 0) + P(\text{ReminderEffective}_k \leq 0) = 1 \qquad \textit{(Equation 5.17)}$$

In line with Claxton et al. (2015a), the value of additional research on telephone reminders as a recruitment strategy ($\text{VOI}_{\text{REMINDER}}$), given *Meta-Analysis 1*, can be estimated through the weighted average of $\text{Effectiveness}_{\text{REMINDER},k}$ conditional upon the iterations for which

$\text{Effectiveness}_{\text{REMINDER},k} \leq 0$, multiplied by $P(\text{ReminderEffective}_k \leq 0)$. In other words, VOI corresponds to the expected number of invited patients not recruited as a result of the effect uncertainty around telephone reminders as a recruitment strategy, multiplied by the probability of telephone reminders not being an effective recruitment strategy:

$$\text{VOI}_{\text{REMINDER}} = \left| \frac{\sum_{k=1}^{k} (\text{EffectivenessREMINDERk} | \text{EffectivenessREMINDERk} < 0)}{k} \right| * P(\text{RemindersEffective}_k \leq 0)$$

$$= 0 \quad \text{when } \text{Effectiveness}_{\text{REMINDER},k} \leq 0 \quad \text{when } \text{Effectiveness}_{\text{REMINDER},k} > 0 \qquad \textit{(Equation 5.18)}$$

The number of invited patients expected to be recruited to randomised trials as a result of receiving telephone reminders during the recruitment stage is expressed as the weighted average of $\text{Effectiveness}_{\text{REMINDER},k}$ across the 5000 iterations given *Meta-Analysis 1*. This is the value of implementation following *Meta-Analysis 1*:

$$\text{ValueOfImplementation}_{\text{REMINDER}} = E(\text{Effectiveness}_{\text{REMINDER},k}) = \frac{\sum_{k=1}^{5000} (\text{EffectivenessREMINDERk})}{k} \qquad \textit{(Equation 5.19)}$$

Following each meta-analysis, the value of additional research ($\text{VOI}_{\text{REMINDER}}$) needs to be compared with the value of implementation ($\text{ValueOfImplementation}_{\text{REMINDER}}$) to assess whether investing in a further SWAT of telephone reminders, costing £5,000, would be cost-effective. Considering that the estimate of the average cost per additional patient recruited is 88 Canadian Dollars (CAD) (which, if converted to GBP through the 2019 USD PPP rate is £48.38) and hence £54.61 at 2022 levels (through the use of GDP deflator from 2019 to 2022) (Kakumanu et al., 2019), this implies that an additional SWAT under a cost of £5,000 would increase the number of patients recruited to all trials, due to reduction in effect uncertainty by £5000/£54.61=92. Also, given an existing level of the value of implementation following each meta-analysis, it is important to estimate the amount the NIHR would need to invest to generate the same improvements in recruitment rates, through additional evidence on other recruitment strategies, if telephone reminders were still not to be adopted as a recruitment strategy across all trials. For instance, if the value of implementation was £10,000, the NIHR would need to invest £54.61*10,000=£546,100 to generate similar additional benefits in terms of improved patient recruitment.

5.3.4. Value of Information (VoI) analysis in the framework of SWATs of pens as a retention strategy

VoI was undertaken for each meta-analysis (i.e. *Meta-Analysis 2.1, Meta-Analysis 2.2, Meta-Analysis 2.3, Meta-Analysis 2.4*) related to pens as a retention strategy. The purpose for doing so was to explore the degree of reduction in the uncertainty of the evidence following the addition of another SWAT, and consequently to estimate the expected benefits and costs should further SWATs be undertaken for evaluating the effectiveness of pens as a retention strategy. Following the introduction of the framework for undertaking VoI related to SWATs in *Section 3.2.6,* the following components were considered for VoI in the framework of SWATs of pens as a retention strategy:

- For each meta-analysis (i=*Meta-Analysis 2.1, Meta-Analysis 2.2, Meta-Analysis 2.3, Meta-Analysis 2.4*), the overall retention rates in the pen plus usual follow-up arm and the no pen (i.e. usual follow-up) arm were reported and expressed as follows:

$$RetentionRate_{PENi} = n_{PENi}/N_{PENi} \quad \textit{(Equation 5.20)}$$

where n_{PEN} is the number of participants who were randomised in the pen arm *and* did not withdraw from the host trials following the receipt of pen (plus usual follow-up), and N_{PEN} is the total number of participants who were randomised in the pen arm.

$$RetentionRate_{NOPENi} = n_{NOPENi}/N_{NOPENi} \quad \textit{(Equation 5.21)}$$

where n_{PEN} is the number of participants who were randomised in the no pen (i.e. usual follow-up) arm *and* did not withdraw from the host trials following usual follow-up, and N_{PEN} is the total number of participants who were randomised in the no pen (i.e. usual follow-up) arm.

- For each meta-analysis (i = *Meta-Analysis 2.1, Meta-Analysis 2.2, Meta-Analysis 2.3, Meta-Analysis 2.4*), the RD in retention rate between the pen (plus usual follow-up) arm and the no pen (i.e. usual follow-up) arm were reported. The corresponding 95% CI lower and upper bounds were also reported.

- For each meta-analysis (i =*Meta-Analysis 2.1, Meta-Analysis 2.2, Meta-Analysis 2.3, Meta-Analysis 2.4*), the baseline retention rate was assumed to be equal to the median retention rates of trials that were funded through NIHR, as the latter is the current main funding body for

SWATs across the UK. Thus, the baseline retention rate was assumed to be 88%, with an interquartile range of between 80% and 97% (Jacques et al., 2022).

Therefore,

$BaselineRetention_i = 0.88$ *(Equation 5.22)*

- In line with the methodology of estimating the additional benefits of health research through meta-analysis (McKenna et al., 2016), an estimate of the population affected by trial methodology-related research is needed. The most relevant estimate up to date could be the population figure of participants involved in clinical research across England, as provided by the NIHR. The figure for year 2020/21 was 1,390,483 (National Institute for Health & Care Research, 2021), adjusted to 347,621 by assuming that 25% of participants across all trials receive a postal questionnaire during follow-up.

- For each meta-analysis (i =*Meta-Analysis 2.1, Meta-Analysis 2.2, Meta-Analysis 2.3, Meta-Analysis 2.4*), estimates of the population being retained in all host trials were obtained. Such estimates were differentiated by intervention arm (i.e. pen vs no pen). The equation representing the expected population of trial participants remaining in trials if they received a pen during follow-up is:

$PopulationRetention_{PENi}$ = Population of trial participants * $JointRetention_i$

= Population of trial participants * $BaselineRetention_i$ * (RD_i +1)

= 347,621 * $BaselineRetention_i$ * (RD_i +1) *(Equation 5.23)*

where $JointRetentionRD_i = BaselineRetention_i * (RD_i + 1)$ *(Equation 5.24)*

$JointRetention_i$ is the product of the baseline retention rate and the risk difference in retention rates between the pen and the no pen arms plus 1. The use of "plus 1" is necessary, since the RD in this case reflects an increase in exposure to a positive outcome, i.e. retention of participants in a randomised trial. 5000 random combinations of $BaselineRetention_i$ and RD_i were generated through random sampling in Microsoft Excel in order to estimate the expected value of additional research (i.e. VOI_{PENS}).

The corresponding equation representing the expected population of trial participants remaining in trials if they did not receive a pen during follow-up is:

$PopulationRetention_{NOPENi}$ = Population of trial participants * $BaselineRetention_i$

$$= 347{,}621 * \text{BaselineRetention}_i \quad \textit{(Equation 5.25)}$$

- Thus, for each meta-analysis (i=*Meta-Analysis 2.1, Meta-Analysis 2.2, Meta-Analysis 2.3, Meta-Analysis 2.4*), the effectiveness of pens as a retention strategy, in terms of the population of trial participants, is expressed as the expected difference between $\text{PopulationRetention}_{PENi}$ and $\text{PopulationRetention}_{NOPENi}$:

$\text{Effectiveness}_{PENSi}$ = (5.25) – (5.23)

$$= \text{PopulationRetention}_{PENi} - \text{PopulationRetention}_{NOPENi} \quad \textit{(Equation 5.26)}$$

In line with McKenna et al. (2016), random sampling with respect to uncertain distribution of $\text{JointRetentionRD}_i$ (=$\text{BaselineRetention}_i$ x (RD_i+1)) was undertaken in Microsoft Excel. Following each meta-analysis, a joint uncertain distribution was generated with 5000 iterations. RD_i was randomly varied between its lower bound and its upper bound from the estimated 95%CI, whereas $\text{BaselineRetention}_i$ was randomly varied between 80% to 97%, to generate 5000 values of $\text{JointRetentionRD}_{i,k}$, where i=1,2,3,4 is the labelled number of meta-analysis and k =1,...,5000 is the labelled number of iteration . Thus, for each k and i,

$$\text{JointRetentionRD}_{i,k} = \text{BaselineRetention}_{i,k} * (RD_{i,k} + 1) \quad \textit{(Equation 5.27)}$$

where $\text{BaselineRetention}_{i,k} = \text{RetentionRate}_{PENi,k} * \text{RetentionRate}_{NOPENi,k.}$ *(Equation 5.28)*

and $RD_{i,k}$ is the random value of RD_i generated in k iteration corresponding to i meta-analysis. Then, for each k simulation and i meta-analysis, $\text{JointRetentionRDi}_{,k}$ was multiplied by the population figure of 347,621 to obtain $\text{PopulationRetention}_{PENi,k.}$

$\text{PopulationRetention}_{PENi,k}$ = Population of trial participants * $\text{JointRetentionRD}_{i,k}$

= Population of trial participants * $\text{BaselineRetention}_{i,k}$ *($RD_{i,k}$+1)

$$= 347{,}621 * \text{BaselineRetention}_{i,k} * (RD_{i,k} + 1) \quad \textit{(Equation 5.29)}$$

Simultaneously, for each k simulation and i meta-analysis, $\text{PopulationRetention}_{NOPENi,k}$ was estimated through a given generated value of $\text{BaselineRetention}_{i,k.}$

$\text{PopulationRetention}_{NOPENi,k}$ = Population of trial participants * $\text{BaselineRetention}_{i,k}$

$$= 347{,}621 * \text{BaselineRetention}_{i,k} \quad \textit{(Equation 5.30)}$$

Finally, for each k simulation and i meta-analysis, the effectiveness of pens as a retention strategy, in terms of the population of trial participants, is expressed as the expected difference between $\text{PopulationRetention}_{PENi,k}$ and $\text{PopulationRetention}_{NOPENi,k}$:

$\text{Effectiveness}_{PENSi,k}$ = (5.29) – (5.30)

$= \text{PopulationRetention}_{PENi,k} - \text{PopulationRetention}_{NOPENi,k}$ *(Equation 5.31)*

where $\text{Effectiveness}_{PENSi,k} < 0$ implies that pens is not an effective retention strategy compared to using no pens in follow-up , $\text{Effectiveness}_{PENSi,k} = 0$ implies that pens is an equally effective retention strategy with using no pens in follow-up, and $\text{Effectiveness}_{PENi,k} > 0$ implies that pens is an effective retention strategy compared to using no pens in follow-up.

Given K=5000 iterations in a given meta-analysis (i=*Meta-Analysis 2.1, Meta-Analysis 2.2, Meta-Analysis 2.3, Meta-Analysis 2.4*), it becomes feasible to estimate the probability that using pens is an effective retention strategy (i.e. $P(\text{PensEffective}_{i,k} > 0)$) and the probability that using pens is not an effective retention strategy (i.e. $P(\text{PensEffective}_{i,k} \leq 0)$). Such probabilities are frequentist and were estimated by measuring the frequency in which $\text{Effectiveness}_{PENi,k} > 0$, $\text{Effectiveness}_{PENi,k} = 0$ and $\text{Effectiveness}_{PENi,k} < 0$ were generated following 5000 iterations for each meta-analysis. The algebraic expressions are provided below:

$$P(\text{PensEffective}_{i,k} > 0) = \frac{\sum_{k=1}^{5000}(Xik)}{k}$$

(Equation 5.32)

where for each meta-analysis and iteration $X_{ik} = 1$ when $\text{Effectiveness}_{PENi,k} > 0$ and $X_{ik} = 0$ when $\text{Effectiveness}_{PENi,k} \leq 0$. Correspondingly,

$$P(\text{PensEffective}_{i,k} \leq 0) = 1 - \frac{\sum_{k=1}^{5000}(Xik)}{k}$$

(Equation 5.33)

By the law of total probability, it is always the case that:

$P(\text{PensEffective}_{i,k} > 0) + P(\text{PensEffective}_{i,k} \leq 0) = 1$ *(Equation 5.34)*

In line with Claxton et al. (2015a), the value of additional research on pens as a retention strategy ($VOI_{PEN,i}$), for each i meta-analysis, can be estimated through the weighted average of $\text{Effectiveness}_{PENi,k}$ conditional upon the iterations for which $\text{Effectiveness}_{PENi,k} \leq 0$, multiplied by $P(\text{PensEffective}_{i,k} \leq 0)$. In other words, VOI corresponds to the expected number of trial participants lost to follow-up as a result of the effect uncertainty around pens as a retention strategy, multiplied by the probability of pens not being an effective retention strategy:

$$VOI_{PEN,i} = \left|\frac{\sum_{k=1}^{k}(EffectivenessPENSi,k|EffectivenessPENSi,k<0)}{k}\right| * P\ (PensEffective_{i,k} \leq 0)$$

when $Effectiveness_{PENSi,k} \leq 0$

$= 0$ when $Effectiveness_{PENSi,k} > 0$ *(Equation 5.35)*

The expected number of trial participants expected to remain in RCTs as a result of receiving a pen during follow-up is expressed as the weighted average of $Effectiveness_{PENi,k}$ across the 5000 iterations for each i meta-analysis. This is the value of implementation following a given i meta-analysis:

$$ValueOfImplementation_{PENi} = E(Effectiveness_{PENi}) = \frac{\sum_{k=1}^{5000}(EffectivenessPENSi,k)}{k}$$

(Equation 5.36)

Following each meta-analysis, the value of additional research (VOI_{PENi}) needs to be compared with the value of implementation ($ValueOfImplementation_{PENi}$) to assess whether investing in a further SWAT of pens, costing £5,000, would be cost-effective. Considering that sending pre-paid envelopes, for participants responding to follow-up questionnaires, would cost £57.90 to retain one additional trial participant on average at 2022 price levels (Murphy et al., 2022b), this implies that an additional SWAT at a cost of £5,000 would reduce the number of participants withdrawing from all trials due to effect uncertainty by £5000/£57.90=86. Also, given an existing level of the value of implementation following each meta-analysis, it is important to estimate the amount the NIHR would need to invest in order to generate the same improvements in retention rates, through additional evidence on other retention strategies if pens were still not to be adopted as a retention strategy across all trials including postal questionnaires for follow-up. For instance, if the value of implementation was 10,000, the NIHR would need to invest £57.90 * 10,000 = £579,000 to generate similar additional benefits in terms of reduced participant attrition.

5.4. Results

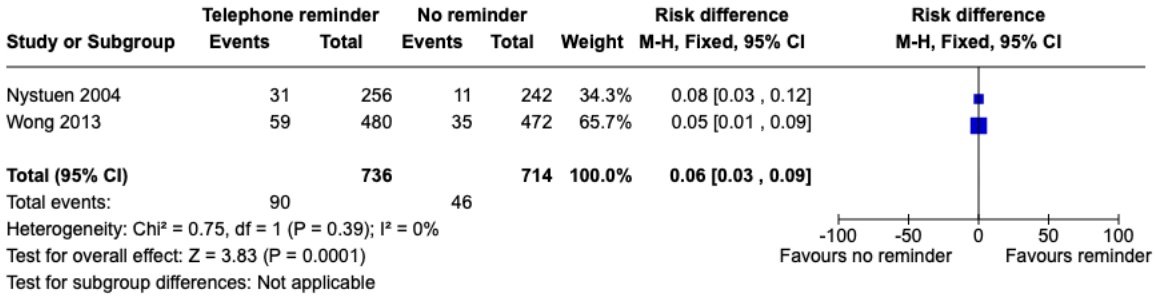
The figures presented in the *Results Section* (*Figure 5.3*, *Figure 5.4*, *Figure 5.5*, *Figure 5.6*, *Figure 5.7)*, and produced with RevMan, show the following: the name of the included studies; the number of participants retained in host trials by intervention group in each SWAT and across all SWATs; the total number of participants having taken part in SWATs by intervention group in each study and across all SWATs; the weighting of each SWAT according to the Mantel- Haenszel method; the reported RD and its 95% CI; a graphical representation of the

RD in each SWAT and across all SWATs; the cumulative sample size by intervention group across all SWATs; the reported Tau^2, Chi^2 and I^2 statistics to test for heterogeneity between the included SWATs; and the Z test of significance of the reported overall effect.

5.4.1. Value of Information (VoI) Analysis following meta-analysis of SWATs of telephone reminders as a recruitment strategy

Following the meta-analysis of the two SWATs undertaken for assessing the effectiveness of telephone reminders as a recruitment strategy versus no reminders, the reported RD does not present uncertainty. The recruitment rates in both trial arms were low, at approximately 12% and 6% for the telephone reminder and the no reminder arms respectively.

Figure 5.3: Risk Difference (RD) of Meta-Analysis 1



Following a random sampling of 5000 random combinations of RD and BaselineRecruitment to generate a joint uncertain distribution of the effectiveness of telephone reminders as a recruitment strategy, i.e. a distribution of the value of implementation, the probability that telephone reminders is (not) an effective strategy was 100% (0%). The value of additional research ($VoI_{REMINDER}$), that is the expected number of patients not recruited to all trials as a result of effect uncertainty, was 0.

Table 5.6: Summary of Findings from Meta-Analysis 1

Input	Value
$RecruitmentRate_{REMINDER}$	0.122
$RecruitmentRate_{NOREMINDER}$	0.064
Risk Difference	0.06
BaselineRecruitment	0.43
Sample Size of Participants	1,450
Range of Risk Difference (95% CI)	(0.03,0.09)
Range of BaselineRecruitment	(0.372, 0.487)
Trial Population	2,600,000
Number of simulations (k)	5000
P (ReminderEffective >0)	1
P (ReminderEffective ≤0)	0
$VoI_{REMINDER}$ (Value of additional research)	0
$ValueOfImplementation_{REMINDER}$ (Value of Implementation)	67,765

If telephone reminders were adopted immediately as a recruitment strategy in all trials following the dissemination of these SWATs, the value of implementation would be equal to 67,765, which means that an excess of 67,765 patients would be expected to be recruited into trials as a result of receiving telephone reminders at the recruitment stage. If a SWAT commissioner decided to invest in additional SWATs on other recruitment strategies, it would have to invest £54.61*67,765= £3,700,647, or in other words fund 740 SWATs, to generate similar additional benefits through additional research. By comparing the value of additional research and the value of implementation, as well as their cost implications from the economic perspective of a SWAT commissioner, it is almost certain that additional SWATs on telephone reminders would not be a cost-effective investment. By using a framework of VoI analysis in the case study of telephone reminders as a recruitment strategy, this chapter contradicts the original results from Trial Forge Guidance (see *Table 5.1*) that originally suggested that additional research on telephone reminders as a recruitment strategy be undertaken.

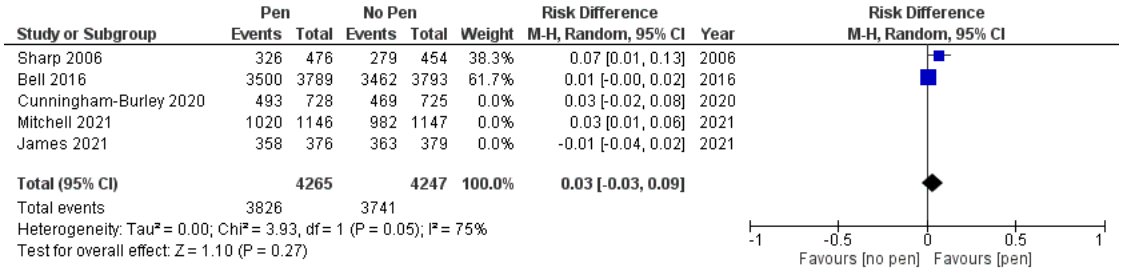
5.4.2. Value of Information (VoI) Analysis following meta-analysis of SWATs of pens as a retention strategy

5.4.2.1. Value of Information (VoI) Analysis following meta-analysis of the first two SWATs of pens as a retention strategy

Following the meta-analysis of the first two SWATs that were undertaken for assessing the effectiveness of pens as a retention strategy versus no pens, the reported RD presented uncertainty. Whereas the overall RD was greater than zero, thus indicating that pens could be an effective retention strategy by 3%, its lower bound was below zero. The retention rates in both trial arms were relatively high, at approximately 90% and 88% for the pen and the no pen arms respectively.

Following a random sampling of 5000 random combinations of RD_1 and $BaselineRetention_1$ to generate a joint uncertain distribution of the effectiveness of pens as a retention strategy, i.e. a distribution of the value of implementation, the probability that pens is (not) an effective strategy was approximately 75% (25%). The value of additional research (VoI_1), that is the expected number of participants withdrawing from all trials as a result of effect uncertainty, was 1,170.

Figure 5.4: Risk Difference (RD) of Meta-Analysis 2.1



If pens were adopted immediately as a retention strategy in all trials following the dissemination of the two SWATs, the value of implementation would be equal to 9,266, which means that an excess of 9,266 participants would be expected not to withdraw from trials as a result of receiving a pen during follow-up. If a SWAT commissioner decided to invest in additional SWATs on other retention strategies, it would have to invest £57.90*9,266= £536,501, or in other words fund 107 SWATs, to generate similar additional benefits through additional research. Nevertheless, it could still be the case that research councils invest in further SWATs of pens since the value of additional research had a figure of 1,170 trial participants, which would require the funding of 1,170/(£5000/£57.90)=1,170/86= 14 SWATs of pens as a retention strategy for effect uncertainty to be reduced significantly. By comparing the value of additional research and the value of implementation, as well as their cost

implications from the economic perspective of a SWAT commissioner, it is possible that additional SWATs on pens would not be a cost-effective investment.

Table 5.7: Summary of Findings from Meta-Analysis 2.1

Input	Value
$RetentionRate_{PEN,1}$	0.897
$RetentionRate_{NOPEN,1}$	0.881
$Risk\ Difference_{PEN,1}$	0.03
$BaselineRetention_1$	0.88
Sample Size of Participants	8,512
Range of Risk $Difference_1$ (95% CI)	(-0.03,0.09)
Range of $BaselineRetention_1$	(0.80,0.97)
Trial Population	347,621
Number of simulations (k)	5000
$P(PensEffective_{1,k} > 0)$	0.7473
$P(PensEffective_{1,k} \leq 0)$	0.2527
$VoI_{PEN,1}$(Value of additional research)	1,170
$ValueOfImplementation_{PEN,1}$ (Value of Implementation)	9,266

5.4.2.2. Value of Information (VoI) Analysis following meta-analysis of the first three SWATs of pens as a retention strategy

Following the meta-analysis of the first three SWATs undertaken for assessing the effectiveness of pens as a retention strategy versus no pens, the reported RD presented uncertainty, but less uncertainty than that observed in *Meta-Analysis 1*. Whereas the overall RD was greater than zero and returned the same value of 0.03, in line with the RD from *Meta-Analysis 1*, its lower bound was still below zero. The retention rates in both trial arms were relatively high, at approximately 87% and 85% for the pen and the no pen arms respectively.

Figure 5.5: Risk Difference (RD) of Meta-Analysis 2.2

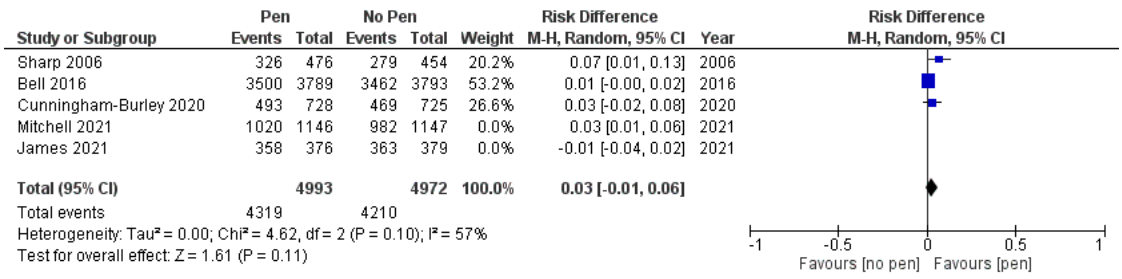


Table 5.8: Summary of Findings from Meta-Analysis 2.2

Input	Value
$RetentionRate_{PEN,2}$	0.865
$RetentionRate_{NOPEN,2}$	0.847
$Risk\ Difference_{PEN,2}$	0.03
$BaselineRetention_2$	0.88
Sample Size of Participants	9,965
Range of Risk Difference$_2$ (95% CI)	(-0.01, 0.06)
Range of $BaselineRetention_2$	(0.80,0.97)
Trial Population	347,621
Number of simulations	5000
$P(PensEffective_{2,k} > 0)$	0.8511
$P(PensEffective_{2,k} \leq 0)$	0.1489
$VoI_{PEN,2}$ (Value of additional research)	232
$ValueOfImplementation_{PEN,2}$ (Value of Implementation)	7,750

Following a random sampling of 5000 random combinations of RD_2 and $BaselineRetention_2$ to generate a joint uncertain distribution of the effectiveness of pens as a retention strategy, i.e. a distribution of the value of implementation, the probability that pens is (not) an effective strategy increased (decreased) significantly from 75% (25%) to 85% (15%). In addition, the value of additional research (VoI_2) fell considerably from 1,170 to 232. However, the value of implementation also fell from 9,266 to 7,750, but remained a significant figure. If a SWAT commissioner decided to invest in additional SWATs on other retention strategies, it would have to invest £57.90*7,750= £448,725, or in other words fund 90 SWATs, to generate similar additional benefits through additional research. In addition, the estimated VoI_2 seems to be low, as only 232/(£5000/£57.90)=232/86=3 further SWATs of pens as a retention strategy would be required for effect uncertainty to be reduced significantly. By comparing the value of additional research and the value of implementation, as well as their cost implications from the economic perspective of a SWAT commissioner, it is highly likely that additional SWATs on pens would not be a cost-effective investment.

5.4.2.3. Value of Information (VoI) Analysis following meta-analysis of the first four SWATs of pens as a retention strategy

Following the meta-analysis of the first four SWATs that were undertaken for assessing the effectiveness of pens as a retention strategy versus no pens, the reported RD slightly fell from

0.03 to 0.02 and presented uncertainty albeit reduced compared to *Meta-Analysis 2*. Whereas the overall RD was greater than zero, thus indicating that pens could be an effective retention strategy by 2%, its lower bound remained below zero. The retention rates in both trial arms were relatively high, at approximately 87% and 86% for the pen and the no pen arms respectively.

Following a random sampling of 5000 random combinations of RD_3 and $BaselineRetention_3$ to generate a joint uncertain distribution of the effectiveness of pens as a retention strategy, i.e. a distribution of the value of implementation, the probability that pens is (not) an effective strategy fell (increased) from 85% (15%) to 80% (20%). The value of additional research (VoI_3) increased slightly from 232 to 317. In addition, the value of implementation fell significantly from 7,750 to 4,764. If a SWAT commissioner decided to invest in additional SWATs on other retention strategies, it would have to invest £57.90*4,764= £275,836, or in other words fund 55 SWATs, to generate similar additional benefits through additional research. Nevertheless, it could still be the case that research councils invest in further SWATs of pens since the value of additional research had a figure of 317 trial participants, which means that 317/(£5000/£57.90)=317/86=4 further SWATs of pens as a retention strategy would be required for effect uncertainty to be reduced significantly. Despite the reduction in the value of implementation and the slight increase in the value of additional research, it is highly likely that additional SWATs on pens would still not be a cost-effective investment given the estimated figures and their cost implications from a SWAT commissioner's perspective.

Figure 5.6: Risk Difference (RD) of Meta-Analysis 2.3

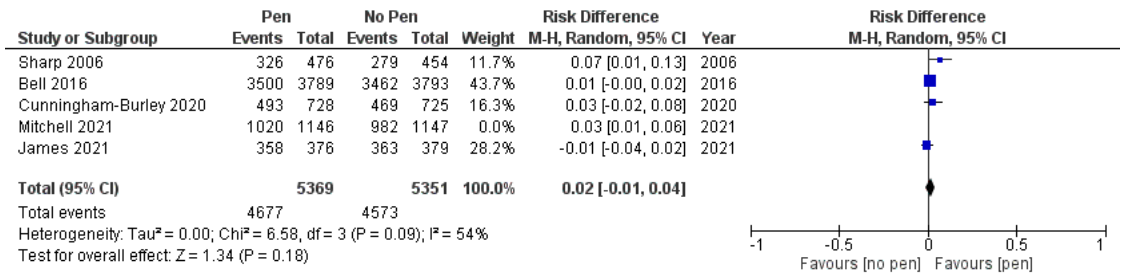


Table 5.9: Summary of Findings from Meta-Analysis 2.3

Input	Value
$RetentionRate_{PEN,3}$	0.871
$RetentionRate_{NOPEN,3}$	0.855
$Risk\ Difference_{PEN,3}$	0.02
$BaselineRetention_3$	0.88
Sample Size of Participants	10,720
Range of $Risk\ Difference_3$ (95% CI)	(-0.01,0.04)
Range of $BaselineRetention_3$	(0.80, 0.97)
Trial Population	347,621
Number of simulations	5000
$P(PensEffective_{3,k} > 0)$	0.7963
$P(PensEffective_{3,k} \leq 0)$	0.2037
$VoI_{PEN,3}$(Value of additional research)	317
$ValueOfImplementation_{PEN,3}$ (Value of Implementation)	4,764

5.4.2.4. Value of Information (VoI) Analysis following meta-analysis of all SWATs of pens as a retention strategy

Following the meta-analysis of all SWATs that were undertaken for assessing the effectiveness of pens as a retention strategy versus no pens, the reported RD did not present uncertainty. The overall RD remained stable at 0.02, in line with RD_3=0.02, thus indicating that pens could be an effective retention strategy, by 2%. In addition, the lower bound was no longer negative as it had been in the previous meta-analyses. The retention rates in both trial arms remained relatively high, at approximately 87% and 86% for the pen and the no pen arms correspondingly.

Figure 5.7: Risk Difference (RD) of Meta-Analysis 2.4

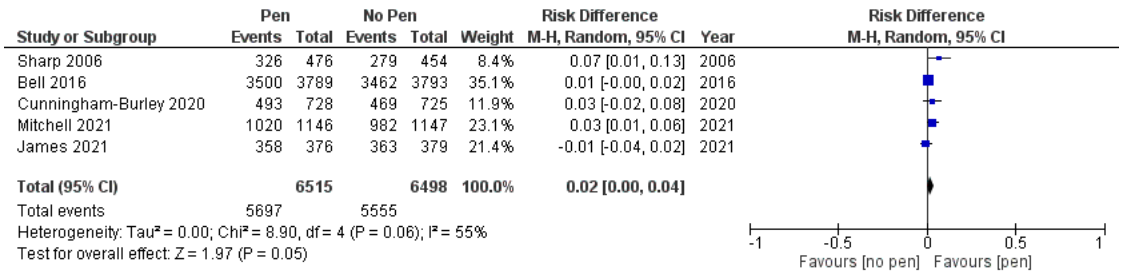


Table 5.10: Summary of Findings from Meta-Analysis 2.4

Input	Value
$RetentionRate_{PEN,4}$	0.874
$RetentionRate_{NOPEN,4}$	0.855
Risk $Difference_{PEN,4}$	0.02
$BaselineRetention_4$	0.88
Sample Size of Participants	13,013
Range of Risk $Difference_4$ (95% CI)	(0.00, 0.04)
Range of $BaselineRetention_4$	(0.80, 0.97)
Trial Population	347,621
Number of simulations	5000
$P(PensEffective_{4,k} > 0)$	0.9770
$P(PensEffective_{4,k} \leq 0)$	0.0230
$VoI_{PEN,4}$(Value of additional research)	0
$ValueOfImplementation_{PENS4}$ (Value of Implementation)	6,227

Following a random sampling of 5000 random combinations of RD_4 and $BaselineRetention_4$ to generate a joint uncertain distribution of the effectiveness of pens as a retention strategy, i.e. a distribution of the value of implementation, the probability that pens is an effective strategy increased from 85% to 98%. The value of additional research (VoI_4) fell considerably from 317 to zero. In addition, the value of implementation increased from 4,764 to 6,227. If a SWAT commissioner decided to invest in additional SWATs on other retention strategies, it would have to invest £57.90*6,227= £360,543, or in other words fund 72 SWATs, to generate similar additional benefits through additional research. Moreover, the reported value of additional research (VoI_4) is zero, which implies that any effect uncertainty with respect to the use of pens as a retention strategy has been reduced significantly, as a result of the evidence provided from all five SWATs. By comparing the value of additional research and the value of implementation, as well as their cost implications from the economic perspective of a SWAT commissioner, it is almost certain that additional SWATs on pens would not be a cost-effective investment. By using a framework of VoI analysis in the case study of pens as a retention strategy, this chapter contradicts the original results from the Cochrane review (Gillies et al., 2021) and Trial Forge Guidance (see *Table 5.2*) that originally suggested that additional research on pens as a retention strategy be undertaken.

5.5. Discussion

5.5.1. Summary of findings

This chapter introduced a quantitative framework, in addition to the already established qualitative one of Trial Forge Guidance 2 (Treweek et al., 2020), to determine whether an additional SWAT of a recruitment or retention strategy would be a cost-effective investment. Such framework, based upon a modified methodology of VoI analysis by Claxton et al. (2015a) and McKenna et al. (2016), consists of the following components: a standard meta-analysis of a recruitment or retention strategy, given the available evidence, to obtain a statistical measure of effect; an estimate of the population of trial participants; a range of the baseline incidence of patient recruitment or participant retention across all trials from the literature; a random sampling of the product of the baseline incidence of patient recruitment or participant retention and a statistical measure of effect; a generated probability of a strategy being (or not being) effective; a probability-weighted effectiveness of a strategy, expressed in terms of the trial population; the value of implementation of a strategy, expressed in terms of the trial population; the value of additional research, expressed in terms of the trial population. Subsequently, the value of implementation and the value of additional research could be compared through the cost of funding an additional SWAT, e.g. £5000 under the PROMETHEUS programme, and the reported cost per additional patient recruited or per additional participant retained from previous SWATs respectively. In other words, the key aim of this methodology is to compare the marginal cost of instantly and universally adopting a recruitment or retention strategy across UK trials with the marginal cost of reducing the effect uncertainty of a recruitment or retention strategy via additional SWATs. The framework presented in this chapter is feasible and applicable for all recruitment and retention strategies associated with SWATs by modifying the aforementioned components for each strategy accordingly.

In the chapter's case studies of telephone reminders as a recruitment strategy and of pens as a retention strategy, there are two key findings. First, whereas Trial Forge Guidance 2 suggests that further SWATs of telephone reminders and pens be undertaken to obtain more precise estimates of its effectiveness as a retention strategy (*Table 5.1*, *Table 5.2*), the modified VoI analysis presented in this chapter concludes the opposite. The estimated value of implementation indicates that 67,765 additional invited patients and 6,227 additional trial participants would be expected to be retained in trials if telephone reminders and pens were universally and instantly adopted as recruitment and retention strategies across UK trials

respectively. The value of additional research, on the other hand, signals that no trial participants across the UK would be expected to withdraw from trials as a result of the existing effect uncertainty of telephone reminders as a recruitment strategy and of pens as a retention strategy.

As the cost per additional invited patient recruited is estimated to be £54.61 (Kakumanu et al., 2019) and the maximum funding the PROMETHEUS programme provided for a new SWAT was up to £5,000, if a SWAT commissioner decided not to adopt telephone reminders as a recruitment strategy across trials and instead invest in additional SWATs on other recruitment strategies, it would have to invest in approximately 740 SWATs to observe similar expected incremental recruitment rates. Correspondingly, as the cost per additional participant retained through envelopes is £57.90 (Murphy et al., 2022b) and the maximum funding the PROMETHEUS programme provided for a new SWAT was up to £5,000, if a SWAT commissioner decided not to adopt pens as a retention strategy across trials and instead invest in additional SWATs on other retention strategies, it would have to invest in approximately 72 SWATs to observe similar expected incremental retention rates. Assuming that the NIHR, which is currently the leading SWAT commissioner, provided a maximum budget of £30,000 to each SWAT (National Institute for Health & Care Research, 2023), the aforementioned figures would correspond to approximately 123 further SWATs for telephone reminders as a recruitment strategy and 12 further SWATs for pens as a retention strategy. Also, the value of additional research figure implies that effect uncertainty has been reduced significantly after conducting two SWATs on telephone reminders as a recruitment strategy and of five SWATs on pens as a retention strategy. Therefore, it is almost certain that additional SWATs on telephone reminders and pens would not be a worthwhile investment; instead, funding SWATs on alternative retention strategies, whose certainty of evidence is lower according to the findings of the Cochrane recruitment review or retention review (Treweek et al., 2018b) (Gillies et al., 2021), could yield a higher return on investment. In other words, the marginal costs of instantly and universally adopting telephone reminders and pens as recruitment and retention strategies, respectively, across UK trials are significantly less than the marginal costs of reducing their effect uncertainty via additional SWATs.

Second, by following the same modified VoI analysis, funding the latest three SWATs of pens was not necessarily a cost-effective decision, from a SWAT commissioner's perspective. In other words, the broader use of pens as a retention strategy may have been justifiable earlier given the estimates of the value of additional research and the value of implementation associated with the evidence from the first two SWATs only; instead, a SWAT commissioner could have allocated resources to the research of alternative retention strategies for which more effect uncertainty exists. Had pens been universally adopted as a retention strategy straight after the first two SWATs that evaluated their effectiveness, an excess of 9,266 trial participants would have been expected not to withdraw from trials according to the estimated value of implementation. Therefore, if a SWAT commissioner decided not to adopt pens as a retention strategy across trials and instead invest in additional SWATs on other retention strategies, it would have to invest in approximately £536,601/£5,000 = 107 SWATs to observe similar expected incremental retention rates. On the other hand, the value of additional research figure stood at 1,170, implying that 14 further SWATs of pens would be needed to reduce uncertainty to an acceptable level. Indeed, the subsequent VoI analyses related to *Meta-Analysis 2.2*, *Meta-Analysis 2.3*, and *Meta-Analysis 2.4* confirm this suggestion; overall, the value of implementation remained at high levels, ranging between 4,764 and 9,266 trial participants, whereas the value of additional research fell almost gradually from 1,170 to zero trial participants.

5.5.2. Strengths and limitations of the study

This chapter recommends that such a modified VoI analysis framework be adopted alongside Trial Forge Guidance 2 for all recruitment and retention strategies, so that trial teams and SWAT commissioners make efficient, evidence-based decisions. Whereas this chapter's framework contradicts the recommendations of Trial Forge Guidance 2 for both telephone reminders as a recruitment strategy and pens as a retention strategy, both approaches have valuable conclusions, advantages and disadvantages.

In the case of telephone reminders, Trial Forge Guidance 2 reasonably highlights the small sample size involved with the cumulative evidence (n=1450), the low number of included SWATs (two), and the moderate GRADE certainty of evidence for randomised trials that have higher baseline recruitment rates exceeding 10%. However, the framework presented in this chapter also supports that further SWATs on telephone reminders as a recruitment strategy may not be a cost-effective investment by a SWAT commissioner, since the value of additional

research is 0 and the value of implementation is 67,765, meaning that approximately additional 740 SWATs would need to be funded to observe similar incremental recruitment rates through research on alternative recruitment strategies. Therefore, whereas the former guidance is more specific and captures with qualitative precision the improvements in future research that need to be made, the latter framework also considers the financial implications of doing so, given the accumulated evidence, for the trial community and SWAT commissioners, something which the original Trial Forge Guidance 2 fails to achieve. Nevertheless, the latter framework is also subject to drawbacks, such as strictly yielding value of additional research estimates of 0 when the lower bound of the RD confidence interval is equal to or greater than 0, irrespective of the sample size included in a meta-analysis, and not capturing sub-groups effectively, due to the fact that data of subgroup-population estimates in each trial, disease, or healthcare setting are hard to collect. Even if the value of additional research is zero, it could still be possible that additional research be worthwhile. Whereas it may not be worthwhile doing more SWATs, the conduct of an additional SWAT could reduce further the width of the confidence intervals and thus produce an improved estimate of the effect difference and lead to more confidence in the results. It is evident that the drawbacks presented in Trial Forge Guidance 2 are addressed by the quantitative design of the chapter's recommended VoI analysis framework, while the drawbacks of the latter are addressed by the qualitative design of the former. The same conclusions can be drawn following the recommendations of Trial Forge Guidance 2 and the VoI analysis framework on pens as a retention strategy.

5.5.3. Conclusion and recommendations for future research

The VoI analysis framework presented in this chapter is encouraged as a guide for decision makers, including SWAT commissioners and trial teams, to maximise the return on investment in future SWATs. It can be feasibly applied to all recruitment and retention strategies to evaluate whether further research on each strategy should be funded, given budget constraints and underlying uncertainty. This approach could accelerate the acquisition of sufficient evidence regarding the most effective recruitment and retention strategies. Thus, research commissioners would be able to fund studies whose findings would have the highest research impact, while ensuring that their allocated budget remains sustainable and resource waste is minimised.

However, given that there are other elements that are needed to be considered in the recruitment and retention of participants, such as those included in Trial Forge Guidance 2 (e.g. population

characteristics, clinical settings where evidence on recruitment/retention strategies arises), it is essential to reach a balance qualitative quality, financial constraints and underlying statistical uncertainty. Therefore, it is crucial that the VoI analysis framework be used in combination with Trial Forge Guidance 2. This combination can lead to more informed, feasible and balanced decisions on SWATs of recruitment and retention strategies and determine which of these strategies should be adopted in the recruitment and retention processes of future trials.

Chapter 6: Concluding remarks

Whilst a RCT is the best method for establishing effectiveness, ironically many of the procedures needed for ensuring a rigorous design are underpinned with weak sources of evidence. 'Custom and practice' prevail for many of the recruitment and retention approaches. The best evidence around retention and recruitment methods is derived from SWATs. Whilst SWATs have been performed for the last 20 or 30 years, they tend to be rare with the relevant Cochrane reviews lamenting at their scarcity (Gillies et al., 2021, Treweek et al., 2018b). However, there has been a sea change in undertaking SWATs with the National Institute for Health & Care Research (NIHR) encouraging SWATs by setting aside a trial related budget solely for the undertaking of SWATs (National Institute for Health & Care Research, 2023). Internationally, the Irish Health Research Board (HRB) and the Canadian Accelerating Clinical Trials (ACT) Consortium fund SWAT research proposals (HRB-Trials Methodology Research Network, 2023, Accelerating Clinical Trials Consortium, 2023).

This thesis makes an important contribution to this new SWAT landscape. It is the first piece of work that systematically uses an economic framework to assess the costs and benefits of improving retention and recruitment. It showed how poor recruitment to the RECOVERY trial led to significant loss of health and utility for Covid-19 patients, and how costly participant attrition can be for trial teams and funders. It also uses economic methods to help improve the prioritisation of SWAT conduct, while highlighting the need for economic evaluations alongside future SWATs of recruitment and retention strategies. In the following chapter the key findings of this thesis are summarised and justified. In addition, the impact and the dissemination of the PhD chapter results are presented. Finally, the direction for future research is discussed.

6.1. Research Question 1: What is the economic impact of poor patient recruitment into RCTs?

Chapter 2 demonstrated, in the case of the RECOVERY trial, that slow patient recruitment prevented the earlier dissemination of a more cost-effective, available treatment for hospitalised COVID-19 patients, i.e. dexamethasone. If recruiting or redeploying two research nurses to each involved hospital increased recruitment rates from 15% to 50%, only £10,641 would need to be invested by a decision maker to generate an additional quality-adjusted life year (QALY). Faster recruitment to the RECOVERY trial could have generated an incremental

net benefit of £13,955,476 (95% CI: £12,457,049, £15,453,904) thus highlighting the magnitude of the foregone population health benefits due to not having updated clinical practice earlier, i.e. early April 2020 instead of mid-June 2020, through faster recruitment. In addition, a study from Knowlson and Torgerson (2020) demonstrated that faster recruitment to the RECOVERY trial would have saved at least 2,600 lives. Note these results are related to patients hospitalised in the UK, thus the reported economic impact of slow recruitment in the RECOVERY trial is underestimated.

Aside from the statistical implications of slow patient recruitment for trials teams, including increased likelihood of Type II error, there are also significant consequences for national healthcare systems, explored for the first time in the study presented in *Chapter 2*. I chose the RECOVERY trial as a case study, in order to highlight how crucial faster patient recruitment would have been for the NHS, while it was facing an increasing number of patient admissions tested positive with Sars-Cov-2, a virus of which the disease had no available evidence-based medical interventions to be treated in the early stages. Given the aforementioned results, I highlighted the importance of identifying strategies that could improve patient recruitment in randomised trials, while emphasising that SWATs would the best study design for achieving this goal. In this chapter, I assumed that recruiting two research nurses was an effective recruitment strategy. However, this is not an evidence-based assumption, despite the fact that it could be realistic for a research nurse to recruit up to four patients every day, meaning it would take only 10.5 days (including five working days in one week and two days off work, followed by another 3.5 working days in the following week) to have reached the desirable recruitment figure of 11,303 patients in the RECOVERY trial. Thus, it is crucial that SWATs identify evidence-based studies that could accelerate patient recruitment in randomised trials.

The methodology and findings of *Chapter 2* were presented orally, and as a poster, at the 6th International Clinical Trials Methodology Conference (ICTMC 2022) in Harrogate, United Kingdom (abstract ID: PS1D-LT11 (Sydes et al., 2023)). I have also written a manuscript related to *Chapter 2* which has been submitted for peer-review to a journal, as of May 2024.

6.2. Research Question 2: What is the economic impact of participant attrition from RCTs?

Chapter 3 demonstrated, in the case of the OTIS trial, that participant attrition generated significant financial costs to the trial team and the study's funder, i.e. of approximately

£10,235. This figure considers the protocol-driven costs, including administration, print and shipping costs, related to recruited participants who were eventually lost to follow-up. In addition, it is possible that NIHR-funded studies may experience an aggregate annual loss of at least £960,000 if the foregone share of the total funding awarded, due to participant loss to follow-up, was 1.42% across all randomised trials funded by the NIHR. Attrition in the OTIS trial generated a resource misallocation of approximately £10,235, an amount that could have been invested by the NIHR in further clinical research instead.

Aside from the statistical implications of participant attrition for trials teams, including attrition bias resulting in misleading results, there are also economic consequences for trial teams and funders. This study is the first research work to identify the monetary costs of attrition from the economic perspective of a trial team in a randomised study. The detailed cost analysis related to participant loss to follow-up can also be used as a guidance for trial teams to estimate the marginal and aggregate cost of participant loss to follow-up in their randomised studies, even before their conduct. Nevertheless, the assumptions and the economic perspective for the costing analysis presented were related to the OTIS trial, thus they may be subject to adjustments should this framework be used for other randomised trials.

I chose the OTIS trial as a case study, in order to highlight that despite its low overall attrition rate, i.e. of 9.8%, the loss to follow-up of 104 recruited participants generated significant financial losses to the trial team, and to the study's funder. Given the aforementioned results, I highlighted the importance of identifying strategies that could improve participant retention in randomised trials, while emphasising that SWATs would the best study design for achieving this goal.

A manuscript related to *Chapter 3* has been submitted for peer-review to an academic journal, as of May 2024.

6.3. Research Question 3: What is the cost-effectiveness of existing recruitment and retention strategies?

Given the significant costs of poor recruitment and attrition to trial teams, funders and healthcare systems, it is strongly recommended that SWATs be promoted as a study design for identifying effective and cost-effective recruitment and retention strategies. Following the establishment of SWATs as the main research design for identifying such strategies, a recent

Cochrane review identified open trial designs to be a more effective recruitment strategy, compared to placebo-controlled designs; it remains uncertain whether the remaining strategies can improve recruitment to randomised trials with certainty (Treweek et al., 2018b). Another Cochrane review did not identify any retention strategy improving retention with high certainty (Gillies et al., 2021).

Whereas these reviews explored the effectiveness of existing recruitment and retention interventions, they did not evaluate their costs and hence their cost-effectiveness. *Chapter 4* presents a systematic review, the first one to critically appraise the evidence surrounding the cost-effectiveness of recruitment and retention strategies, whose evidence arises from SWATs.

Recruitment strategies which could be cost-effective include financial incentives, trial-branded pens, telephone reminders and pre-notification leaflets. Adding a branded pen to a trial invitation pack or a telephone reminder versus an SMS reminder could be cost-effective strategies, with their incremental cost-effectiveness ratios (ICERs) being low, at $21.41 and $23.37 respectively. However, as their lower bound odds ratios (ORs) are less than 1, their cost-effectiveness is statistically insignificant. In addition, their GRADE certainty of evidence is moderate, implying these strategies would benefit from further SWATs to determine their cost-effectiveness, with less uncertainty. Providing financial incentives may be an effective yet costly strategy, with $476.57 required to recruit an additional patient.

Retention strategies which could be cost-effective include vouchers and trial-branded pens. Providing vouchers of up to £10 during follow-up could be a cost-effective retention strategy, with an estimated ICER of $15.89; it costs only $15.89 for an additional participant to be retained in a host trial. Providing a trial-branded pen may also be a cost-effective strategy, with an ICER of $6.98, yet not statistically significant since its lower bound OR=1 (hence its lower bound effectiveness is zero). Also, providing a trial-branded pen before recruitment may be a cost-effective strategy, with an ICER of $0.40, which is statistically significant. However, the GRADE certainty of evidence for both strategies is moderate, meaning that additional SWATs of these strategies would be beneficial for making more certain inferences about their cost-effectiveness.

The thesis review, also published as a peer-reviewed study (Gkekas et al., 2023), has also introduced recommendations on how costs should be reported and economic evaluations undertaken alongside future SWATs. Costs should be reported transparently, expressed in unit

terms and stratified into different types of direct (print, administrative and print) and indirect costs, including the costs of obtaining outcome data. The study also recommends that a cost-effectiveness threshold be set for recruitment and retention strategies in future research, as such a threshold is not currently in place. Finally, it suggests that "*economic evaluations be carried out alongside all future SWATs*" (Gkekas et al., 2023: p.94). These findings also justify the decision for carrying out a flexible review that also considered studies with a high Cochrane risk of bias, or unpublished/non-peer-reviewed studies, given the absence of cost reporting or economic evaluations alongside most published SWATs.

The results and conclusions from the systematic review were also presented as a poster at the 6th International Clinical Trials Methodology Conference (ICTMC 2022) in Harrogate, United Kingdom (abstract ID: P-253 (Sydes et al., 2023)). Furthermore, an update to the Cochrane review of strategies for improving recruitment in randomised trials is currently being undertaken (Treweek et al., 2018b), in which I participate as a reviewer and cost and cost-effectiveness outcome data will be also reported if available; the current version of the Cochrane review did not collect cost and cost-effectiveness outcome data from SWATs related to recruitment strategies (Treweek et al., 2018b). This thesis informed the decision to assess costs in the Cochrane review, and hence the methodology and data extraction form of the cost-effectiveness review (Gkekas et al., 2023) are being considered to report these outcomes. The Cochrane review update of strategies to improve recruitment in randomised trials is expected to be published in 2024. It will also include novel evidence surrounding the cost-effectiveness of recruitment strategies that has been published after March 2021, the date on which the search of records for the cost-effectiveness review ended (Gkekas et al., 2023).

6.4. Research Question 4: How could VoI analyses related to retention or recruitment interventions inform decision makers on whether additional SWATs are needed for improving the evidence on the (cost-) effectiveness of such interventions?

Whereas the Cochrane reviews (Treweek et al., 2018b; Gillies et al., 2021) and the cost-effectiveness review presented in *Chapter 4* (Gkekas et al., 2023) included results for potential recruitment and retention strategies whose evidence arose primarily from SWATs, the uncertainty regarding their generalisability in improving trial efficiency originated from the absence of replications of SWATs surrounding their effectiveness and cost-effectiveness.

Thus, the authors of the Cochrane review recommended that "*the research community...establish a process for prioritising which… interventions are most in need of evaluation*" (Treweek et al., 2018b: p.22).

To address this issue, Trial Forge Guidance 2 has been created to assist trial teams and researchers in deciding whether an additional SWAT on a given strategy should be undertaken (Treweek et al., 2020). The review presented in *Chapter 4* also applied this guidance to explore whether the existing recruitment and retention strategies, for which cost-effectiveness data were available, should be explored further. The study concluded that "*future SWATs should replicate existing recruitment and retention strategies, rather than evaluate novel ones*" (Gkekas et al., 2023: p.94), following the application of Trial Forge Guidance 2 criteria. However, such a guidance does not provide a financial threshold to determine whether the existing level of evidence is sufficient given the budget constraints a SWAT commissioner may face.

To address this challenge, *Chapter 5* introduced and applied a VoI analysis framework to one recruitment and one retention strategy, telephone reminders and pens respectively (Claxton et al., 2015a, McKenna et al., 2016). The purpose of this approach was to compare the marginal cost of instantly and universally adopting a recruitment or retention strategy across UK trials with the marginal cost of reducing its effect uncertainty via additional SWATs, in order for SWAT commissioners to maximise their returns on investment on future SWATs. Whereas Trial Forge Guidance 2 suggests that further SWATs of telephone reminders and pens be undertaken to obtain more precise estimates of its effectiveness as a retention strategy, the modified VoI analysis presented in Chapter 5 of this thesis comes to opposite conclusions. The estimated value of implementation shows that 67,765 additional invited patients would be expected to be recruited in trials, if telephone reminders were universally and instantly adopted as recruitment strategy across UK trials. This implies that, if reminders were not adopted universally as a recruitment strategy, the SWAT commissioner would need to invest in additional 740 SWATs to generate similar incremental recruitment rates. Moreover, 6,227 additional trial participants would be expected to be retained in trials, if pens were universally and instantly adopted as a retention strategy across UK trials. This implies that, if pens were not adopted universally as a retention strategy, the SWAT commissioner would need to invest in additional 72 SWATs to generate similar incremental retention rates. On the other hand, the value of additional research, given its methodological structure and assumptions, implies that

no patients taking part in randomised trials across the UK would be expected to withdraw from trials as a result of the existing effect uncertainty of telephone reminders as a recruitment strategy and of pens as a retention strategy.

Nevertheless, such a guidance is also subject to limitations. For instance, the value of additional research is zero when the lower bound of the risk difference (RD) confidence interval is greater than zero, independently from the sample size and the number of studies included in the meta-analysis of a given recruitment or retention strategy. Whereas it may not be worthwhile doing more SWATs, the conduct of an additional SWAT could reduce further the width of the confidence intervals and thus lead to more confidence in the results, something which the VoI analysis presented in *Chapter 5* cannot capture. On the other hand, the framework presented in *Chapter 5* is the first attempt to consider the financial implications of undertaking another SWAT of an existing strategy, from the economic perspective of a SWAT commissioner, something which Trial Forge Guidance 2 fails to do. Therefore, by comparing the results from the VoI analysis and Trial Forge Guidance 2, it strongly recommends that such a framework be used as a guide for decision makers and included in an updated version of Trial Forge Guidance 2, so that both qualitative and economic aspects be considered when prioritising future SWATs on certain recruitment and retention strategies.

The methodology and results of this study were presented at an Economic Evaluation seminar at the Centre for Health Economics (CHE) of the University of York in December 2022.This seminar was also attended by the authors of the studies that introduced the VoI analysis framework I adopted for evaluating the cost-effectiveness of undertaking further SWATs (Claxton et al., 2015a, McKenna et al., 2016). Their feedback on my research work enabled me to transition the VoI analysis framework in the context of clinical studies to that of SWATs.

The conclusions of this chapter have motivated the submission of a grant application called "*Trial Forge Guidance 2 Extension: Reducing research waste by considering the cost-effectiveness of undertaking further SWATs on interventions*" to the Irish Health Research Board- Trials Methodology Research Network (HRB-TMRN), of which I was a co-applicant. The grant application was approved for funding, for an amount of €10,000. This project aims: 1) to apply VoI analysis to a range of recruitment and retention strategies; 2) to develop a VoI tool for trial teams conducting SWATs and; 3) to update Trial Forge Guidance 2. The project findings are also expected to be presented at the 7th International Clinical Trials Methodology Conference (ICTMC 2024) in Edinburgh, United Kingdom.

6.5. Direction for future research on improving trial efficiency

Slow recruitment of patients to and attrition of participants in randomised trials may generate noticeable costs for trial teams, funders and healthcare systems. Given that SWATs is the principal study design to evaluate existing or novel recruitment and retention strategies to address these challenges, direction for future research on trial efficiency should be focused on the conduct of SWATs.

Whereas SWATs are becoming increasingly popular, and major funders such as the NIHR in the UK (National Institute for Health & Care Research, 2023), the HRB in Ireland (TMRN 2023) and the ACT in Cananda (ACT 2023) provide special funding for the conduct of SWATs, few SWATs have considered the costs and the cost-effectiveness of the strategies under evaluation. The cost-effectiveness review in *Chapter 4* identified only 17 out of 139 included in the recruitment and retention reviews (Treweek et al., 2018b) (Gillies et al., 2021). Thus, it is highly recommended that future SWATs include economic evaluations, by breaking down the costs related to recruitment and retention strategies into direct (print, design and shipping) and indirect costs. Health economics should be more intensely incorporated within the research on SWATs, given that recommended recruitment and retention strategies should not only be effective, but also cost-effective given the costs they may generate to trial teams and funders. It could also be worthwhile for trial teams to estimate the financial costs they may incur due to expected participant losses to follow-up when they submit a grant application to a research funder. Such an estimation could be in line with the costing analysis I presented in *Chapter 3*, which explored the financial costs of attrition in the OTIS trial from the economic perspective of a trial team.

In addition, given that the NIHR is the main SWAT commissioner in the UK, providing grants of up to £30,000 for eligible SWATs embedded within randomised trials, it is important that additional research on recruitment and retention strategies be prioritised according to the criteria of both Trial Forge Guidance 2 (Treweek et al., 2020) and the VoI analysis framework introduced in *Chapter 5*, which also considers the budget implications of additional research for SWAT commissioners. The framework introduced in the thesis will be used for a funded project, of which I will be the principal researcher. It will develop a VoI tool for as many recruitment and retention strategies as possible, so that trial teams and SWAT commissioners prioritise with more confidence which SWATs of which interventions should be undertaken and funded. Trial efficiency can be achieved faster and more efficiently through evidence-

based investment decisions on future SWAT research of recruitment and retention strategies, which could be achieved through VoI analysis, in addition to the existing Trial Forge Guidance 2.

As randomised trials are irrefutably the gold-standard design for health and clinical research, leading to the provision of more effective and cost-effective interventions, diagnostics, treatments and healthcare services for patients and the population overall, they deserve to be more efficient and to minimise the dire consequences of slow patient recruitment and participant attrition - for the benefit of patients, trial teams, national healthcare systems and society as a whole. The most suitable design to achieve this goal is SWATs, capable of producing credible evidence surrounding the effectiveness and the cost-effectiveness of recruitment and retention strategies under evaluation. For SWATs to maximise their impact, health economic methods and approaches should be integrated more closely in future research. This implies that trial teams should undertake economic evaluations alongside future SWATs, as well as determine which SWATs of which interventions should be prioritised for research for trial efficiency to be improved as early and cost-effectively as possible.

References

ACCELERATING CLINICAL TRIALS CONSORTIUM. 2023. *RFA Funding Successes* [Online]. Available: https://www.act-aec.ca/successes [Accessed 23/09/2023].

ADAMS, M., CAFFREY, L. & MCKEVITT, C. 2015. Barriers and opportunities for enhancing patient recruitment and retention in clinical research: findings from an interview study in an NHS academic health science centre. *Health Res Policy Syst,* 13**,** 8.

ÁGUAS, R., MAHDI, A., SHRETTA, R., HORBY, P., LANDRAY, M. & WHITE, L. 2021. Potential health and economic impacts of dexamethasone treatment for patients with COVID-19. *Nature Communications,* 12**,** 915.

AKL, E. A., BRIEL, M., YOU, J. J., SUN, X., JOHNSTON, B. C., BUSSE, J. W., MULLA, S., LAMONTAGNE, F., BASSLER, D., VERA, C., ALSHURAFA, M., KATSIOS, C. M., ZHOU, Q., CUKIERMAN-YAFFE, T., GANGJI, A., MILLS, E. J., WALTER, S. D., COOK, D. J., SCHÜNEMANN, H. J., ALTMAN, D. G. & GUYATT, G. H. 2012. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. *BMJ : British Medical Journal,* 344**,** e2809.

ARUNDEL, C., JEFFERSON, L., BAILEY, M., COCKAYNE, S., HICKS, K., LOUGHREY, L., RODGERS, S. & TORGERSON, D. J. 2017. A randomized, embedded trial of pre-notification of trial participation did not increase recruitment rates to a falls prevention trial. *J Eval Clin Pract,* 23**,** 73-78.

ASHBY, R., TURNER, G., CROSS, B., MITCHELL, N. & TORGERSON, D. 2011. A randomized trial of electronic reminders showed a reduction in the time to respond to postal questionnaires. *J Clin Epidemiol,* 64**,** 208-12.

AVENELL, A., GRANT, A. M., MCGEE, M., MCPHERSON, G., CAMPBELL, M. K. & MCGEE, M. A. 2004. The effects of an open design on trial participant recruitment, compliance and retention--a randomized controlled trial comparison with a blinded, placebo-controlled design. *Clin Trials,* 1**,** 490-8.

BELL, K., CLARK, L., FAIRHURST, C., MITCHELL, N., LENAGHAN, E., BLACKLOCK, J., CUSHNAGHAN, J., COOPER, C., GITTOES, N., O'NEILL, T. W., SHEPSTONE, L. & TORGERSON, D. J. 2016. Enclosing a pen reduced time to response to questionnaire mailings. *J Clin Epidemiol,* 74**,** 144-50.

BLACK, N. 1996. Why we need observational studies to evaluate the effectiveness of health care. *Bmj,* 312**,** 1215-8.

BOWER, P., BRUETON, V., GAMBLE, C., TREWEEK, S., SMITH, C. T., YOUNG, B. & WILLIAMSON, P. 2014. Interventions to improve recruitment and retention in clinical trials: a survey and workshop to assess current practice and future priorities. *Trials,* 15**,** 399.

BRACKEN, K., HAGUE, W., KEECH, A., CONWAY, A., HANDELSMAN, D. J., GROSSMANN, M., JESUDASON, D., STUCKEY, B., YEAP, B. B., INDER, W., ALLAN, C., MCLACHLAN, R., ROBLEDO, K. P. & WITTERT, G. 2019. Recruitment of men to a multi-centre diabetes prevention trial: an evaluation of traditional and online promotional strategies. *Trials,* 20**,** 366.

BRADSHAW, L. E., MONTGOMERY, A. A., WILLIAMS, H. C., CHALMERS, J. R. & HAINES, R. H. 2020. Two-by-two factorial randomised study within a trial (SWAT) to evaluate strategies for follow-up in a randomised prevention trial. *Trials,* 21**,** 529.

BRIGGS, A., CLAXTON, K. & SCULPHER, M. 2006. *Decision Modelling for Health Economic Evaluation,* New York, United States of America, Oxford University Press.

BRUETON, V. 2022. Retaining trial participants: an individualised approach is needed. *BMJ,* 376**,** o115.

BRUETON, V. C., TIERNEY, J. F., STENNING, S., MEREDITH, S., HARDING, S., NAZARETH, I. & RAIT, G. 2014. Strategies to improve retention in randomised trials: a Cochrane systematic review and meta-analysis. *BMJ Open,* 4**,** e003821.

BRUNSDON, D., BIESTY, L., BROCKLEHURST, P., BRUETON, V., DEVANE, D., ELLIOTT, J., GALVIN, S., GAMBLE, C., GARDNER, H., HEALY, P., HOOD, K., JORDAN, J., LANZ, D., MAESO, B., ROBERTS, A., SKENE, I., SOULSBY, I., STEWART, D., TORGERSON, D., TREWEEK, S., WHITING, C., WREN, S., WORRALL, A. & GILLIES, K. 2019. What are the most important unanswered research questions in trial retention? A James Lind Alliance Priority Setting Partnership: the PRioRiTy II (Prioritising Retention in Randomised Trials) study. *Trials,* 20**,** 593.

CALLEN, T. 2012. Purchasing power parity: weights matter. *Finance & Development*.

CARTA, A. & CONVERSANO, C. 2021. Cost utility analysis of Remdesivir and Dexamethasone treatment for hospitalised COVID-19 patients - a hypothetical study. *BMC Health Serv Res,* 21**,** 986.

CHINN, S. 2000. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med,* 19**,** 3127-31.

CLARK, L., ARUNDEL, C., COLEMAN, E., DOHERTY, L., PARKER, A., HEWITT, C., BEARD, D., BOWER, P., BROCKLEHURST, P., COOPER, C., CULLIFORD, L., DEVANE, D., EMSLEY, R., ELDRIDGE, S., GALVIN, S., GILLIES, K., MONTGOMERY, A., SUTTON, C., TREWEEK, S. & TORGERSON, D. 2022. The PROMoting the USE of SWATs (PROMETHEUS) programme: Lessons learnt and future developments for SWATs. *Research Methods in Medicine & Health Sciences,* 0**,** 26320843221089632.

CLARK, L., RONALDSON, S., DYSON, L., HEWITT, C., TORGERSON, D. & ADAMSON, J. 2015. Electronic prompts significantly increase response rates to postal questionnaires: a randomized trial within a randomized trial and meta-analysis. *J Clin Epidemiol,* 68**,** 1446-50.

CLARKE, M., SAVAGE, G., MAGUIRE, L. & MCANENEY, H. 2015. The SWAT (study within a trial) programme; embedding trials to improve the methodological design and conduct of future research. *Trials,* 16**,** P209.

CLAXTON, K., GRIFFIN, S., KOFFIJBERG, H. & MCKENNA, C. 2015a. How to estimate the health benefits of additional research and changing clinical practice. *Bmj,* 351**,** h5987.

CLAXTON, K., MARTIN, S., SOARES, M., RICE, N., SPACKMAN, E., HINDE, S., DEVLIN, N., SMITH, P. C. & SCULPHER, M. 2015b. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol Assess,* 19**,** 1-503, v-vi.

CLAXTON, K., PALMER, S., LONGWORTH, L., BOJKE, L., GRIFFIN, S., SOARES, M., SPACKMAN, E. & ROTHERY, C. 2016. A Comprehensive Algorithm for Approval of Health Technologies With, Without, or Only in Research: The Key Principles for Informing Coverage Decisions. *Value Health,* 19**,** 885-891.

COCHRANE, A., WELCH, C., FAIRHURST, C., COCKAYNE, S. & TORGERSON, D. J. 2020. An evaluation of a personalised text message reminder compared to a standard text message on postal questionnaire response rates: an embedded randomised controlled trial. *F1000Res,* 9**,** 154.

COCKAYNE, S., FAIRHURST, C., ADAMSON, J., HEWITT, C., HULL, R., HICKS, K., KEENAN, A. M., LAMB, S. E., GREEN, L., MCINTOSH, C., MENZ, H. B.,

REDMOND, A. C., RODGERS, S., TORGERSON, D. J., VERNON, W., WATSON, J., KNAPP, P., RICK, J., BOWER, P., ELDRIDGE, S., MADURASINGHE, V. W. & GRAFFY, J. 2017. An optimised patient information sheet did not significantly increase recruitment or retention in a falls prevention study: an embedded randomised recruitment trial. *Trials,* 18**,** 144.

COCKAYNE, S., PIGHILLS, A., ADAMSON, J., FAIRHURST, C., CROSSLAND, S., DRUMMOND, A., HEWITT, C. E., RODGERS, S., RONALDSON, S. J., MCCAFFERY, J., WHITESIDE, K., SCANTLEBURY, A., ROBINSON-SMITH, L., COCHRANE, A., LAMB, S. E., BOYES, S., GILBODY, S., RELTON, C. & TORGERSON, D. J. 2021. Home environmental assessments and modification delivered by occupational therapists to reduce falls in people aged 65 years and over: the OTIS RCT. *Health Technol Assess,* 25**,** 1-118.

COOK, J., COOK, J., BONGARD, E., HENEGHAN, C. & BUTLER, C. 2021. Conditional versus Non-Conditional Incentives to Maximise Return of Participant Completed Questionnaires in Clinical Trials: A Cluster Randomised Study Within a Trial. Research Square.

CUNNINGHAM-BURLEY, R., ROCHE, J., FAIRHURST, C., COCKAYNE, S., HEWITT, C., ILES-SMITH, H. & TORGERSON, D. J. 2020. Enclosing a pen to improve response rate to postal questionnaire: an embedded randomised controlled trial. *F1000Res,* 9**,** 577.

CURTIS, L. A. & BURNS, A. 2020. *Unit Costs of Health & Social Care 2020,* PSSRU, University of Kent.

DEPARTMENT OF HEALTH AND SOCIAL CARE 2019. Health Trust Reference Costs 2017-18. *In:* DATA.GOV.UK (ed.).

DORLING, J., HEWER, O., HURD, M., BARI, V., BOSIAK, B., BOWLER, U., KING, A., LINSELL, L., MURRAY, D., OMAR, O., PARTLETT, C., ROUNDING, C., TOWNEND, J., ABBOTT, J., BERRINGTON, J., BOYLE, E., EMBLETON, N., JOHNSON, S., LEAF, A., MCCORMICK, K., MCGUIRE, W., PATEL, M., ROBERTS, T., STENSON, B., TAHIR, W., MONAHAN, M., RICHARDS, J., RANKIN, J. & JUSZCZAK, E. 2020. Two speeds of increasing milk feeds for very preterm or very low-birthweight infants: the SIFT RCT. *Health Technol Assess,* 24**,** 1-94.

DRUMMOND, M. F., SCULPHER, M., CLAXTON, K., STODDART, G. L. & TORRANCE, G. W. 2015. *Methods for the Economic Evaluation of Health Care Programmes,* New York, US, Oxford University Press.

EDWARDS, L., SALISBURY, C., HORSPOOL, K., FOSTER, A., GARNER, K. & MONTGOMERY, A. A. 2016. Increasing follow-up questionnaire response rates in a randomized controlled trial of telehealth for depression: three embedded controlled studies. *Trials,* 17**,** 107.

EDWARDS, R. T., CHARLES, J. M. & LLOYD-WILLIAMS, H. 2013. Public health economics: a systematic review of guidance for the economic evaluation of public health interventions and discussion of key methodological issues. *BMC Public Health,* 13**,** 1001.

FLETCHER, B., GHEORGHE, A., MOORE, D., WILSON, S. & DAMERY, S. 2012. Improving the recruitment activity of clinicians in randomised controlled trials: a systematic review. *BMJ Open,* 2**,** e000496.

FOUAD, M. N., JOHNSON, R. E., NAGY, M. C., PERSON, S. D. & PARTRIDGE, E. E. 2014. Adherence and retention in clinical trials: a community-based approach. *Cancer,* 120 Suppl 7**,** 1106-12.

FREE, C., HOILE, E., ROBERTSON, S. & KNIGHT, R. 2010. Three controlled trials of interventions to increase recruitment to a randomized controlled trial of mobile phone based smoking cessation support. *Clin Trials,* 7**,** 265-73.

FRIEDMAN, L. M., FURBERG, C. D., DEMETS, D. L., REBOUSSIN, D. M. & GRANGER, C. B. 2015. *Fundamentals of clinical trials*, Springer.

GATES, S., WILLIAMS, M. A., WITHERS, E., WILLIAMSON, E., MT-ISA, S. & LAMB, S. E. 2009. Does a monetary incentive improve the response to a postal questionnaire in a randomised controlled trial? The MINT incentive study. *Trials,* 10**,** 44.

GILLIES, K., KEARNEY, A., KEENAN, C., TREWEEK, S., HUDSON, J., BRUETON, V. C., CONWAY, T., HUNTER, A., MURPHY, L., CARR, P. J., RAIT, G., MANSON, P. & ACEVES-MARTINS, M. 2021. Strategies to improve retention in randomised trials. *Cochrane Database Syst Rev,* 3**,** Mr000032.

GKEKAS, A., EVANS, A., PARKER, A., RONALDSON, S. J. & TORGERSON, D. J. 2023. A systematic review of economic evaluations alongside studies within a trial (SWATs) for improving recruitment and retention in randomised controlled trials. *Research Methods in Medicine & Health Sciences,* 4**,** 94-112.

GRAY, A. M., CLARKE, P. M., WOLSTENHOLME, J. L. & WORDSWORTH, S. 2011. *Applied Methods of Cost-effectiveness Analysis in Health Care,* Oxford, United Kingdom, Oxford University Press.

GREAT ORMOND STREET HOSPITAL. 2021. *A diary from an ICU research nurse during the pandemic* [Online]. Available: https://www.gosh.nhs.uk/news/diary-icu-research-nurse-during-pandemic/ [Accessed 10/05/2024].

GRIFFIN, J., LALL, R., BRUCE, J., WITHERS, E., FINNEGAN, S. & LAMB, S. E. 2019. Comparison of alternative falls data collection methods in the Prevention of Falls Injury Trial (PreFIT). *J Clin Epidemiol,* 106**,** 32-40.

GRIMSBY, G. & JACOBS, M. A. 2016. Response to Editorial re 'Barriers to Transition in Young Adults with Neurogenic Bladder'. *J Pediatr Urol,* 12**,** 260.

HANCOCKS, H., KING, J., GUDE, A., TREWEEK, S., INGRAM, W. & THOMPSON, T. The effect of the method of invitation on recruitment of participants from GP practices to a trial of a smoking reduction intervention: a study within a trial. International clinical trials methodology conference, 2019. 143.

HARDY, P., BELL, J. L. & BROCKLEHURST, P. 2016. Evaluation of the effects of an offer of a monetary incentive on the rate of questionnaire return during follow-up of a clinical trial: a randomised study within a trial. *BMC Med Res Methodol,* 16**,** 82.

HARITON, E. & LOCASCIO, J. J. 2018. Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. *Bjog,* 125**,** 1716.

HEMMINKI, E., HOVI, S. L., VEERUS, P., SEVÓN, T., TUIMALA, R., RAHU, M. & HAKAMA, M. 2004. Blinding decreased recruitment in a prevention trial of postmenopausal hormone therapy. *J Clin Epidemiol,* 57**,** 1237-43.

HENSHALL, C., NARENDRAN, P., ANDREWS, R. C., DALEY, A., STOKES, K. A., KENNEDY, A. & GREENFIELD, S. 2018. Qualitative study of barriers to clinical trial retention in adults with recently diagnosed type 1 diabetes. *BMJ Open,* 8**,** e022353.

HEYMANS, M. W. & TWISK, J. W. R. 2022. Handling missing data in clinical research. *Journal of Clinical Epidemiology,* 151**,** 185-188.

HIGGINS, J. P., ALTMAN, D. G., GØTZSCHE, P. C., JÜNI, P., MOHER, D., OXMAN, A. D., SAVOVIC, J., SCHULZ, K. F., WEEKS, L. & STERNE, J. A. 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Bmj,* 343**,** d5928.

HORBY, P., LIM, W. S., EMBERSON, J. R., MAFHAM, M., BELL, J. L., LINSELL, L., STAPLIN, N., BRIGHTLING, C., USTIANOWSKI, A., ELMAHI, E., PRUDON, B., GREEN, C., FELTON, T., CHADWICK, D., REGE, K., FEGAN, C., CHAPPELL, L. C., FAUST, S. N., JAKI, T., JEFFERY, K., MONTGOMERY, A., ROWAN, K., JUSZCZAK, E., BAILLIE, J. K., HAYNES, R. & LANDRAY, M. J. 2021. Dexamethasone in Hospitalized Patients with Covid-19. *N Engl J Med,* 384**,** 693-704.

HRB-TRIALS METHODOLOGY RESEARCH NETWORK. 2023. *Study within a Trial (SWAT) Funding* [Online]. Available: https://www.hrb-tmrn.ie/research-innovation/study-within-a-trial-swat-funding/ [Accessed 23/09/2023].

HVIDBERG, M. F., BRINTH, L. S., OLESEN, A. V., PETERSEN, K. D. & EHLERS, L. 2015. The Health-Related Quality of Life for Patients with Myalgic Encephalomyelitis / Chronic Fatigue Syndrome (ME/CFS). *PLoS One,* 10**,** e0132421.

JACQUES, R. M., AHMED, R., HARPER, J., RANJAN, A., SAEED, I., SIMPSON, R. M. & WALTERS, S. J. 2022. Recruitment, consent and retention of participants in randomised controlled trials: a review of trials published in the National Institute for Health Research (NIHR) Journals Library (1997-2020). *BMJ Open,* 12**,** e059230.

JAMES, S., PARKER, A., COCKAYNE, S., RODGERS, S., FAIRHURST, C., TORGERSON, D. J., RHODES, S. & COTTERILL, S. 2021. Including a pen and/or cover letter, containing social incentive text, had no effect on questionnaire response rate: a factorial randomised controlled Study within a Trial [version 2; peer review: 2 approved]. *F1000Research,* 9.

JENNINGS, C. G., MACDONALD, T. M., WEI, L., BROWN, M. J., MCCONNACHIE, L. & MACKENZIE, I. S. 2015. Does offering an incentive payment improve recruitment to clinical trials and increase the proportion of socially deprived and elderly participants? *Trials,* 16**,** 80.

JOLLY, K., SIDHU, M., BOWER, P. & MADURASINGHE, V. 2019. Improving recruitment to a study of telehealth management for COPD: a cluster randomised controlled 'study within a trial' (SWAT) of a multimedia information resource. *Trials,* 20**,** 453.

KAIER, K., HEISTER, T., MOTSCHALL, E., HEHN, P., BLUHMKI, T. & WOLKEWITZ, M. 2019. Impact of mechanical ventilation on the daily costs of ICU care: a systematic review and meta regression. *Epidemiol Infect,* 147**,** e314.

KAKUMANU, S., MANNS, B. J., TRAN, S., SAUNDERS-SMITH, T., HEMMELGARN, B. R., TONELLI, M., TSUYUKI, R., IVERS, N., SOUTHERN, D., BAKAL, J. & CAMPBELL, D. J. T. 2019. Cost analysis and efficacy of recruitment strategies used in a large pragmatic community-based clinical trial targeting low-income seniors: a comparative descriptive analysis. *Trials,* 20**,** 577.

KEARNEY, A., HARMAN, N. L., ROSALA-HALLAS, A., BEECHER, C., BLAZEBY, J. M., BOWER, P., CLARKE, M., CRAGG, W., DUANE, S., GARDNER, H., HEALY, P., MAGUIRE, L., MILLS, N., ROOSHENAS, L., ROWLANDS, C., TREWEEK, S., VELLINGA, A., WILLIAMSON, P. R. & GAMBLE, C. 2018. Development of an online resource for recruitment research in clinical trials to organise and map current literature. *Clinical Trials,* 15**,** 533-542.

KEDING, A., BRABYN, S., MACPHERSON, H., RICHMOND, S. J. & TORGERSON, D. J. 2016. Text message reminders to improve questionnaire response rates. *J Clin Epidemiol,* 79**,** 90-95.

KENYON, S., PIKE, K., JONES, D., TAYLOR, D., SALT, A., MARLOW, N. & BROCKLEHURST, P. 2005. The effect of a monetary incentive on return of a postal health and development questionnaire: a randomised trial [ISRCTN53994660]. *BMC Health Serv Res,* 5**,** 55.

KHADJESARI, Z., MURRAY, E., KALAITZAKI, E., WHITE, I. R., MCCAMBRIDGE, J., THOMPSON, S. G., WALLACE, P. & GODFREY, C. 2011. Impact and costs of incentives to reduce attrition in online trials: two randomized controlled trials. *J Med Internet Res,* 13**,** e26.

KITTERMAN, D. R., CHENG, S. K., DILTS, D. M. & ORWOLL, E. S. 2011. The prevalence and economic impact of low-enrolling clinical studies at an academic medical center. *Acad Med,* 86**,** 1360-6.

KNOWLSON, C. & TORGERSON, D. J. 2020. Effects of rapid recruitment and dissemination on Covid-19 mortality: the RECOVERY trial. *F1000Res,* 9**,** 1017.

KUNTZ, K., SAINFORT, F., BUTLER, M., TAYLOR, B., KULASINGAM, S., GREGORY, S., MANN, E., ANDERSON, J. M. & KANE, R. L. 2013. AHRQ Methods for Effective Health Care. *Decision and Simulation Modeling in Systematic Reviews.* Rockville (MD): Agency for Healthcare Research and Quality (US).

LAND, S. R., RITTER, M. W., COSTANTINO, J. P., JULIAN, T. B., CRONIN, W. M., HAILE, S. R., WOLMARK, N. & GANZ, P. A. 2007. Compliance with patient-reported outcomes in multicenter clinical trials: methodologic and practical approaches. *J Clin Oncol,* 25**,** 5113-20.

LEON, A. C., MALLINCKRODT, C. H., CHUANG-STEIN, C., ARCHIBALD, D. G., ARCHER, G. E. & CHARTIER, K. 2006. Attrition in randomized controlled clinical trials: methodological issues in psychopharmacology. *Biol Psychiatry,* 59**,** 1001-5.

LIÉNARD, J. L., QUINAUX, E., FABRE-GUILLEVIN, E., PIEDBOIS, P., JOUHAUD, A., DECOSTER, G. & BUYSE, M. 2006. Impact of on-site initiation visits on patient recruitment and data quality in a randomized trial of adjuvant chemotherapy for breast cancer. *Clin Trials,* 3**,** 486-92.

LIU, Y., PENCHEON, E., HUNTER, R. M., MONCRIEFF, J. & FREEMANTLE, N. 2018. Recruitment and retention strategies in mental health trials - A systematic review. *PLoS One,* 13**,** e0203127.

LYNCH, T. R., HEMPEL, R. J., WHALLEY, B., BYFORD, S., CHAMBA, R., CLARKE, P., CLARKE, S., KINGDON, D. G., O'MAHEN, H., REMINGTON, B., RUSHBROOK, S. C., SHEARER, J., STANTON, M., SWALES, M., WATKINS, A. & RUSSELL, I. T. 2020. Refractory depression – mechanisms and efficacy of radically open dialectical behaviour therapy (RefraMED): findings of a randomised trial on benefits and harms. *The British Journal of Psychiatry,* 216**,** 204-212.

MACLENNAN, G., MCDONALD, A., MCPHERSON, G., TREWEEK, S. & AVENELL, A. 2014. Advance telephone calls ahead of reminder questionnaires increase response rate in non-responders compared to questionnaire reminders only: The RECORD phone trial. *Trials,* 15**,** 13.

MAN, M. S., RICK, J. & BOWER, P. 2015. Improving recruitment to a study of telehealth management for long-term conditions in primary care: two embedded, randomised controlled trials of optimised patient information materials. *Trials,* 16**,** 309.

MAN, M. S., TILBROOK, H. E., JAYAKODY, S., HEWITT, C. E., COX, H., CROSS, B. & TORGERSON, D. J. 2011. Electronic reminders did not improve postal questionnaire response rates or response times: a randomized controlled trial. *J Clin Epidemiol,* 64**,** 1001-4.

MARQUES, E., JOHNSON, E. C., GOOBERMAN-HILL, R., BLOM, A. W. & NOBLE, S. 2013. Using resource use logs to reduce the amount of missing data in economic evaluations alongside trials. *Value Health,* 16**,** 195-201.

MARSH, P. & KENDRICK, D. 1999. Using a diary to record near misses and minor injuries--which method of administration is best? *Inj Prev,* 5**,** 305-9.

MCCABE, C., CLAXTON, K. & CULYER, A. J. 2008. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics,* 26**,** 733-44.

MCDONALD, A. M., KNIGHT, R. C., CAMPBELL, M. K., ENTWISTLE, V. A., GRANT, A. M., COOK, J. A., ELBOURNE, D. R., FRANCIS, D., GARCIA, J., ROBERTS, I. & SNOWDON, C. 2006. What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies. *Trials,* 7**,** 9.

MCKENNA, C., GRIFFIN, S., KOFFIJBERG, H. & CLAXTON, K. 2016. Methods to place a value on additional evidence are illustrated using a case study of corticosteroids after traumatic brain injury. *J Clin Epidemiol,* 70**,** 183-90.

MCKENNA, C., SOARES, M., CLAXTON, K., BOJKE, L., GRIFFIN, S., PALMER, S. & SPACKMAN, E. 2015. Unifying Research and Reimbursement Decisions: Case Studies Demonstrating the Sequence of Assessment and Judgments Required. *Value Health,* 18**,** 865-75.

MEDICAL RESEARCH COUNCIL. 2020. *Studies Within A Trial (SWAT)* [Online]. Available: http://www.methodologyhubs.mrc.ac.uk/resources/swat [Accessed 02/01/2021].

MILLER, N. L., MARKOWITZ, J. C., KOCSIS, J. H., LEON, A. C., BRISCO, S. T. & GARNO, J. L. 1999. Cost effectiveness of screening for clinical trials by research assistants versus senior investigators. *J Psychiatr Res,* 33**,** 81-5.

MITCHELL, A. S., COOK, L., DEAN, A., FAIRHURST, C., NORTHGRAVES, M., TORGERSON, D. J. & REED, M. 2021. Using pens as an incentive for questionnaire return in an orthopaedic trial: an embedded randomised controlled retention trial [version 2; peer review: 1 approved, 1 approved with reservations]. *F1000Research,* 9.

MURPHY, E., GILLIES, K. & SHIELY, F. 2022a. How do trial teams plan for retention during the design stage of the trial? A scoping review protocol. *Trials,* 23**,** 944.

MURPHY, E., SHIELY, F. & TREWEEK, S. 2022b. How much is the lack of retention evidence costing trial teams in Ireland and the UK? *Trials,* 23**,** 396.

NAKASH, R. A. 2007. *A study of response and non-response to postal questionnaire follow-up in clinical trials.* University of Warwick.

NATIONAL HEALTH SERVICE. 2020. *Annual pay scales 2020/21* [Online]. Available: https://www.nhsemployers.org/articles/annual-pay-scales-202021 [Accessed 22/10/2023].

NATIONAL HEALTH SERVICE. 2023. *COVID-19 Hospital Activity: Reported admissions and diagnoses by age category* [Online]. Available: https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-hospital-activity/ [Accessed 02/05/2023].

NATIONAL INSTITUTE FOR HEALTH & CARE RESEARCH. 2020. *UK's Chief Medical Officers urge recruitment of more COVID-19 patients for vital scientific research* [Online]. Available: https://www.nihr.ac.uk/news/uks-chief-medical-officers-urge-recruitment-of-more-covid-19-patients-for-vital-scientific-research/24754 [Accessed 25/10/2023].

NATIONAL INSTITUTE FOR HEALTH & CARE RESEARCH. 2021. *Annual Statistics* [Online]. Available: https://www.nihr.ac.uk/about-us/who-we-are/our-research-performance/annual-statistics.htm [Accessed 10/05/2023].

NATIONAL INSTITUTE FOR HEALTH & CARE RESEARCH. 2023. *Methodological Sub-Studies: Studies within a trial or project (SWAT) and studies within a review (SWAR)* [Online]. Available: https://www.nihr.ac.uk/documents/methodological-sub-studies-studies-within-a-trial-or-project-swat-and-studies-within-a-review-swar/21512 [Accessed 12/04/2024].

NATIONAL INSTITUTE FOR HEALTH & CARE RESEARCH. 2024. *NIHR Open Data-Funded Portfolio* [Online]. Available: https://nihr.opendatasoft.com/ [Accessed 10/05/2024].

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE. 2023. *Dexamethasone* [Online]. Available: https://bnf.nice.org.uk/drug/dexamethasone.html [Accessed 15/05/2023].

NORDIC COCHRANE CENTRE, T. C. C. 2014. Review Manager (RevMan). Copenhagen, Denmark.

NYSTUEN, P. & HAGEN, K. B. 2004. Telephone reminders are effective in recruiting nonresponding patients to randomized controlled trials. *J Clin Epidemiol,* 57**,** 773-6.

PAGE, M. J., MCKENZIE, J. E., BOSSUYT, P. M., BOUTRON, I., HOFFMANN, T. C., MULROW, C. D., SHAMSEER, L., TETZLAFF, J. M., AKL, E. A., BRENNAN, S. E., CHOU, R., GLANVILLE, J., GRIMSHAW, J. M., HRÓBJARTSSON, A., LALU, M. M., LI, T., LODER, E. W., MAYO-WILSON, E., MCDONALD, S., MCGUINNESS, L. A., STEWART, L. A., THOMAS, J., TRICCO, A. C., WELCH, V. A., WHITING, P. & MOHER, D. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev,* 10**,** 89.

PARKER, A., YOUNG, B., HEMPEL, R. J., RUSSELL, I. T., WAHEED, W. & BOWER, P. 2016. What can we learn from trial decliners about improving recruitment? Qualitative study. *Trials,* 17**,** 494.

PARKER, A., YOUNG, B., WAHEED, W., SMALL, N. & BOWER, P. 2015. Factors affecting recruitment into depression trials: Systematic review, meta-synthesis and conceptual framework. *J Affect Disord,* 172**,** 274-90.

PETERSON, J. C., PIRRAGLIA, P. A., WELLS, M. T. & CHARLSON, M. E. 2012. Attrition in longitudinal randomized controlled trials: home visits make a difference. *BMC Med Res Methodol,* 12**,** 178.

RESNIK, D. B. 2008. Randomized controlled trials in environmental health research: ethical issues. *J Environ Health,* 70**,** 28-30.

REYNOLDS, S. A., O'CONNOR, L., MCGEE, A., KILCOYNE, A. Q., CONNOLLY, A., MOCKLER, D., GUINAN, E. & O'NEILL, L. 2023. Recruitment rates and strategies in exercise trials in cancer survivorship: a systematic review. *J Cancer Surviv*.

RILEY, R. D., HIGGINS, J. P. & DEEKS, J. J. 2011. Interpretation of random effects meta-analyses. *Bmj,* 342**,** d549.

ROGERS, A., FLYNN, R. W. V., MACKENZIE, I. S. & MACDONALD, T. M. 2019. Does the provision of a DVD-based audio-visual presentation improve recruitment in a clinical trial? A randomised trial of DVD trial invitations. *BMC Med Res Methodol,* 19**,** 24.

SAKPAL, T. V. 2010. Sample size estimation in clinical trial. *Perspect Clin Res,* 1**,** 67-9.

SALDANHA, I. J., SKELLY, A. C., LEY, K. V., WANG, Z., BERLINER, E., BASS, E. B., DEVINE, B., HAMMARLUND, N., ADAM, G. P., DUAN-PORTER, D., BAÑEZ, L. L., JAIN, A., NORRIS, S. L., WILT, T. J., LEAS, B., SIDDIQUE, S. M., FIORDALISI, C. V., PATINO-SUTTON, C. & VISWANATHAN, M. 2022. AHRQ Methods for Effective Health Care. *Inclusion of Nonrandomized Studies of Interventions in Systematic Reviews of Intervention Effectiveness: An Update.* Rockville (MD): Agency for Healthcare Research and Quality (US).

SERVELLITA, V., MORRIS, M. K., SOTOMAYOR-GONZALEZ, A., GLIWA, A. S., TORRES, E., BRAZER, N., ZHOU, A., HERNANDEZ, K. T., SANKARAN, M., WANG, B., WONG, D., WANG, C., ZHANG, Y., REYES, K. R., GLASNER, D., DENG, X., STREITHORST, J., MILLER, S., FRIAS, E., RODGERS, M., CLOHERTY, G., HACKETT, J., JR., HANSON, C., WADFORD, D., PHILIP, S.,

TOPPER, S., SACHDEV, D. & CHIU, C. Y. 2022. Predominance of antibody-resistant SARS-CoV-2 variants in vaccine breakthrough cases from the San Francisco Bay Area, California. *Nat Microbiol,* 7**,** 277-288.

SEVERI, E., FREE, C., KNIGHT, R., ROBERTSON, S., EDWARDS, P. & HOILE, E. 2011. Two controlled trials to increase participant retention in a randomized controlled trial of mobile phone-based smoking cessation support in the United Kingdom. *Clin Trials,* 8**,** 654-60.

SHARP, L., COCHRAN, C., COTTON, S. C., GRAY, N. M. & GALLAGHER, M. E. 2006. Enclosing a pen with a postal questionnaire can significantly increase the response rate. *J Clin Epidemiol,* 59**,** 747-54.

SHEINSON, D., DANG, J., SHAH, A., MENG, Y., ELSEA, D. & KOWAL, S. 2021. A Cost-Effectiveness Framework for COVID-19 Treatments for Hospitalized Patients in the United States. *Adv Ther,* 38**,** 1811-1831.

ST SAUVER, J. L., LOPES, G. S., ROCCA, W. A., PRASAD, K., MAJERUS, M. R., LIMPER, A. H., JACOBSON, D. J., FAN, C., JACOBSON, R. M., RUTTEN, L. J., NORMAN, A. D. & VACHON, C. M. 2021. Factors Associated With Severe COVID-19 Infection Among Persons of Different Ages Living in a Defined Midwestern US Population. *Mayo Clin Proc,* 96**,** 2528-2539.

STARR, K., MCPHERSON, G., FORREST, M. & COTTON, S. C. 2015. SMS text pre-notification and delivery of reminder e-mails to increase response rates to postal questionnaires in the SUSPEND trial: a factorial design, randomised controlled trial. *Trials,* 16**,** 295.

STUART, M., SPENCER, D. A., MCLACHLAN, C. J. & FORDE, C. 2021. COVID-19 and the uncertain future of HRM: Furlough, job retention and reform. 31**,** 904-17.

SUTHERLAND, E., JANSSEN, B. & CABASES, J. 2021. *Coronavirus (COVID-19) Infection Survey technical article: waves and lags of COVID-19 in England, June 2021* [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19infectionsurveytechnicalarticle/wavesandlagsofcovid19inenglandjune2021 [Accessed 22/02/2022].

SYDES, M., GILLIES, K., WILLIAMSON, P. & AL., E. 2023. ICTMC 2022: 6th International Clinical Trials Methodology Conference -- book of abstracts. *Zenodo*.

SZENDE, A., JANSSEN, B. & CABASES, J. 2014. Self-Reported Population Health: An International Perspective based on EQ-5D. Dordrecht (NL): Springer.

TACCONELLI, E. 2010. Systematic reviews: CRD's guidance for undertaking reviews in health care. *The Lancet Infectious Diseases,* 10**,** 226.

TAQUET, M., DERCON, Q., LUCIANO, S., GEDDES, J. R., HUSAIN, M. & HARRISON, P. J. 2021. Incidence, co-occurrence, and evolution of long-COVID features: A 6-month retrospective cohort study of 273,618 survivors of COVID-19. *PLoS Med,* 18**,** e1003773.

TORGERSON, D. J. & TORGERSON, C. J. 2008. Designing Randomised Trials in Health, Education and the Social Sciences: An Introduction. London: Palgrave Macmillan UK.

TRANBERG, M., BECH, B. H., BLAAKÆR, J., JENSEN, J. S., SVANHOLM, H. & ANDERSEN, B. 2018. Preventing cervical cancer using HPV self-sampling: direct mailing of test-kits increases screening participation more than timely opt-in procedures - a randomized controlled trial. *BMC Cancer,* 18**,** 273.

TREVENA, L., IRWIG, L. & BARRATT, A. 2006. Impact of privacy legislation on the number and characteristics of people who are recruited for research: a randomised controlled trial. *J Med Ethics,* 32**,** 473-7.

TREWEEK, S., ALTMAN, D. G., BOWER, P., CAMPBELL, M., CHALMERS, I., COTTON, S., CRAIG, P., CROSBY, D., DAVIDSON, P., DEVANE, D., DULEY, L., DUNN, J., ELBOURNE, D., FARRELL, B., GAMBLE, C., GILLIES, K., HOOD, K., LANG, T., LITTLEFORD, R., LOUDON, K., MCDONALD, A., MCPHERSON, G., NELSON, A., NORRIE, J., RAMSAY, C., SANDERCOCK, P., SHANAHAN, D. R., SUMMERSKILL, W., SYDES, M., WILLIAMSON, P. & CLARKE, M. 2015. Making randomised trials more efficient: report of the first meeting to discuss the Trial Forge platform. *Trials,* 16**,** 261.

TREWEEK, S., BEVAN, S., BOWER, P., BRIEL, M., CAMPBELL, M., CHRISTIE, J., COLLETT, C., COTTON, S., DEVANE, D., EL FEKY, A., GALVIN, S., GARDNER, H., GILLIES, K., HOOD, K., JANSEN, J., LITTLEFORD, R., PARKER, A., RAMSAY, C., RESTRUP, L., SULLIVAN, F., TORGERSON, D., TREMAIN, L., VON ELM, E., WESTMORE, M., WILLIAMS, H., WILLIAMSON, P. R. & CLARKE, M. 2020. Trial Forge Guidance 2: how to decide if a further Study Within A Trial (SWAT) is needed. *Trials,* 21**,** 33.

TREWEEK, S., BEVAN, S., BOWER, P., CAMPBELL, M., CHRISTIE, J., CLARKE, M., COLLETT, C., COTTON, S., DEVANE, D., EL FEKY, A., FLEMYNG, E., GALVIN, S., GARDNER, H., GILLIES, K., JANSEN, J., LITTLEFORD, R., PARKER, A., RAMSAY, C., RESTRUP, L., SULLIVAN, F., TORGERSON, D., TREMAIN, L., WESTMORE, M. & WILLIAMSON, P. R. 2018a. Trial Forge Guidance 1: what is a Study Within A Trial (SWAT)? *Trials,* 19**,** 139.

TREWEEK, S., GALLANT, S. & ANDERSON, A. S. 2021. SWAT 76 evaluation: randomised evaluation of sending pre-notification cards to trial participants before a face-to-face primary outcome measurement to increase attendance. *F1000Res,* 10**,** 84.

TREWEEK, S., PITKETHLY, M., COOK, J., FRASER, C., MITCHELL, E., SULLIVAN, F., JACKSON, C., TASKILA, T. K. & GARDNER, H. 2018b. Strategies to improve recruitment to randomised trials. *Cochrane Database Syst Rev,* 2**,** Mr000013.

TRIAL FORGE. 2023. *A systematic approach to making trials more efficient* [Online]. Available: https://www.trialforge.org [Accessed 25/10/2023].

UK HEALTH SECURITY AGENCY. 2022. *Deaths in United Kingdom (UK Coronavirus Dashboard)* [Online]. Available: https://coronavirus.data.gov.uk/details/deaths [Accessed 07/06/2022].

VEKARIA, B., OVERTON, C., WIŚNIOWSKI, A., AHMAD, S., APARICIO-CASTRO, A., CURRAN-SEBASTIAN, J., EDDLESTON, J., HANLEY, N. A., HOUSE, T., KIM, J., OLSEN, W., PAMPAKA, M., PELLIS, L., RUIZ, D. P., SCHOFIELD, J., SHRYANE, N. & ELLIOT, M. J. 2021. Hospital length of stay for COVID-19 patients: Data-driven methods for forward planning. *BMC Infect Dis,* 21**,** 700.

VILLARRUEL, A. M., JEMMOTT, J. B., 3RD & JEMMOTT, L. S. 2006. A randomized controlled trial testing an HIV prevention intervention for Latino youth. *Arch Pediatr Adolesc Med,* 160**,** 772-7.

WALTERS, S. J., BONACHO DOS ANJOS HENRIQUES-CADBY, I., BORTOLAMI, O., FLIGHT, L., HIND, D., JACQUES, R. M., KNOX, C., NADIN, B., ROTHWELL, J., SURTEES, M. & JULIOUS, S. A. 2017. Recruitment and retention of participants in randomised controlled trials: a review of trials funded and published by the United Kingdom Health Technology Assessment Programme. *BMJ Open,* 7**,** e015276.

WARLOW, C. 2003. Clinical trials for tomorrow funded by the MRC. *Bmj,* 327**,** 240-1.

WATSON, J. M. & TORGERSON, D. J. 2006. Increasing recruitment to randomised trials: a review of randomised controlled trials. *BMC Med Res Methodol,* 6**,** 34.

WELLS, A. A., PALINKAS, L. A., WILLIAMS, S.-L. L. & ELL, K. 2015. Retaining Low-Income Minority Cancer Patients in a Depression Treatment Intervention Trial: Lessons Learned. *Community Mental Health Journal,* 51**,** 715-722.

WHITESIDE, K., FLETT, L., MITCHELL, A., FAIRHURST, C., COCKAYNE, S., RODGERS, S. & TORGERSON, D. 2019. Using pens as an incentive for trial recruitment of older adults: An embedded randomised controlled trial. *F1000Res,* 8**,** 315.

WILSON, E. C. 2015. A practical guide to value of information analysis. *Pharmacoeconomics,* 33**,** 105-21.

WISE, J. & COOMBES, R. 2020. Covid-19: The inside story of the RECOVERY trial. *Bmj,* 370**,** m2670.

WONG, A. D., KIRBY, J., GUYATT, G. H., MOAYYEDI, P., VORA, P. & YOU, J. J. 2013. Randomized controlled trial comparing telephone and mail follow-up for recruitment of participants into a clinical trial of colorectal cancer screening. *Trials,* 14**,** 40.

WORLD BANK. 2020. *Inflation, GDP deflator (annual %) - United Kingdom* [Online]. Available: https://data.worldbank.org/indicator/NY.GDP.DEFL.KD.ZG?locations=GB [Accessed 01/03/2022].

WORLD BANK. 2023. *GDP deflator (base year varies by country) - United Kingdom* [Online]. Available: https://data.worldbank.org/indicator/NY.GDP.DEFL.ZS?locations=GB [Accessed 01/12/2023].

YORK HEALTH ECONOMICS CONSORTIUM. 2016a. *Decision Tree* [Online]. Available: https://yhec.co.uk/glossary/decision-tree [Accessed 23/09/2021].

YORK HEALTH ECONOMICS CONSORTIUM. 2016b. *Probabilistic/Stochastic Sensitivity Analysis* [Online]. Available: https://yhec.co.uk/glossary/probabilistic-stochastic-sensitivity-analysis/ [Accessed 23/09/2021].

Appendix

Supplemental Material 2.1. Estimation of pathway probabilities

- The randomisation point of the hospitalised COVID-19 patients to the RECOVERY trial was assumed to accurately represent the initial decision node at any hospital in real clinical settings.
- The probabilities associated with "admission" can be found in *Table 1* of the research article of the RECOVERY trial, which was also used to estimate all input probabilities but long COVID (Horby et al., 2021).
- The probabilities of the health outcomes related to invasive ventilation at the point of entry/randomisation can be found in *Figure 3* of the research article. For instance, 95 out of 324 invasively ventilated patients (at the randomisation point) died in the case of dexamethasone (P (death | invasive ventilation, Dexamethasone) = 0.293).
- The remaining probabilities were manually computed according to the available data from the RECOVERY trial.

For example, for the Dexamethasone arm, the following steps were followed:

Step 1: 365 out of 1279 COVID-19 inpatients receiving non-invasive ventilation at randomisation, died and/or received invasive ventilation (Figure S2, Supplementary Appendix).

Step 2: 298 out of 1279 inpatients receiving non-invasive ventilation at randomisation, died (*Figure 3*, research article).

Step 3: From *Figure 3*, 89 out of 501 inpatients receiving no ventilation at randomisation, died.

Step 4: From *Table 2.2*, 25/501 inpatients in acute wards went for non-invasive ventilation and/or invasive ventilation later on (p=0.05). 20 received non-invasive ventilation and 9 received invasive ventilation. According to the rules of probability, 5 of them went straightaway to invasive ventilation ((P (invasive ventilation | acute hospital ward, Dexamethasone) = 0.018), and 16 received non-invasive ventilation immediately ((P (non-invasive ventilation | acute hospital ward, Dexamethasone) = 0.032).

Step 5: Since 9 patients not receiving ventilation at randomisation received invasive ventilation, 101 out of 1279 patients receiving non-invasive ventilation at randomisation

received invasive ventilation, according to *Table 2* of the research article. Therefore, ((P (invasive ventilation | non-invasive ventilation, Dexamethasone)) = 0.079.

Step 6: Considering Steps 1, 2 and 5, as well as the rules of probability, 34 out of 1279 inpatients, receiving non-invasive ventilation at randomisation, received invasive ventilation and died. Therefore, from Step 2, 264 out of 1279 inpatients receiving non-invasive ventilation at randomisation, died, without receiving any invasive ventilation. Therefore, ((P (death | non-invasive ventilation, Dexamethasone) = 0.206.

To summarise:

1. 264 out of 1279 inpatients receiving non-invasive ventilation at randomisation, died, without receiving any invasive ventilation.
2. 34 out of 1279 inpatients receiving non-invasive ventilation at randomisation, died, after having received invasive ventilation
3. 67 out of 1279 inpatients receiving non-invasive ventilation at randomisation, survived, after having received invasive ventilation
4. 914 out of 1279 inpatients receiving non-invasive ventilation at randomisation, surviving, without receiving any invasive ventilation. ((P (survival | non-invasive ventilation, Dexamethasone) =0.715.

Step 7: Given the yielded probabilities, P(death| acute hospital ward, dexamethasone) was computed manually.

1. From Figure 3, we know that 89 out of 501 patients receiving no ventilation, eventually died.
2. Patient deaths given no ventilation at randomisation, followed by non-invasive ventilation= 16 * (P (death | non-invasive ventilation, Dexamethasone) = 0.206) = 3 (see Step 4).
3. Patient deaths given no oxygen at randomisation, followed by invasive ventilation= 9 * ((P (death | invasive ventilation, Dexamethasone) = 0.293) = 3 (see Step 4).
4. Therefore, P (death | acute hospital ward, dexamethasone) = 0.166 and P (survival | acute hospital ward, Dexamethasone) = 0.784, the sum of which is indeed equal to 0.95 (see Step 4).

The same computation process was followed for the No Dexamethasone group.

Supplemental Material 3.1. Types and subtypes of participant loss to follow-up in the OTIS trial

Table 5 in the published study of the OTIS trial provides details about rates of participant loss to follow-up and questionnaire response rates (Cockayne et al., 2021).

- Before 4-month questionnaire, 16 participants were lost to follow-up in the intervention group, whereas 12 participants were lost to follow-up in the usual care group (Cockayne et al., 2021). It is also reported that 11 participants in the intervention group and 25 participants in the usual care group did not respond to the 4-month follow-up questionnaire (Cockayne et al., 2021).
- Before 8-month questionnaire, nine more participants were lost to follow-up in the intervention group and 15 more participants were lost to follow-up in the usual care group (Cockayne et al., 2021). In addition, six participants did not respond to the 8-month follow-up questionnaire in the intervention group and 28 participants in the usual care group did not respond to the 8-month follow-up questionnaire (Cockayne et al., 2021).
- Before 12-month questionnaire, four more participants were lost to follow-up in the intervention group and 12 more participants were lost to follow-up in the usual care group (Cockayne et al., 2021). In addition, 12 participants did not respond to the 12-month follow-up questionnaire in the intervention group and 24 participants in the usual care group did not respond to the 12-month follow-up questionnaire (Cockayne et al,. 2021).

The sample sizes of participants related to all attrition types (i.e. 1,2,3,4) were obtained directly from the aforementioned study findings. Nevertheless, due to data unavailability assumptions about the sample sizes of participants related to the subtypes of attrition (i.e. 2a, 2b, 3a, 3b, 3c, 3d,4a, 4b) by intervention group were made.

- With regards to Subtypes 2a and 2b, it was assumed that all participants lost to follow-up before the 8-month follow-up period had not responded to the 4-month follow-up questionnaire, if the number of participants not responding to the 4-month questionnaire was not lower than the number of participants lost to follow-up. In the intervention group, this means all nine participants lost to follow-up before the 8-month questionnaire had not responded to the 4-month follow-up questionnaire were assumed to be lost to follow-up (Subtype 2b). Thus, the number of participants lost to follow-up, belonging to Subtype 2a, is zero, in addition to two participants who did not respond to the 4-month questionnaire but were considered not to have been lost to follow-up before the 12-month questionnaire.

In the usual care group, since 25 participants did not respond to the 4-month questionnaire and 15 participants were lost to follow-up before the 8-month questionnaire, all participants lost to follow-up before the 8-month questionnaire were assumed not to have responded to the 4-month questionnaire (Subtype 2b), in addition to 10 participants who did not respond to the 4-month questionnaire but were considered not to have been lost to follow-up before the 12-month questionnaire.

- Moreover, in the intervention group, since all participants lost to follow-up before the 8-month questionnaire were assumed not to have responded to the 4-month questionnaire, and since the number of participants not responding to the 4-month questionnaire was greater than the number of participants lost to follow-up by two participants, the cumulative sample size for subgroups 3b and 3d is two. In the control group, the corresponding cumulative sample size for subgroups 3b and 3d is 10, given that 10 participants not responding to the 4-month questionnaire were regarded not to have been lost to follow-up before the 12-month follow-up period.
- Similarly, it was assumed that all participants lost to follow-up before the 12-month follow-up period had not responded to the 8-month follow-up questionnaire. In the intervention group, since four participants were lost to follow-up before the 12-month questionnaire and six participants did not respond to the 8-month questionnaire, the cumulative sample size for subgroups 3b and 3c is four. Also, it seems that two out of six participants who did not respond to the 8-month questionnaire were not lost to follow-up before the 12-month questionnaire; we assume that 50% of such participants, i.e. one participant, also did not respond to the 12-month questionnaire (Subtype 4b). Given that the cumulative sample size for subgroups 3b and 3d is two, it means that two participants fall to subtype 3b and two participants fall to subtype 3c. In the usual care group, since 12 participants were lost to follow-up before the 12-month questionnaire and 28 participants did not respond to the 8-month questionnaire, the cumulative sample size for subtypes 3b and 3c is 12. Also, it seems that 16 out of 28 participants who did not respond to the 8-month questionnaire were not lost to follow-up before the 12-month questionnaire; we assume that 50% of such participants, i.e. eight participants, also did not respond to the 12-month questionnaire (Subtype 4b). Given that the cumulative sample size for subgroups 3b and 3d is 10, it means that 10 participants fall to subtype 3b and two participants fall to Subtype 3c.
- Finally, since 12 participants in the intervention group and 24 participants in the control group were lost to follow-up due to not responding to the 12-month questionnaire, and one participant in the intervention group and 8 participants in the control group fall under

Subtype 4b, 11 participants from the intervention group are expected to have been lost from follow-up under Subtype 4a, with the figure in the control group being 16.

Supplemental Material 3.2. Estimation of the monthly falls calendar returns by post by type and subtype of attrition

Table 4 in the published study of the OTIS trial provide details about rates of monthly falls calendars return rates (Cockayne et al., 2021).

Between the baseline and the 4-month follow-up questionnaire, 66 participants did not respond to the monthly falls calendars. Between the baseline and the 8-month follow-up questionnaire, 91 participants did not respond to the monthly falls calendars. Between the baseline and the 12-month follow-up questionnaire, 130 participants did not respond to the monthly falls calendars.

To estimate the number of participants lost to follow-up responding to monthly falls calendars, we assumed that participants not sending a falls calendar in a month x wouldn't send any subsequent falls calendars. We also assume that such participants had sent all falls calendars related to the previous months prior to the falls calendar related to the month affected. Finally, we assume that participants not sending falls calendars are also lost to follow-up as the calendar return rates were similar to the attrition rates.

Therefore, for the 4-month follow-up period, given that:

1) 26 participants did not send a falls calendar at month 0

2) 12 participants did not send a falls calendar at month 1

3) 6 participants did not send a falls calendar at month 2

4) 9 participants did not send a falls calendar at month 3

5) 13 participants did not send a falls calendar at month 4

Considering the total figure of 66 participants not responding to the monthly falls calendars at month 4, which coincides with the 4-month follow-up questionnaire in terms of time, 39.39% (26 out of 66) of these participants did not respond at month 0, 18.18% (12 out of 66) of these participants did not respond at month 1, 9.09% (6 out of 66 participants) did not respond at

month 2, 13.64% (9 out of 66 participants) did not respond at month 3 and 19.70% (13 out of 66 participants) did not respond at month 4.

Therefore, given our assumptions the average rate of monthly falls calendar returns for participants lost to follow-up before the 4-month questionnaire is the weighted average of:

39.39%*0+ 18.18%*1+ 9.09%*2+ 13.64%*3+ 19.70%*4= 1.56 monthly calendar fall returns

Correspondingly, for the 8-month follow-up period, given that:

1) 11 participants did not send a falls calendar at month 5

2) 5 participants did not send a falls calendar at month 6

3) -1 participants did not send a falls calendar at month 7

4) 10 participants did not send a falls calendar at month 8

Considering the total figure of 25 participants not responding to the monthly falls calendars at month 8, which coincides with the 8-month follow-up questionnaire in terms of time, 44% (11 out of 25) of these participants did not respond at month 5, 20% (5 out of 25) of these participants did not respond at month 6, -4% (-1 out of 25 participants) did not respond at month 7 and 40% (10 out of 25 participants) did not respond at month 8.

Therefore, the average rate of monthly falls calendar returns for participants lost to follow-up before the 8-month questionnaire is the weighted average of:

44%*5+ 20%*6- 4%*7+ 40%*8= 6.32 monthly calendar fall returns

Following the same approach, it can be shown that the average rate of monthly falls calendar returns for participants lost to follow-up before the 12-month questionnaire is the weighted average of:

15.38%*9+ 30.77%10+41.03%*11+12.82%*12= 10.51 monthly calendar fall returns

Supplemental Material 3.3. Randomised trials funded by the NIHR in 2016/2017

Table 3.S1: Pilot and full-scale randomised trials funded by the NIHR HTA in 2016/2017 (from 01/04/2016 to 31/03/2017) (National Institute for Health & Care Research, 2024)

Project ID	**Award Value**	**Aggregate funding allocated by the NIHR HTA to randomised trials in 2016/17**	**Cost per randomised trial funded by the NIHR HTA in 2016/17**
15/118/01	£ 467,437.65	**£67,553,088.13**	**£1,535,297.46**
15/08/40	£ 475,575.73		
15/44/01	£ 488,196.87		
14/49/149 (OTIS Trial)	**£ 722,096.59**		
15/57/160	£ 811,989.21		
14/49/154	£ 853,259.81		
13/115/62	£ 935,504.72		
13/82/04	£ 965,630.44		
14/192/109	£ 1,011,672.23		
14/49/159	£ 1,016,564.59		
13/155/05	£ 1,047,206.83		
13/142/04	£ 1,118,326.00		
14/28/02	£ 1,147,564.92		
15/106/04	£ 1,207,594.53		
15/110/02	£ 1,222,886.63		
15/80/39	£ 1,242,497.00		
15/26/06	£ 1,249,382.00		
14/192/97	£ 1,259,771.62		
14/166/01	£ 1,289,952.72		
14/222/02	£ 1,412,681.43		
14/216/01	£ 1,457,334.00		
14/08/60	£ 1,461,432.00		
14/192/71	£ 1,474,733.93		
15/38/04	£ 1,566,143.53		
13/34/64	£ 1,569,080.49		
14/224/04	£ 1,631,592.55		
14/192/110	£ 1,647,360.00		
15/160/02	£ 1,715,379.38		
14/49/84	£ 1,789,594.60		
13/115/29	£ 1,801,970.92		

15/57/02	£ 1,807,350.81		
13/115/48	£ 1,807,618.00		
13/146/02	£ 1,819,018.42		
14/68/09	£ 1,837,662.40		
14/22/23	£ 1,842,732.00		
14/192/53	£ 1,908,190.00		
15/102/04	£ 1,922,993.79		
15/35/03	£ 1,984,963.63		
15/166/08	£ 2,051,482.24		
14/68/08	£ 2,142,216.63		
15/43/07	£ 2,387,728.27		
15/57/66	£ 2,522,709.93		
15/57/39	£ 3,012,766.73		
15/57/143	£ 4,445,242.36		

Supplemental Material 4.1. Data extraction of the included studies

Study	Jennings et al. (2015a)
	Host Trial
Name	The Standard Care Versus Celecoxib Outcome Trial (SCOTLSSS)
Design	Randomised Controlled Trial (RCT); Prospective Randomised Open Blinded End point (PROBE) design.
Location	United Kingdom
Setting	Primary care
Population	People aged 60 or over taking long-term NSAIDS for arthritis.
Intervention(s)	200-400 mg of Celecoxib daily in divided doses.
Comparator(s)	Traditional NSAIDs prescribed for the treatment of arthritis.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised; 1:1 simple fixed randomisation; no stratification, patients blinded.
Strategy(-ies) and study objective	Invitation letter with a fixed payment of £100 to improve recruitment to the host trial.
Comparator(s)	An invitation letter with no fixed payment incentive.
Frequency of strategy	Once, if response is received. Twice if there was no response to the first invitation letter.
Measure(s) of benefit	Increase in positive response to recruitment. Increase in consented patients with incentive.
Costs	Incremental cost per additional patient recruited. Incremental cost per additional consented patient.
Type of economic evaluation	Cost-effectiveness analysis.
Numbers and/or proportions of participants in the intervention and control groups	Offered incentive payment: n= 84 (46.4%) Not offered incentive payment: n=97 (53.6%) Total sample size: n= 181 (100 %)
Results	Primary: Incremental cost per patient recruited (ICER): £1,276 Incremental cost per consented patient (ICER): £1,613 Other: Increase in positive response to recruitment: 8.5% (34/84 in the intervention group vs 31/97 in the control group) Increase in consented patients with incentive: 6.2% (26/84 in the intervention group vs 24/97 in the control group)
Perspective adopted	Trial team.

Study	Jennings et al. (2015b)
	Host Trial
Name	Febuxostat versus Allopurinol Streamlined Trial (FAST)
Design	Randomised Controlled Trial (RCT); Prospective Randomised Open Blinded End point (PROBE) design.
Location	United Kingdom
Setting	Primary care
Population	People aged 60 or over with chronic hyperuricaemia in conditions where urate deposition has already occurred.
Intervention(s)	Febuxostat
Comparator(s)	Allopurinol
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised; 1:1 simple fixed randomisation; no stratification, patients blinded.
Strategy(-ies) and study objective	Invitation letter with a fixed payment of £100 to improve recruitment to the host trial.
Comparator(s)	An invitation letter with no fixed payment incentive.
Frequency of strategy	Once, if response is received. Twice if there was no response to the first invitation letter.
Measure(s) of benefit	Increase in positive response to recruitment. Increase in consented patients with incentive.
Costs	Incremental cost per additional patient recruited. Incremental cost per additional consented patient.
Type of economic evaluation	Cost-effectiveness analysis.
Numbers and/or proportions of participants in the intervention and control groups	Offered incentive payment: n= 158 (47.6%) Not offered incentive payment: n=174 (52.4%) Total sample size: n= 332 (100 %)
Results	Primary: Incremental cost per patient recruited (ICER): £933 Incremental cost per consented patient (ICER): £763 Other: Increase in positive response to recruitment: 12.0% (68/158 in the intervention group vs 54/174 in the control group) Increase in consented patients with incentive: 13.1% (58/158 in the intervention group vs 41/174 in the control group)
Perspective adopted	Trial team.

Study	Jennings et al. (2015c)
	Host Trial
Name	Prevention and Treatment of Hypertension with Algorithm Guided Therapy study 1(PATHWAY 1)
Design	Randomised Controlled Trial (RCT); open, parallel, double-blind
Location	United Kingdom
Setting	Primary care
Population	People aged 18-79 with newly diagnosed hypertension.
Intervention(s)	Monotherapy as initial hypertension treatment.
Comparator(s)	Dual therapy as initial hypertension treatment.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised; 1:1 simple fixed randomisation; no stratification, patients blinded.
Strategy(-ies) and study objective	Invitation letter with a fixed payment of £100 to improve recruitment to the host trial.
Comparator(s)	An invitation letter with no fixed payment incentive.
Frequency of strategy	Once, if response is received. Twice if there was no response to the first invitation letter.
Measure(s) of benefit	Increase in positive response to recruitment. Increase in consented patients with incentive.
Costs	Incremental cost per additional patient recruited. Incremental cost per additional consented patient.
Type of economic evaluation	Cost-effectiveness analysis.
Numbers and/or proportions of participants in the intervention and control groups	Offered incentive payment: n= 46 (49.5%) Not offered incentive payment: n=47 (50.5 %) Total sample size: n= 93 (100 %)
Results	Primary: Incremental cost per patient recruited (ICER): N/A Incremental cost per consented patient (ICER): N/A Other: Increase in positive response to recruitment: -4% (5/46 in the intervention group vs 7/47 in the control group) Increase in consented patients with incentive: -2% (3/46 in the intervention group vs 4/47 in the control group)
Perspective adopted	Trial team.

Study	Jennings et al. (2015d)
	Host Trial
Name	Prevention and Treatment of Hypertension with Algorithm Guided Therapy study 2 (PATHWAY 2)
Design	Randomised Controlled Trial (RCT); open, parallel, double-blind
Location	United Kingdom
Setting	Primary care
Population	People aged 18-79 with uncontrolled blood pressure.
Intervention(s)	Spironolactone or doxazosin or bisoprolol
Comparator(s)	Placebo drug
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised; 1:1 simple fixed randomisation; no stratification, patients blinded.
Strategy(-ies) and study objective	Invitation letter with a fixed payment of £100 to improve recruitment to the host trial.
Comparator(s)	An invitation letter with no fixed payment incentive.
Frequency of strategy	Once, if response is received. Twice if there was no response to the first invitation letter.
Measure(s) of benefit	Increase in positive response to recruitment. Increase in consented patients with incentive.
Costs	Incremental cost per additional patient recruited. Incremental cost per additional consented patient.
Type of economic evaluation	Cost-effectiveness analysis.
Numbers and/or proportions of participants in the intervention and control groups	Offered incentive payment: n= 101 (48.1%) Not offered incentive payment: n=109 (51.9%) Total sample size: n= 210 (100 %)
Results	Primary: Incremental cost per patient recruited (ICER): £7,248 Incremental cost per consented patient (ICER): N/A Other: Increase in positive response to recruitment: 1.4% (19/210 in the intervention group vs 19/210 in the control group) Increase in consented patients with incentive: -4.3 % (4/210 in the intervention group vs 9/210 in the control group)
Perspective adopted	Trial team.

Study	Jennings et al. (2015e)
	Host Trial
Name	Prevention and Treatment of Hypertension with Algorithm Guided Therapy study 3 (PATHWAY 3)
Design	Randomised Controlled Trial (RCT); open, parallel, double-blind
Location	United Kingdom
Setting	Primary care
Population	People aged 18-80 with at least one component of the metabolic syndrome.
Intervention(s)	Single-agent diuretic therapy for low-renin hypertension.
Comparator(s)	Combination diuretic therapy for low-renin hypertension.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised; 1:1 simple fixed randomisation; no stratification, patients blinded.
Strategy(-ies) and study objective	Invitation letter with a fixed payment of £100 to improve recruitment to the host trial.
Comparator(s)	An invitation letter with no fixed payment incentive.
Frequency of strategy	Once, if response is received. Twice if there was no response to the first invitation letter.
Measure(s) of benefit	Increase in positive response to recruitment. Increase in consented patients with incentive.
Costs	Incremental cost per additional patient recruited. Incremental cost per additional consented patient.
Type of economic evaluation	Cost-effectiveness analysis.
Numbers and/or proportions of participants in the intervention and control groups	Offered incentive payment: n= 92 (46.2 %) Not offered incentive payment: n=107 (53.8%) Total sample size: n= 199 (100 %)
Results	Primary: Incremental cost per patient recruited (ICER): £1,249 Incremental cost per consented patient (ICER): £2,381 Other: Increase in positive response to recruitment: 8.7 % (26/92 in the intervention group vs 21/107 in the control group) Increase in consented patients with incentive: 4.2% (9/92 in the intervention group vs 6/107 in the control group)
Perspective adopted	Trial team.

Study	Free et al. (2010a)
	Host Trial
Name	Txt2stop
Design	Randomised Controlled Trial (RCT); pilot, single-blind
Location	United Kingdom
Setting	Community
Population	People aged 16 and above who are smokers and willing to stop smoking in next month.
Intervention(s)	A composite mobile phone-based smoking cessation support.
Comparator(s)	Simple, short, generic text messages.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised; single-blind, with concealed allocation.
Strategy(-ies) and study objective	Research staff sending a text message regarding the newly available online registration facility.
Comparator(s)	Research staff calling the participants' mobile numbers to register them for the trial (no text message).
Frequency of strategy	Once, before recruitment to the Txt2stop trial.
Measure(s) of benefit	1) Registration to the Txt2stop trial at two weeks by eligible participants. 2) Registration using the online facility by two weeks by eligible participants. 3) Completed registrations at two weeks, including both eligible and ineligible participants. 4) All completed registrations at 2 weeks using the online registration facility.
Costs	Incremental cost of sending a text message for each eligible participant.
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	SMS messages: n= 470 (50.2%) Calls: n=467 (49.8%) Total sample size: n=937 (100%)
Results	Secondary: Incremental cost per participant registered = £2.34 (based on a cost of £0.05 for each SMS message sent and 15 min of data manager's time) Unit cost of SMS: £0.05 Other: 1) 3.6% (17/470) of participants who were sent the text message regarding the new online registration facility were registered successfully, for the trial within 2 weeks, compared with 1.1% (5/467) of the control group. The risk difference is 2.5% (95% confidence interval: 0.6–4.5). 2) None of the intervention group registered successfully online, compared with 0.2% (1/467) of the control group. The risk difference is -0.2 (95% confidence interval: -0.6 – 0.2). 3) 4.5% (21/470) of the intervention group and 1.5% (7/467) of the control group attempted to register for the trial (eligible and ineligible participants). The risk difference is 2.9% (95% confidence interval: 0.7–5.0). 4) Of these combined eligible and ineligible registrations, 0.6% (3/470) of the intervention group and 0.4% (2/467) of the control group. The risk difference is 0.2 (95% confidence interval: - 0.7–1.1).
Perspective adopted	Trial team.

Study	Free et al. (2010b)
	Host Trial
Name	Txt2stop
Design	Randomised Controlled Trial (RCT); pilot, single-blind
Location	United Kingdom
Setting	Community
Population	People aged 16 and above who are smokers and willing to stop smoking in next month.
Intervention(s)	A composite mobile phone-based smoking cessation support.
Comparator(s)	Simple, short, generic text messages.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised; single-blind, with concealed allocation.
Strategy(-ies) and study objective	Letter containing study and consent information and a £5 note.
Comparator(s)	Participants received the normal trial procedures.
Frequency of strategy	Once, before recruitment to the Txt2stop trial.
Measure(s) of benefit	1) Randomisation into the Txt2stop trial within 2 weeks. 2) Consent to be randomised into the Txt2stop trial within 2 weeks.
Costs	Incremental cost per participant randomised to the Txt2stop trial
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	Letter and £5 note: n= 245 (49.9%) Usual recruitment: n=246 (50.1%) Total sample size: n=491 (100%)
Results	Secondary: Incremental cost per participant randomised: £121.5 Other: 1) 4.5% (11/246) of participants who were sent the letter with £5 were randomized into the Txt2stop trial compared to 0.4% (1/245) of those who were not sent anything. The risk difference is 4.0% (95% confidence interval: 1.4–6.7). 2) 5.3% (13/246) of participants who were sent the letter with £5 gave their consent to be randomized into the Txt2stop trial, compared with 0.4% (1/245) of the control group. The risk difference is 4.9 (95% confidence interval: 2.0–7.7).
Perspective adopted	Trial team.

Study	Free et al. (2010c)
	Host Trial
Name	Txt2stop
Design	Randomised Controlled Trial (RCT); pilot, single-blind
Location	United Kingdom
Setting	Community
Population	People aged 16 and above who are smokers and willing to stop smoking in next month.
Intervention(s)	A composite mobile phone-based smoking cessation support.
Comparator(s)	Simple, short, generic text messages.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised; single-blind, with concealed allocation.
Strategy(-ies) and study objective	Nudge intervention, which is a series of four text messages over 1 week containing quotes from existing participants.
Comparator(s)	Participants received the normal trial procedures.
Frequency of strategy	Four times over one week.
Measure(s) of benefit	1) Randomisation into the Txt2stop trial within 2 weeks. 2) Consent to be randomised into the Txt2stop trial within 2 weeks.
Costs	Incremental cost per participant randomised to the Txt2stop trial
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	Nudge intervention: n= 406 (50%) Usual recruitment: n=405 (50%) Total sample size: n=811 (100%)
Results	Secondary: Incremental cost per participant randomised= £6.8 Unit cost of SMS: £0.05 Other: 1) 3.5% (14/405) of those who were sent the series of text messages containing quotes were randomized into the trial, and none of the 406 people in the control group were randomized into the trial. The risk difference is 3.5 (95% confidence interval: 1.7–5.2). 2) 4.2% (17/405) of those who were sent the series of text messages containing quotes gave their consent to join the trial, and none of the 406 people in the control group gave their consent to join the trial at 2 weeks. The risk difference is 4.2 (95% confidence interval: 2.2–6.1).
Perspective adopted	Trial team.

Study	Bell et al. (2016)
	Host Trial
Name	SCOOP trial
Design	Randomised Controlled Trial (RCT)
Location	United Kingdom
Setting	Primary care
Population	Females aged 70-85 who are not currently on prescription medication to prevent osteoporotic fractures before randomisation.
Intervention(s)	Screening of osteoporosis for prevention of fractures.
Comparator(s)	Usual care
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised controlled trial; two-arm, 1:1 randomisation
Strategy(-ies) and study objective	Trial-branded pen with the 60-month follow-up questionnaire, to improve the retention of host trial participants.
Comparator(s)	60-month follow-up questionnaire alone.
Frequency and timing of intervention	Once. Reminder notices were sent approximately 18 days after the initial questionnaire if no response had been received by that time. After continued nonresponse, a follow-up telephone call was administered approximately12 days after the follow-up reminder notice. After three attempts to contact participants by telephone, the participant was considered a non-responder at that timepoint. The retention period was 60 months.
Measure(s) of benefit	Questionnaire return rate. Proportion of reminders sent to both groups.
Costs	Incremental cost per participant retained. Reported unit costs of trial-branded pen.
Type of economic evaluation	Cost-effectiveness analysis
Numbers and/or proportions of participants in the intervention and control groups	Pen Group: n=3826 (50%) No Pen Group: n=3829 (50%) Total sample size: n=7655 (100%)
Results	Primary: Incremental cost per participant retained using trial-branded pen=£32.03 *If a trial-branded pen indeed induces a 1.1% improvement in retention rates, the number of participants required to be sent a pen to achieve one additional returned questionnaire relative to not being sent a pen is 91 (1/0.011590.9); therefore, the cost per additional participant retained is approximately £36 (91*£0.4). However, a trial-branded pen saved on the cost of preparing and sending reminder mailings, which is estimated at £1.97 per reminder. Consequently, approximately two fewer reminder mailings are required per retained participant reducing the cost per retained participant to an estimated £32.03.* Secondary: Unit postage cost: £0.22 Unit purchase of pen: £0.18 Unit paper and printing cost: £1.97 Unit postage cost: £0.95 Two minutes of secretarial time: £0.52 Other: Difference in retention rates: OD=1.16 (95% CI: 0.98-1.37) Retention rate (pen): 92.4% (N=3500/3789) Retention rate (no pen): 91.3% (N=3462/3793) Reminders sent (pen):22.5% (N=853/3789) Reminders sent (no pen): 24.8% (N=941/3793)
Perspective adopted	Trial team

Study	Miller et al. (1999)
	Host Trial
Name	Not provided.
Design	2 Randomised Controlled Trials (RCTs)
Location	United States
Setting	Secondary care
Population	Participants aged 18-75 with DSM-IV dysthymic disorder, double depression (major depression superimposed on antecedent dysthymia), or chronic major depression.
Intervention(s)	Psychotherapy, antidepressant medication, or both.
Comparator(s)	Psychotherapy, antidepressant medication, or both.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Quasi-randomised controlled trial.
Strategy(-ies) and study objective	Phone screening by a Senior Investigator, to see if screening by trained research assistants is more cost-effective than by senior investigators for recruitment to the two host depression trials. Following the screening, a proportion of subjects come for an interview in person, after which they might be randomised to a host depression trial.
Comparator(s)	Phone screening by a trained Research Assistant.
Frequency of strategy	Once.
Measure(s) of benefit	Proportion of participants recruited to the two host trials.
Costs	Mean phone screen length, approximate staff hourly rate, cost per phone screen, cost per positive SCID, cost per randomised subject. Incremental cost per additional randomised participant.
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible).
Numbers and/or proportions of participants in the intervention and control groups	Senior Investigator Group: n =162 (46.7%) Research Assistant Group: n= 185 (53.3%) Total sample size: n=347 (100%)
Results	Secondary: Incremental cost per positive SCID= $26.99 (=$66.71-$39.72) Incremental cost per phone screen=$7.56 (=$13.58-$6.02) Incremental staff hourly costs=$37.05(=$53.2-$16.15) Other: Recruitment rate to host trials: OR=0.59 [95% confidence interval: OR= 0.26-1.35] Recruitment rate (Senior Investigator Group): 42% (N=66/185) Recruitment rate (Research Assistant Group): 35.7% (N=68/162)
Perspective adopted	Trial team

Study	Cunningham-Burley et al. (2020)
	Host Trial
Name	SSHeW
Design	Randomised Controlled Trial (RCT); two-arm, 1:1 randomisation
Location	United Kingdom
Setting	Secondary care
Population	NHS staff who are subject to a Trust dress code.
Intervention	A pair of 5-star GRIP rated slip resistant footwear.
Comparator	No footwear during the trial.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised controlled trial; two-arm, parallel, 1:1 randomisation
Strategy(-ies) and study objective	Trial-branded pen with the 14-week follow-up questionnaire, to improve the retention of host trial participants.
Comparator(s)	14-week follow-up questionnaire alone.
Frequency and timing of intervention	Once. A reminder questionnaire was sent three weeks after the initial questionnaire if no response had been received; no additional pens were sent with reminders. The retention period was 14 weeks.
Measure(s) of benefit	Questionnaire return rate. Proportion of reminders sent to both groups.
Costs	Reported unit costs of a standard questionnaire pack and trial-branded pen.
Type of economic evaluation	Cost-effectiveness analysis
Numbers and/or proportions of participants in the intervention and control groups	Pen Group: n=733 (50%) No Pen Group: n=733 (50%) Total sample size: n=1466 (100%)
Results	Primary: Incremental cost per additional participant retained using trial-branded pen=£10 *If a trial-branded pen indeed induces a 3% improvement in retention rates, the number of participants required to be sent a pen to achieve one additional returned questionnaire relative to not being sent a pen is 33 (1/0.03); therefore, the cost per additional participant retained is approximately £10.56 (33*£0.32). However, 91 participants would need to be sent a pen to prevent one reminder mailing and therefore to save £2.42. Hence, one reminder is required per three retained participants, and the cost per retained participant is approximately £10.* Other: Difference in retention rates: OD=1.15 (95% CI: 0.92-1.43) Retention rate (pen): 67.7% (N=496/733) Retention rate (no pen): 64.7% (N=474/733) Reminders sent (pen):46.6% (N=342/733) Reminders sent (no pen): 50.9% (N=373/733)
Perspective adopted	Trial team

Study	Cochrane et al. (2020)
	Host Trial
Name	Occupational Therapist Intervention Study (OTIS)
Design	Cohort, pragmatic, two-arm, open Randomised Controlled Trial (RCT)
Location	United Kingdom
Setting	Primary care
Population	People aged 65 or older who either have experienced a fall in the last year or feel worried about falling in the future.
Intervention(s)	Occupational therapist at-home visit.
Comparator(s)	Usual healthcare from GP and a falls prevention leaflet.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised; 1:1 randomisation, stratified by OTIS allocation groups.
Strategy(-ies) and study objective	Personalised text message, as a retention intervention to encourage participants to send their post-randomisation postal questionnaire.
Comparator(s)	Generalised text message.
Frequency and timing of intervention(s)	Once or twice. The texts were sent 4 days after the postal questionnaire became available. In the event of no response, reminder letters were additionally sent to unresponsive participants. The retention period was 4 months.
Measure(s) of benefit	Questionnaire return rate.
Costs	Incremental cost per participant retained using personalised message. Mean and median unit costs of personalised texts, generalised texts and reminder letters.
Type of economic evaluation	Cost reporting
Numbers and/or proportions of participants in the intervention and control groups	Personalised texts: n=139 (49.1%) Generalised texts: n=144 (50.9%) Total sample size: n= 283 (100%)
Results	Secondary: Mean unit cost of personalised text=£0.096 Mean unit cost of generalised text=£0.048 Mean unit cost of reminder letters=£0.17 Incremental cost per participant retained using personalised message=£0.04 Other: Difference in retention rates: OD=0.64 (95% CI: 0.10-3.88) Retention rate (personalised text): 97.8% (N=136/139) Retention rate (generalised text): 98.6% (N=142/144)
Perspective adopted	Trial team

Study	Clark et al. (2015)
	Host Trial
Name	DOC Trial
Design	Randomised Controlled Trial (RCT); two-arm, 1:1 block randomisation
Location	United Kingdom
Setting	Primary care
Population	Smokers who are aged 35 or more, who are invited to undertake a series of case-finding tools, which comprise lung function tests and several symptom-based case-finding questionnaires, for the potential identification of COPD.
Intervention(s)	After recruitment, the participants receive the case-finding tests straightaway.
Comparator(s)	After recruitment, the participants receive the case-finding tests 6 months later.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT); two-arm, 1:1 simple randomisation
Strategy(-ies) and study objective	Participants receive an electronic prompt (i.e. SMS or e-mail) to return the study questionnaire. This intervention aims to improve retention to the host trial. Retention period: 2-6 months after randomisation to host trial.
Comparator(s)	Participants do not receive any electronic prompt.
Frequency and timing of intervention	One to three times. Two reminder letters were sent in an attempt to encourage response. The first reminder letter was sent 2 weeks after the follow-up questionnaire, and the second reminder was sent 2 weeks later (i.e. 4 weeks after the follow-up questionnaire). The follow-up questionnaire was sent to participants between 2 and 6 months (depending on study site) after the date of randomization. The response period was up to 2 months after the follow-up questionnaire was sent. The retention period was 2-6 months (depending on site) after randomisation.
Measure(s) of benefit	Questionnaire return rate.
Costs	Unit cost of sending automatic electronic prompts
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	Electronic prompt: n=226 (51.7%) No Electronic prompt: n=211 (48.3%) Total sample size: n=437 (100%)
Results	Secondary: Unit cost of sending automatic electronic prompts= £0.08 Other: Retention rate (Electronic prompt): 69.5% (N=157/226) Retention rate (No Electronic prompt): 60.7% (N=128/211)
Perspective adopted	Trial team

Study	Hardy et al. (2016)
	Host Trial
Name	BUMPES
Design	Randomised Controlled Trial (RCT)
Location	United Kingdom
Setting	Secondary care
Population	Adult women who are nulliparous, have a single cephalic presentation, greater than or equal to 37 weeks' gestation, intend spontaneous vaginal birth, are in second stage of labour and with an effective mobile epidural in situ.
Intervention(s)	*Intervention group 1*: Women allocated to an 'upright position' would aim to be in positions where their pelvis is in as vertical a plane as possible during the second stage of labour. *Intervention group 2*: Women allocated to a 'lying-down position' would aim to be in positions where their pelvis is in as horizontal a plane as possible during the second stage of labour.
Comparator(s)	No comparator.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT): parallel, randomisation stratified by host trial allocation and by the centre.
Strategy(-ies) and study objective	*Intervention group* 1 received an incentive cover letter sent with the first mailout of the questionnaire containing details of a promise of a £10 gift voucher (redeemable at some shops) on the return of a completed questionnaire. The covering letter included a sentence explaining that the voucher was to thank participants for their time and effort. All reminder letters included a sentence about the incentive. *Intervention group 2* received a cover letter sent at first mailout did not mention the incentive. If the questionnaire was not returned, all reminder letters detailed the promise of a £10 gift voucher on the return of a completed questionnaire.
Comparator(s)	No comparator.
Frequency and timing of intervention	Once. The retention period was 12 months.
Measure(s) of benefit	Questionnaire return rate.
Costs	Total cost of the vouchers; cost of vouchers per participant
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	Incentive cover letter: n=508 (49.5%) Incentive reminder letter n=518 (50.5%) Total sample size: n=1026 (100%)
Results	Secondary: Additional cost of vouchers per participant in incentive cover letter: £4.56 (95% confidence interval: £4.02- £5.11) Cost of vouchers per participant in first cover letter: £7.53 Cost of vouchers per participant in reminder letter: £2.97 Other: Retention rate (incentive cover letter): 74.2% (N=377/508) Retention rate (incentive reminder letter): 71.8% (N=372/518)
Perspective adopted	Trial team

Study	Gates et al. (2009)
	Host Trial
Name	Managing Injuries of the Neck Trial (MINT)
Design	Cluster Randomised Controlled Trial (RCT)
Location	United Kingdom
Setting	Secondary care
Population	Participants who attended Emergency Departments (EDs) with an acute whiplash injury of whiplash-associated disorder grades I-III were eligible for Step 1. People who attended EDs with whiplash injuries and had persistent symptoms 3 weeks after ED attendance were eligible for Step 2.
Intervention(s)	In Step 1, the intervention was a psycho-educational intervention (WBA/active management advice). In Step 2, the intervention was a package of up to six physiotherapy treatments.
Comparator(s)	Usual Care Advice (UCA)
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Quasi-randomised
Strategy(-ies) and study objective	£5 gift voucher, redeemable at a range of shops with their questionnaire, and a covering letter including a sentence explaining that the voucher is to thank participants for their time and effort. This is an intervention to potentially improve retention to the host trial. The retention period in this trial is between 4 and 8 months.
Comparator(s)	No gift voucher, and a standard covering letter.
Frequency and timing of intervention	Once. In case of no follow-up, a second copy of the questionnaire was sent after two weeks, followed by up to three attempts to make contact by telephone to request return of the questionnaire. Finally, participants were offered the option to provide the most important outcome data by telephone. The retention periods were 4 and 8 months.
Measure(s) of benefit	Questionnaire return rate.
Costs	Cost per additional questionnaire returned.
Type of economic evaluation	Cost-effectiveness analysis
Numbers and/or proportions of participants in the intervention and control groups	Incentive: n=1070 (51.7%) No incentive: n=1074 (49.9%) Total sample size: n=2144 (100%)
Results	Primary: Incremental cost per additional questionnaire returned using incentive =£67.29 Secondary: Unit costs of mailing (£1.25), follow-up phone call (£0.90), successful telephone data collection (£6.54), unsuccessful telephone data collection (£1.09) and returned questionnaire (£0.52). Other: Retention rate (incentive): 52.3% (N=560/1070) Retention rate (no incentive): 45.9% (N=493/1074)
Perspective adopted	Trial team

Study	James et al. (2020)
	Host Trial
Name	Occupational Therapist Intervention Study (OTIS)
Design	Cohort, pragmatic, two-arm, open Randomised Controlled Trial (RCT)
Location	United Kingdom
Setting	Primary care
Population	People aged 65 or older who either have experienced a fall in the last year or feel worried about falling in the future.
Intervention(s)	Occupational therapist at-home visit.
Comparator(s)	Usual healthcare from GP and a falls prevention leaflet.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT): 2x2 factorial, 1:1:1:1 simple block randomisation
Strategy(-ies) and study objective	*Intervention group 1* received a branded pen and a standard cover letter. *Intervention group 2* received a branded pen and a social incentive cover letter. *Intervention group 3* received no pen and a social incentive cover letter.
Comparator(s)	The control group received no pen and a standard cover letter.
Frequency and timing of intervention	Once. The retention period was 12 months.
Measure(s) of benefit	Questionnaire return rate.
Costs	Reported unit costs of pen, letter printing and paper, staff, postage and incentive payment cash.
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	Group 1: n=195 (25%) Group 2: n=195 (25%) Group 3: n=195 (25%) Control Group: n=194 (25%) Total sample size: n=779 (100%)
Results	Secondary: Unit pen cost: £0.19 Unit letter cost printing and paper: £0.05 Staff letter unit cost (no incentive cover letter): £0.37 Staff letter unit cost (with incentive cover letter): £0.79 Unit postage cost (inc. pen): £0.83 Unit postage cost (ex. pen): £0.61 Other: Retention rate (Group 1): 95.8% (N=187/195) Retention rate (Group 2): 95.8% (N=187/195) Retention rate (Group 3): 95.8% (N=187/195) Retention rate (Group 4): 95.8% (N=186/194)
Perspective adopted	Trial team

Study	Kenyon et al. (2005)
	Host Trial
Name	MRC ORACLE Children Study
Design	Cohort, pragmatic, two-arm, open Randomised Controlled Trial (RCT)
Location	United Kingdom
Setting	Primary care
Population	Children whose mothers joined the MRC ORACLE Trial. Their mothers have had preterm, prelabour rupture of the fetal membranes (pROM).
Intervention(s)	Antibiotics to generate health benefits for the neonate, and hence to reduce childhood disability (from MRC ORACLE Trial)
Comparator(s)	Placebo drug
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT): parallel, simple randomisation
Strategy(-ies) and study objective	The parents of the survived children associated with the MRC ORACLE Children Study receive a monetary incentive (£5 voucher redeemable at high street stores) together with the reminder questionnaire associated with their child(ren)'s health.
Comparator(s)	No monetary incentive. The same reminder questionnaire associated with child(ren)'s health was sent.
Frequency and timing of intervention	Once. The retention time was up to 6 weeks after the follow-up questionnaire was sent. The retention period was 84 months after the original trial.
Measure(s) of benefit	Questionnaire return rate.
Costs	Cost per additional questionnaire returned
Type of economic evaluation	Cost-effectiveness analysis
Numbers and/or proportions of participants in the intervention and control groups	Monetary Incentive: n=369 (51.1%) No incentive: n=353 (48.9%) Total sample size: n=722 (100%)
Results	Primary: Cost per additional questionnaire returned: £67 (assuming a 3% retention rate) Other: Retention rate (Monetary Incentive): 42.3% (N=156/369) Retention rate (No incentive): 30.6% (N=108/353)
Perspective adopted	Trial team

Study	Khadjesari et al. (2011a)
	Host Trial
Name	Down Your Drink (DYD) Study
Design	Randomised Controlled Trial; 2-arm, randomisation stratified by age and gender
Location	United Kingdom
Setting	Secondary care
Population	People who visited DownYourDrink while browsing the web, and who had an AUDIT-C score greater than 5.
Intervention(s)	Online DYD psychologically enhanced intervention to reduce drinking problems
Comparator(s)	Similarly designed website with no psychologically enhanced intervention.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT): parallel, randomisation stratified by DYD group
Strategy(-ies) and study objective	Participants who did not complete the first follow-up questionnaire within 1 week received either a £5 Amazon.co.uk voucher, £5 donation to Cancer Research UK, or entry in a £250 prize draw in the second prompt for completion of questionnaires.
Comparator(s)	Participants who did not complete the first follow-up questionnaire within 1 week did not receive any incentive. They only received another prompt for completion of questionnaires.
Frequency and timing of intervention	At least once. In case the participants did not return any completed questionnaires within 1 week of their allocated interventions, they remained in the same trial arm and received a final prompt for the completion of the questionnaires. The retention period was 3 months after the original trial. The retention time was six weeks after an initial questionnaire was sent in the follow-up study.
Measure(s) of benefit	Questionnaire return rate.
Costs	Cost per additional questionnaire returned. Unit costs associated with the interventions and the trial capital resources.
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible).
Numbers and/or proportions of participants in the intervention and control groups	£5 Amazon voucher: n=206 (51.1%) £5 Charity donation: n=204 (48.9%) £250 prize draw: n=205 (100%) Incentives (collectively): n= 615 No incentive: n=611 Total sample size: n=1226 (100%)
Results	Secondary: Cost per additional questionnaire returned: £110.15 Unit cost of setting up the database: £0.67 Unit cost of time sending confirmatory incentive email (per response to questionnaires): £0.95 Other: Retention rate (Incentives (collectively): 29% (N=175/615) Retention rate (No incentive): 27% (N=162/611) Difference in Retention rate (incentive vs no incentive): 2% (95% CI: -3% to 7%) Retention rate (£5 Amazon voucher): 32% (N=66/206) Retention rate (£5 Charity donation): 27% (N=55/204) Retention rate (£250 prize draw): 26% (N=54/205)
Perspective adopted	Trial team

Study	Khadjesari et al. (2011b)
	Host Trial
Name	Down Your Drink (DYD) Study
Design	Randomised Controlled Trial; 2-arm, randomisation stratified by age and gender
Location	United Kingdom
Setting	Secondary care
Population	People who visited DownYourDrink while browsing the web, and who had an AUDIT-C score greater than 5.
Intervention(s)	Online DYD psychologically enhanced intervention to reduce drinking problems
Comparator(s)	Comparator website with no psychologically enhanced intervention.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT): parallel, randomisation stratified by DYD group
Strategy(-ies) and study objective	Participants received a £10 Amazon.co.uk voucher in the first prompt for completion of questionnaires.
Comparator(s)	Participants did not receive any incentive. They only received another prompt for completion of questionnaires.
Frequency and timing of intervention	At least once. In case the participants did not return any completed questionnaires within 1 week of their allocated interventions, they remained in the same trial arm and received a second prompt for the completion of the questionnaires. In case the participants did not return any completed questionnaires within 1 week of the second prompt, they remained in the same trial arm and received a third prompt for the completion of the questionnaires. The retention period was 12 months after the original trial. The retention time was six weeks after an initial questionnaire was sent in the follow-up study.
Measure(s) of benefit	Questionnaire return rate.
Costs	Cost per additional questionnaire returned. Unit costs associated with the interventions and the trial capital resources.
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible).
Numbers and/or proportions of participants in the intervention and control groups	£10 Amazon voucher: n=1296 (50%) No incentive: n=1295(50%) Total sample size: n=2591 (50%)
Results	Secondary: Cost per additional questionnaire returned: £52 Unit cost of setting up the database: £0.46 Unit cost of time sending confirmatory incentive email (per response to questionnaires): £0.95 Other: Retention rate (£10 Amazon voucher): 37% (N=476/1296) Retention rate (No incentive): 28% (N=364/1295) Difference in Retention rate (£10 Amazon voucher vs no incentive): 9% (95% CI: 5% to 12%)
Perspective adopted	Trial team

Study	Marsh et al. (1999)
	Host Trial
Name	Not provided.
Design	Cluster Randomised Controlled Trial (RCT)
Location	United Kingdom
Setting	Primary care
Population	All children aged 3-12 months registered with 36 participating practices in Nottingham.
Intervention(s)	A package of safety advice at child health surveillance consultations at 6-9, 12-15, and 18-24 months; provision of low-cost safety equipment to families on means tested state benefits; and home safety checks and first aid training by health visitors. The aim was to reduce the frequency and severity of medically attended injuries.
Comparator(s)	Low-cost equipment was provided by the health visitor for families receiving means tested state benefits. Equipment comprised stair gates and fireguards, cupboard locks, and smoke alarms (50p each).
SWAT (embedded trial)	
Design (randomised or quasi-randomised)	Quasi randomised; cluster randomised.
Strategy(-ies) and study objective	*Intervention group 1* received postal administration with financial incentive (£2 voucher to spend in a local children's store) once the completed diary had been received or postal group without financial incentive. *Intervention group 2* received telephone administration with financial incentive (£2 voucher to spend in a local children's store) once the completed diary had been received or telephone group without financial incentive. Participants included parents of children aged 3-12 months registered with the practices participating in the main trial.
Comparator(s)	Control group were selected from four practices and their matched control practices were selected for the clinic visits.
Frequency and timing of intervention	The retention period is unclear.
Measure(s) of benefit	Return rate of diaries.
Costs	Average cost per returned diary. Unit fixed costs are also reported.
Type of economic evaluation	Cost reporting
Numbers and/or proportions of participants in the intervention and control groups	Group 1: n=204 (47%) Group 2: n=130 (30%) Control Group: n=100 (23%) Total sample size: n=434 (100%)
Results	Secondary: Average cost per returned diary without financial incentive (Group 1): £1.73 Average cost per returned diary with financial incentive (Group 1): £3.54 Average cost per returned diary without financial incentive (Group 2): £14.98 Average cost per returned diary with financial incentive (Group 2): £10.23 Average cost per returned diary without financial incentive (Control Group): £38.48 Average cost per returned diary with financial incentive (Control Group): £35.24 Unit cost of postage: £0.2 Unit cost of telephone calls: £1 (for an average call of 10 minutes) Unit travel cost: £0.41/mile Other: Retention rate (Group 1/ incentive): 61% (N= 61/102) Retention rate (Group 1/ no incentive): 54% (N=55/102) Retention rate (Group 2/ incentive): 69% (N=45/65) Retention rate (Group 2/ no incentive): 45% (N=29/65) Retention rate (Group 3/ incentive): 44% (N=22/50) Retention rate (Group 3/ no incentive): 38% (N=19/50)
Perspective adopted	Trial team

Study	Treweek et al. (2021)
	Host Trial
Name	ActWELL
Design	Randomised Controlled Trial (RCT); four-centre, 1:1 parallel group
Location	United Kingdom
Setting	Primary care
Population	Women aged 50-70 who are overweight and attending routine breast screening in four Scottish breast screening service centres.
Intervention(s)	Two face-to-face visits with a lifestyle coach and a further 9 phone calls over 12 months. The participants will be given a diet and physical activity programme with the aim of weight management and change in physical activity. This will be delivered in the community by Breast Cancer Now volunteer Lifestyle Coaches.
Comparator(s)	Usual care.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT); two-arm, 1:1 parallel group, stratified by service centre
Strategy(-ies) and study objective	A pre-notification card was sent around one month before the face-to-face primary outcome measurement visit. The text on the card was not developed using formal behavioural change theory but did target factors thought to influence attendance.
Comparator(s)	No pre-notification card.
Frequency and timing of intervention	Once. The retention period is 12 months.
Measure(s) of benefit	The difference in the proportion of participants attending the host trial primary outcome measurement visit (i.e. retention).
Costs	Detailed unit costs associated with the pre-notification card.
Type of economic evaluation	Cost per additional questionnaire returned: £110.15
Numbers and/or proportions of participants in the intervention and control groups	Pre-notification card: n=274(49.1%) No card: n=284(50.9%) Total sample size: n=558 (50%)
Results	Secondary: Direct costs of printing the cards: £72 Second class (i.e. delivery within two days) postage costs: £120 Total costs of cards: £192 Other: Retention rate (pre-notification card): 84% (231/274) Retention rate (no card): 81% (230/284) Risk difference in retention rates: 3.3% (95% CI: -3% to 9.6%)
Perspective adopted	Trial team

Study	Whiteside et al. (2019)
	Host Trial
Name	Occupational Therapist Intervention Study (OTIS)
Design	Cohort, pragmatic, two-arm, open Randomised Controlled Trial (RCT)
Location	United Kingdom
Setting	Community setting
Population	People aged 65 or older who either have experienced a fall in the last year or feel worried about falling in the future.
Intervention(s)	Occupational therapist at-home visit.
Comparator(s)	Usual healthcare from GP and a falls prevention leaflet.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT); two-arm, 2:1 block randomisation stratified by GP practice.
Strategy(-ies) and study objective	Branded pen with trial invitation pack.
Comparator(s)	No pen in the trial invitation pack.
Frequency and timing of intervention	Once, before randomisation to the host trial. The retention period was related to those remaining in the host trial at 3-months.
Measure(s) of benefit	Randomisation rate, i.e. recruitment. Proportion of participants who remained in the trial at 3 months post randomisation, i.e. retention.
Costs	Cost per participant recruited in both groups.
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	Branded pen: n=620 (33.3%) No pen: n=1242 (66.7%) Total sample size: n= 1862 (100%)
Results	Secondary: Unit cost of pens: £0.32 Unit cost of invitation packs (inc. printing, packaging and postage costs): £2.53 Other: Recruitment rate (pen): 4.5% (N=28/620) Recruitment rate (no pen): 4.3% (N=54/1242) Difference in recruitment rates (pen vs no pen): 0.17% (95% CI: -1.82% to 2.16%) Retention rate (pen):4.4% (N=27/620) Retention rate (no pen): 3.9% (N=49/1242)
Perspective adopted	Trial team

Study	Jolly et al. (2019)
	Host Trial
Name	PSM COPD study
Design	Pragmatic, multicentre Randomised Controlled Trial (RCT)
Location	United Kingdom
Setting	Primary care
Population	People aged 18 or older who are on the practice COPD register and have mild dyspnoea.
Intervention(s)	Telephone health coaching to support self-management.
Comparator(s)	Usual care.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Cluster Randomised Controlled Trial (RCT): pragmatic, 1:1 block randomisation stratified by area
Strategy(-ies) and study objective	The practices recruiting participants for the host trial accessed standard printed patient information materials, as well as a multimedia information resource that was developed by patient and public involvement (PPI) contributors and researchers. The resource includes both study-specific information and generic information about trials.
Comparator(s)	The practices recruiting participants for the host trial accessed standard printed patient information materials, with no extra multimedia information resource.
Frequency and timing of intervention	Once randomised, the practices sent out letters to the associated COPD patients. The retention period was 6 and 12 months.
Measure(s) of benefit	Response rates to invitation letter. Randomisation rates to the host trial, i.e. recruitment rate. Retention at 6 months' follow-up Retention at 12 months' follow-up
Costs	Incremental cost per participant recruited. Incremental cost per participant retained.
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	PPI intervention: n=2280 (54.1%) Standard material: n=1934 (45.9%) Total sample size: n=4214 (100%)
Results	Secondary: Average unit cost of multimedia resource: £0.59 (=£2500/4223) Other: Recruitment rate (PPI intervention): 10.8% (N=247/2280) Recruitment rate (standard material): 9.6% (N=185/1934) Retention rate, 6 months (PPI intervention): 10.1%(N=231/2280) Retention rate, 6 months (standard material): 8.8 %(N=171/1934) Retention rate, 12 months (PPI intervention): 9.8 %(N=223/2280) Retention rate, 12 months (standard material): 8.2% (N= 159/1934)
Perspective adopted	Trial team

Study	Bracken et al. (2019)
	Host Trial
Name	Testosterone for Diabetes Mellitus (T4DM) trial
Design	Randomised Controlled Trial (RCT): multicentre, double-blind
Location	Australia
Setting	Primary care
Population	Men aged 50-74 years, obese or overweight, with prediabetes or newly diagnosed type 2 diabetes, and a low serum testosterone.
Intervention(s)	Intramuscular injection of testosterone undecanoate (1000 mg), 6 weeks, and then every 3 months for 2 years.
Comparator(s)	Placebo injection
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT); two-arm, parallel, 1:1 block randomisation stratified by age group
Strategy(-ies) and study objective	Individuals who were eligible on a pre-screening questionnaire who did not attend a further screening assessment within 4 weeks were randomized to receive an SMS reminder. The SMS reminder text provided key enrolment information as well as including a peripheral cue based on the concept of social proof.
Comparator(s)	Individuals who were eligible on a pre-screening questionnaire who did not attend a further screening assessment within 4 weeks were randomized to receive a telephone reminder. Staff members were provided with a reminder call script, so that the content of the calls was as similar as possible across participants.
Frequency of strategy	The reminder was received 4 weeks after completing the T4DM pre-screening questionnaire if the participants had not attended laboratory screening within 4 weeks of pre-screening. The SMS reminder was sent within 2 days of randomisation and the telephone reminder was conducted within 4 days of randomisation. The SMS reminder intervention was delivered once, whereas the telephone reminders could occur once or twice, dependent upon whether the participants responded to the phone call. If a participant could not be reached on the second attempt, a voicemail message was left, if possible. The recruitment period was 4 weeks after the reminders were received.
Measure(s) of benefit	Attendance rates for laboratory screening
Costs	Reported costs of each intervention. Incremental cost-effectiveness ratio of phone call reminders
Type of economic evaluation	Cost-effectiveness analysis.
Numbers and/or proportions of participants in the intervention and control groups	SMS reminder: 354 (49.9%) Phone reminder:355 (50.1%) Total sample size:709 (100%)
Results	Primary: Incremental cost-effectiveness ratio of phone call reminder: AU$112.05 Incremental cost-effectiveness ratio of phone call reminder (65-74 yrs): AU$31.45 Secondary: Unit cost of telephone reminder: AU$6.21 Unit cost of SMS reminder: AU$0.53 Cost of staff time per SMS reminder: AU$0.35 Cost of staff time per telephone reminder: AU$5.67 Other: Attendance rate (telephone reminder): 23% (N=82/355) Attendance rate (SMS reminder):18% (N=64/354) Difference in response to telephone vs SMS reminders: RR=1.29 (95% CI: 0.96-1.73)
Perspective adopted	Trial team

Study	Arundel et al. (2017)
	Host Trial
Name	REFORM
Design	Cohort Randomised Controlled Trial (cRCT); two-arm, pragmatic, open, multicentre
Location	United Kingdom, Ireland
Setting	Primary care and secondary care
Population	Patients aged 65 years and over who have attended routine podiatry services within the past 6 months, have had one fall in the past 12 months; or one fall in the past 24 months requiring hospital attention; or report a fear of falling on their baseline questionnaire that is, have worried about falling at least some of the time, in the past 4 weeks.
Intervention(s)	Routine podiatry care, a falls prevention leaflet and a multifaceted podiatry intervention. The intervention aims to prevent falls.
Comparator(s)	Routine podiatry care and a falls prevention leaflet.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT); 1:1 randomisation
Strategy(-ies) and study objective	A pre-notification leaflet, 2–3 weeks before the trial recruitment pack.
Comparator(s)	Trial recruitment pack only.
Frequency of strategy	Once.
Measure(s) of benefit	Proportion of participants randomized for the host trial.
Costs	Design, printing, postage and staff costs associated with the leaflet and the trial recruitment pack.
Type of economic evaluation	Cost-effectiveness analysis.
Numbers and/or proportions of participants in the intervention and control groups	Leaflet: n=1436 (33.3%) No leaflet: n=2878 (66.7%) Total sample size: n=4314 (100%)
Results	Primary: Incremental cost-effectiveness ratio (ICER): £224.29 Secondary: Unit cost of leaflet: £4.36 Unit cost for control group: £2.79 Cost per recruited participant (no leaflet): £63 (=£2.79* 100/4.4) Cost per recruited participant (with leaflet): £85.5 (=£4.36* 100/5.1) Total design cost: £16/h *30 hours =£480 Unit printing cost: £0.26 Total packing cost: 16/h *8 hours =£128 Total supply cost (i.e. envelopes and second-class stamps): £956.35 Total labelling cost: £19.38*16 hours= £310 Total incremental cost of using a pre-notification leaflet= £2251.30 Other: Recruitment rate (with leaflet):5.1% (N=73/1436) Recruitment rate (no leaflet): 4.4% (N=126/2878) Conclusion: Compared with the cost of randomizing an additional control group participant into REFORM (£63), the intervention was three and a half times more expensive than the control, per additional randomization obtained. Therefore, the intervention is not cost-effective at the point estimate difference, and for the intervention to be cost-effective, it would need to increase the recruitment rate by nearly 2.5% (i.e. £1.57/£63 × 100).
Perspective adopted	Trial team

Study	Rogers et al. (2019)
	Host Trial
Name	Febuxostat versus Allopurinol Streamlined Trial (FAST)
Design	Randomised Controlled Trial (RCT): prospective, open-label, blinded.
Location	United Kingdom
Setting	Primary care
Population	People aged 60 or over, taking allopurinol for chronic gout, and with additional cardiovascular risk factors.
Intervention(s)	After an allopurinol lead-in phase where the dose of allopurinol is optimised to achieve European League against Rheumatism (EULAR) urate targets (serum urate <357 µmol/L), patients continue with optimal dose of allopurinol
Comparator(s)	Participants receive febuxostat.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT); one-blinded, 1:1 randomisation.
Strategy(-ies) and study objective	DVD presentation containing an audio-visual presentation explaining the background and operation of FAST, and a standard invitation pack.
Comparator(s)	Standard invitation pack only.
Frequency of strategy	Once.
Measure(s) of benefit	Randomisation rate to host trial.
Costs	Incremental costs of sending a DVS invitation pack, and incremental costs per randomised participant. Cost breakdown available.
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	DVD and invitation pack: n=498 (48.4%) Invitation pack only: n=531 (51.6%) Total sample size: n=1029 (100%)
Results	Secondary: Incremental cost of sending a DVD invitation pack: £31.48 Incremental cost of sending a DVD invitation pack per recruited participant: £138.32 Unit video production costs: £29.94 Unit DVD manufacture and packaging costs: £0.9 Unit postage costs: £0.64 Other: Randomisation rate (DVD and invitation pack): 22.8% (N=114/498) Randomisation rate (Invitation pack only): 24.3% (N=129/531)
Perspective adopted	Trial team

Study	Hancocks et al. (2019)
	Host Trial
Name	Trial of physical Activity-assisted Reduction of Smoking (TARS) study
Design	Randomised Controlled Trial (RCT): pragmatic, multicentred, parallel, two group.
Location	United Kingdom
Setting	Primary care
Population	People aged 18 or over, who are smokers and smoke at least 10 cigarettes per day (for at least one year).
Intervention(s)	A tailored individual health trainer face-to-face and/or telephone support to reduce smoking and increase PA as an aid to smoking reduction.
Comparator(s)	A brief written/electronic advice to reduce or quit smoking.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT), involving six GP practices; 1:1:1 randomisation
Strategy(-ies) and study objective	Group 1: Full invitation pack from a GP. Group 2: Single-page invitation from a GP. These interventions were designed to boost recruitment to the host trial.
Comparator(s)	Text message invitation
Frequency of strategy	Once.
Measure(s) of benefit	Recruitment rate to host trial.
Costs	Invitation costs per recruited participant associated with each intervention. Unit costs of each intervention.
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	Full invitation invitation: n=459 (36.2%) Single-page invitation: n=459 (36.2%) Text message invitation: n= 349 (27.6%) Total sample size: n=1267 (100%)
Results	Secondary: Cost per recruited participant (full invitation pack): £35.31 Cost per recruited participant (single-page pack): £36.06 Unit cost of full invitation pack: £1 Unit cost of single-page invitation: £0.55 Unit cost of text message invitation: £0 Other: Recruitment rate (full invitation pack): 2.8% (N=13/459) Recruitment rate (single-page invitation): 1.5% (N=7/459) Recruitment rate (text message invitation): 0.3% (N=1/349)
Perspective adopted	Trial team

Study	Cook et al. (2021)
	Host Trial
Name	Antivirals for influenza-Like Illness? An rCt of Clinical and Cost effectiveness in primary Care (ALIC^{4}E)
Design	Randomised Controlled Trial (RCT): pragmatic, multicountry, adaptive, two group, phase IV
Location	15 European countries
Setting	Primary care
Population	People aged 1 or over, who have sudden onset of self-reported fever, with at least one respiratory symptom (cough, sore throat, running or congested nose) and one systemic symptom (headache, muscle ache, sweats or chills or tiredness) during a period of increased influenza activity.
Intervention(s)	Antiviral treatment and usual care.
Comparator(s)	Usual care.
	SWAT (embedded trial)
Design (randomised or quasi-randomised)	Randomised Controlled Trial (RCT); matched pair cluster randomised parallel group design by trial sites.
Strategy(-ies) and study objective	Unconditional monetary incentive of £20 given to participants at recruitment, as an intervention to boost retention in the host trial.
Comparator(s)	Conditional monetary incentive of £20 given to participants only once a questionnaire had been returned.
Frequency and timing of intervention	Once. The total retention period is 28 days.
Measure(s) of benefit	Diary return rate, i.e. retention rate.
Costs	Unit direct and indirect costs, cost per diary received in the two arms
Type of economic evaluation	Cost reporting.
Numbers and/or proportions of participants in the intervention and control groups	Unconditional monetary incentive: n=220 (63.8%) Conditional monetary incentive: n=125 (36.2%) Total sample size: n=345 (100%)
Results	Secondary: Cost per diary received (Unconditional monetary incentive): £44.85 Cost per diary received (Conditional monetary incentive): £21.75 Direct cost per diary received (Unconditional monetary incentive): £33.85 Direct cost per diary received (Conditional monetary incentive): £20 Indirect cost per diary received (Unconditional monetary incentive): £11 Indirect cost per diary received (Conditional monetary incentive): £1.75 Other: Retention rate (Unconditional monetary incentive): 58% (N=127/220) Retention rate (Conditional monetary incentive): 73% (N=91/125)
Perspective adopted	Trial team

Study	Dorling et al. (2020)
	Host Trial
Name	The Speed of Increasing milk Feeds Trial (SIFT)
Design	Randomised Controlled Trial (RCT): parallel, multicentre, two group
Location	United Kingdom and Ireland
Setting	Secondary care
Population	Infants born at <32 weeks' gestation or a weight of <1500g, who were receiving <30 ml/kg/day of milk at trial enrolment.
Intervention(s)	Daily feed increments of 30 ml/kg/day to improve survival/reduce the risk of neurodevelopmental disability.
Comparator(s)	Daily feed increments of 18 ml/kg/day
SWAT (embedded trial)	
Design (randomised or quasi-randomised)	Randomised Controlled Trial; 1:1 permuted block randomisation, stratified by original SIFT allocation
Strategy(-ies) and study objective	After group. The first paper follow-up letter to parents would include a promise of an incentive (£15 (€15 for Irish residents) gift voucher redeemable at high-street shops) after receipt of a completed form, to improve questionnaire response return rate, i.e. to improve retention.
Comparator(s)	Before group. The first paper letter to parents would enclose the incentive (£15 (€15 for Irish residents) gift voucher redeemable at high-street shops) before the receipt of a completed form.
Frequency and timing of intervention	Once. The retention period was 24 months. It is uncertain how frequently the parents were contacted via text and/or e-mail to give reminders during the follow-up.
Measure(s) of benefit	Rate of questionnaire return
Costs	Postage, receipt of material via prepaid Freepost, cost of envelopes, Supplemental Materials (e.g. sticker sets sent with questionnaires for infants to play with) and value of gift vouchers. It did not include Freepost licence fee, printing, telephone calls and trial staff time. All costs for participants were calculated in GBP.
Type of economic evaluation	Cost reporting (cost-effectiveness analysis feasible)
Numbers and/or proportions of participants in the intervention and control groups	After group: n=464 (50.3%) Before group: n=459 (49.7%) Total sample size: n=923 (100%)
Results	Secondary: Incremental cost per increase in response per infant (before group): £1.35 Incremental cost per increase in response per infant (after group): £2.95 Incremental cost per infant (before group): £14.8 Incremental cost per infant (after group): £11.84 Mean cost of the incentive strategy per infant (before group): £17.97 Mean cost of the incentive strategy per infant (after group): £15 Other: Questionnaire return rate (before group): 83% (N=381/459) Questionnaire return rate (after group): 76.1% (N=353/464)
Perspective adopted	Trial team

Supplemental Material 4.2. Search Strategy

Searches were conducted for recruitment-related interventions from 12 February 2015 until 3 March 2021, and for retention-related interventions from 1 March 2020 until 3 March 2021. The following search strategy was applied to Ovid (MEDLINE), Ovid (PsycInfo) and OVID (Embase) up to March 3rd.

1. ((minimi* or prevent* or lessen* or decreas* or reduc*) adj2 (attrition or drop*-out* or dropout* or withdr*w* or missing data)).ab,ti.
2. ((increas* or encourag* or maximi* or promot* or improv*) adj2 (retention or follow-up or followup or completion or data collection or data return)).ab,ti.
3. ((strateg* or intervention* or method* or technique*) adj3 (retention or attrition or drop*-out* or dropout* or follow-up or followup)).ab,ti.
4. (Complian* adj2 (follow-up or followup)).ab,ti.
5. ((loss or lost) adj2 (follow-up or followup)).ab,ti.
6. ((difficult* or problem* or challeng* or success* or feasibl*) adj3 (retain* or retention)).ab,ti.
7. (retention adj2 rate*).ab,ti.
8. (attrition adj2 rate*).ab,ti.
9. ((Dropout* or Drop-out*) adj2 rate*).ab,ti.
10. (Completion adj2 rate*).ab,ti.
11. ((Follow-up or followup) adj2 rate*).ab,ti.
12. (Incomplete adj2 (follow-up or followup)).ab,ti.
13. (questionnaire* adj3 (response* adj2 method*)).ab,ti.
14. (questionnaire* adj3 (response adj2 technique*)).ab,ti.
15. questionnaire response rate*.ab,ti.
16. ((Strateg* or increas* or encourag* or maximi* or promot* or improv* or influenc* or success*) adj2 (questionnaire* adj3 response*)).ab,ti.
17. ((incentiv* or reminder*) adj3 (retention or retain or respon*e*)).ab,ti.
18. (retention adj4 training).ab,ti.
19. (Trial site adj2 (retention or retain*)).ab,ti.
20. exp "Lost to Follow-Up"/
21. exp Patient Dropouts/
22. (Patient retention or Dropout* or Drop*-out* or attrition).kw.
23. ((survey* or questionnaire*) and (respon*e* or return* or rate*)).ti.
24. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23
25. Randomized controlled trial.pt.
26. Controlled clinical trial.pt.
27. Randomi*ed.tw.
28. Placebo.tw.
29. Clinical trials as topic.sh.
30. Randomly.tw.
31. Trial*.tw.
32. 25 or 26 or 27 or 28 or 29 or 30 or 31
33. 24 and 32
34. limit 33 to english language
35. exp animals/ not humans.sh.
36. 34 not 35
37. limit 36 to comment, editorial, news and letter.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonym

38. 36 not 37
39. ((Stud* or trial*) adj (within or in) adj4 (trial* or stud*)).ab,ti.
40. ((Contain* or incorporat* or includ* or nest* or embed*) adj4 RCT).ab,ti.
41. ((Contain* or incorporat* or includ* or nest* or embed*) adj4 (trial* or stud*)).ab,ti.
42. ((incorporat* or contain* or embed* or nest* or includ*) adj1 (within or in or into)).ab,ti.

43. substud*.ab,ti.
44. sub-stud*.ab,ti.
45. SWAT.ab,ti.
46. SWATs.ab,ti.
47. 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46
48. 47 and 38
49. limit 38 to yr="2020 - 2022"
50. limit 48 to yr="2020 - 2022"
51. Patient Selection/
52. ((participat* or recruit* or enrol*) adj4 trial?).ab,ti.
53. 51 or 52
54. Randomized controlled trial.pt.
55. Controlled clinical trial.pt.
56. Randomi*ed.tw.
57. Placebo.tw.
58. Clinical trials as topic.sh.
59. Randomly.tw.
60. Trial*.tw.
61. 54 or 55 or 56 or 57 or 58 or 59 or 60
62. 53 and 61
63. limit 62 to english language
64. exp animals/ not humans.sh.
65. 63 not 64
66. limit 65 to comment, editorial, news and letter.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

67. 65 not 66
68. ((Stud* or trial*) adj (within or in) adj4 (trial* or stud*)).ab,ti.
69. ((Contain* or incorporat* or includ* or nest* or embed*) adj4 RCT).ab,ti.
70. ((Contain* or incorporat* or includ* or nest* or embed*) adj4 (trial* or stud*)).ab,ti.
71. ((incorporat* or contain* or embed* or nest* or includ*) adj1 (within or in or into)).ab,ti.
72. substud*.ab,ti.
73. sub-stud*.ab,ti.
74. SWAT.ab,ti.
75. SWATs.ab,ti.
76. 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75
77. 67 and 76
78. limit 67 to yr="2015 - 2022"
79. limit 77 to yr="2015 - 2022"
80. 79 or 50
81. 78 or 49
82. 81 not 80

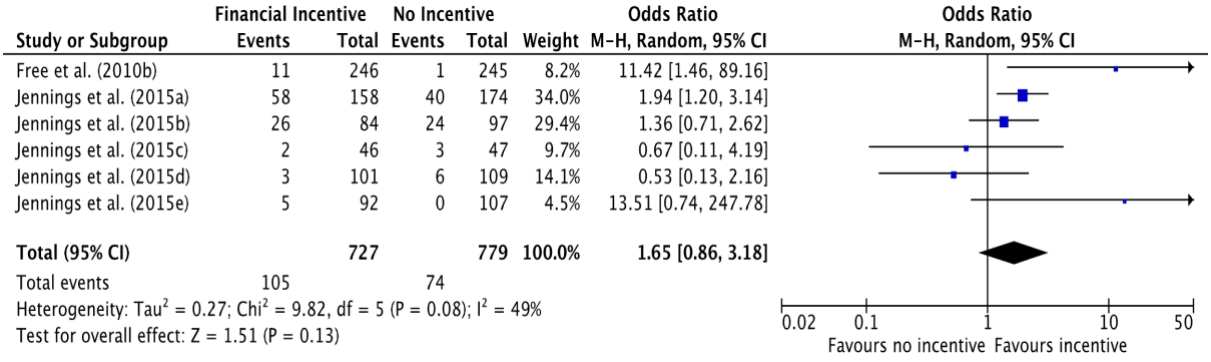
Supplemental Material 4.3. Meta-analysis tables and figures of recruitment strategies

Strategy: Financial incentive

Table 4.S1: Meta-analysis; recruitment strategy (Financial incentive versus no financial incentive)

Study or SWAT	Jennings et al. (2015a)	Jennings et al. (2015b)	Jennings et al. (2015c)	Jennings et al. (2015d)	Jennings et al. (2015e)	Free et al. (2010b)	Overall
Number of participants recruited to the host trial in the intervention group	58/158	26/84	2/46	3/101	5/92	11/246	105/727
Number of participants recruited to the host trial in the control group	40/174	24/97	3/47	6/109	0/107	1/245	74/779
Incremental unit cost of the strategy (2019 $ PPP)	$144.72	$144.72	$144.72	$144.72	$144.72	$7.15	$133.44
Total sample size (n=)	332	181	93	210	199	491	1506
Odds ratio (total)	1.65 (0.86,3.18)						
Effect size (total), i.e. incremental recruitment rate (total)	0.28(-0.08,0.64)						
Incremental cost per additional participant recruited (ICER)	**$476.57 ($ 208.50, N/A)**						
GRADE certainty of evidence	Moderate						
I^2 Statistic	49%						

Figure 4.S1: Meta-analysis; recruitment strategy (Financial incentive versus no financial incentive)

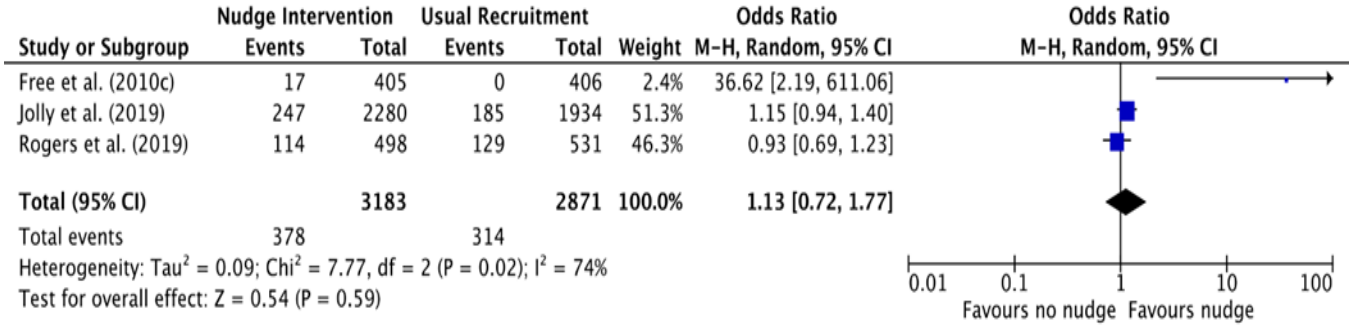


Strategy: Nudge intervention

Table 4.S2: Meta-analysis; recruitment strategy (Nudge intervention versus usual recruitment)

Study or SWAT	**Free et al. (2010c)**	**Jolly et al. (2019)**	**Rogers et al. (2019)**	**Overall**
Number of participants recruited to the host trial in the intervention group	17/405	247/2280	114/498	378/3183(11.86%)
Number of participants recruited to the host trial in the control group	0/406	185/1934	129/531	314/2871(10.94%)
Incremental unit cost of the strategy (2019 $ PPP)	$0.34	$1.74	$45.76	$22.00
Total sample size (n=)	811	4214	1029	6054
Odds ratio (total)	1.13(0.72,1.77)			
Effect size (total), i.e. incremental recruitment rate (total)	0.07(-0.18, 0.32)			
Incremental cost per additional participant recruited (ICER)	**$314.29 ($68.75, N/A)**			
GRADE certainty of evidence	Very low			
I^2 Statistic	74%			

Figure 4.S2: Meta-analysis; recruitment strategy (Nudge intervention versus usual recruitment)



Recruitment strategies consisting of a single SWAT

Table 4.S3: Meta-analyses of recruitment strategies consisting of a single SWAT

Study or SWAT	**Miller et al. (1999)**	**Free et al. (2010a)**	**Arundel et al. (2017)**	**Bracken et al. (2019)**	**Hancocks et al. (2019)**	**Whiteside et al. (2019)**
Strategy	**Screening for the host trial undertaken by a senior investigator vs screening undertaken by a research assistant**	**Primary text message vs primary call and no text message**	**Pre-notification leaflet vs no leaflet**	**Telephone reminder vs SMS reminder**	**Invitation packs from a GP vs no invitation packs**	**Branded pen vs no pen**
Number of participants recruited to the host trial in the intervention group	22/185	17/470	73/1436	80/355	20/918	28/620
Number of participants recruited to the host trial in the control group	28/162	5/467	126/2878	62/354	1/349	54/1242
Incremental unit cost of the strategy (2019 $ PPP)	$37.05	$3.04	$2.25	$3.98	$1.13	$0.47
Total sample size (n=)	347	937	4314	709	1267	1862
Odds ratio	0.19 (0.11,0.32)	3.47 (1.27,9.48)	1.17 (0.87, 1.57)	1.37 (0.95,1.98)	7.75 (1.04,57.97)	1.04(0.65,1.66)
Effect size (total), i.e. incremental recruitment rate (total)	-0.92 (-1.22,-0.63)	0.69 (0.13,1.24)	0.09 (-0.08 , 0.25)	0.17 (-0.03, 0.38)	1.13 (0.02- ,2.24)	0.02 (-0.24,0.28)
Incremental cost per additional participant recruited (ICER)	**N/A (ineffective)**	**$4.41 ($2.45, $23.38)**	**$25.97 ($9.00,N/A)**	**$23.37 ($10.47, N/A)**	**$1.00 ($0.50,$57.47)**	**$21.41 ($1.68 ,N/A)**
GRADE certainty of evidence	Low	Moderate	Moderate	Moderate	Very Low	Moderate

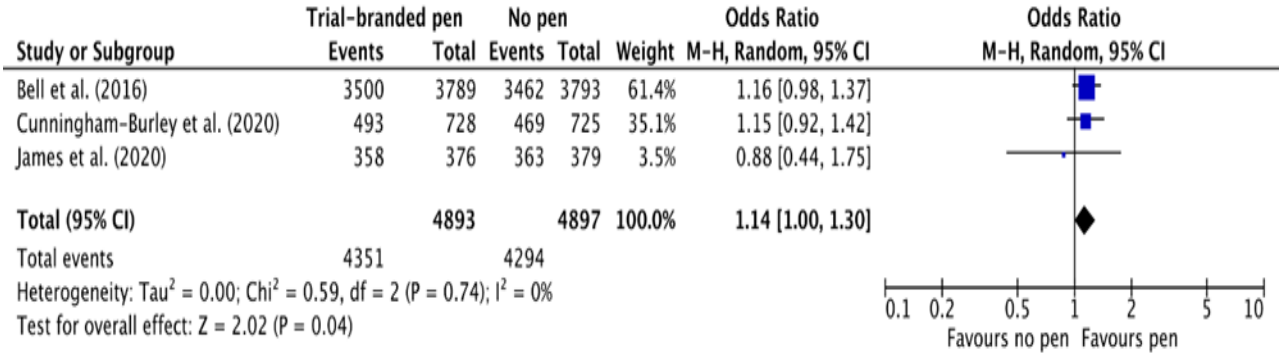
Supplemental Material 4.4. Meta-analysis tables and figures of retention strategies

Strategy: Trial-branded pen

Table 4.S4: Meta-analysis; retention strategy (Trial-branded pen versus no trial-branded pen)

Study or SWAT	Bell et al. (2016)	Cunningham-Burley et al. 2020)	James et al. (2020)	Overall
Number of participants retained in the host trial in the intervention group	3500/3789	493/728	358/376	4351/4893(88.92%)
Number of participants retained in the host trial in the control group	3462/3793	469/725	363/379	4294/4897 (87.69%)
Incremental unit cost of the intervention (2019 $ PPP)	$0.57	$0.47	$0.28	$0.52
Total sample size (n=)	7582	1453	755	9790
Odds ratio (total)	1.14 (1.00,1.30)			
Effect size (total), i.e. incremental retention rate (total)	0.07 (0.00,0.14)			
Incremental cost per additional participant retained (ICER)	**$6.98 ($3.63,N/A)**			
GRADE certainty of evidence	Moderate			
I^2 Statistic	0%			

Figure 4.S3: Meta-analysis; retention strategy (Trial-branded pen versus no trial-branded pen)

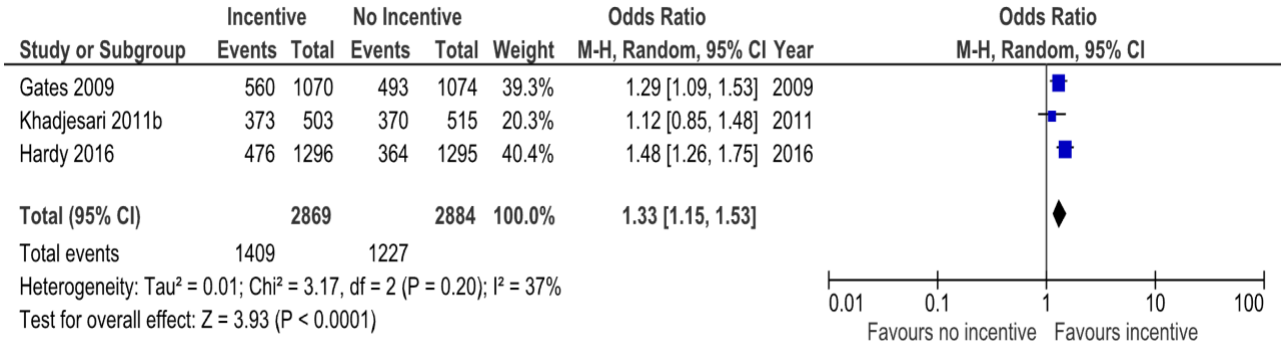


Strategy: Financial incentive

Table 4.S5: Meta-analysis; retention strategy (Financial incentive versus no incentive)

Study or SWAT	Gates et al. (2009)	Khadjesari et al. (2011b)	Hardy et al. (2016)	Overall
Number of participants retained in the host trial in the intervention group	560/1070	476/1296	373/503	1409/2869(49.11%)
Number of participants retained in the host trial in the control group	493/1074	364/1295	370/515	1227//2884(44.80%)
Incremental unit cost of the strategy (2019 $ PPP)	$6.74	$14.33	$6.54	$8.20
Total sample size (n=)	2144	2591	1018	5753
Odds ratio (total)	1.33 (1.15,1.53)			
Effect size (total), i.e. incremental retention rate (total)	0.16 (0.08,0.23)			
Incremental cost per additional participant retained (ICER)	**$15.89 ($10.65,$32.42)**			
GRADE	Moderate			
I^2 Statistic	37%			

Figure 4.S4: Meta-analysis; retention strategy (Financial incentive versus no incentive)

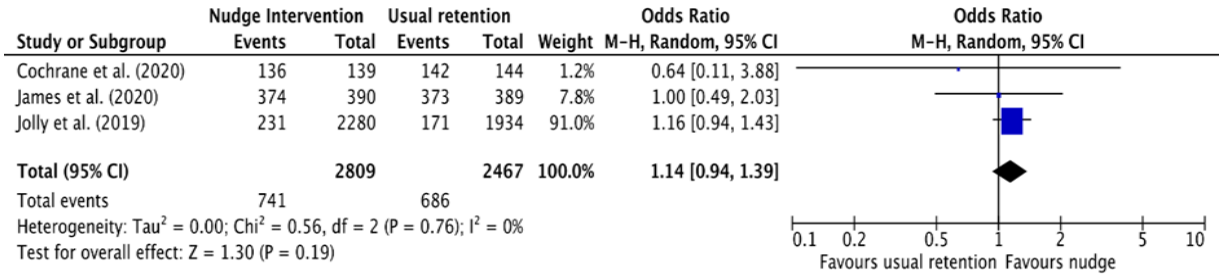


Strategy: Nudge intervention

Table 4.S6: Meta-analysis; retention strategy (Nudge intervention versus usual retention)

Study or SWAT	Cochrane et al. (2020)	James et al. (2020)	Jolly et al. (2019)	Overall
Number of participants retained in the host trial in the intervention group	136/139	374/390	231/2280	741/2809(26.38%)
Number of participants retained in the host trial in the control group	142/144	373/389	171/1934	686/2467(27.81%)
Incremental unit cost of the strategy (2019 $ PPP)	$0.07	$0.68	$0.86	$0.84
Total sample size (n=)	283	779	4214	5276
Odds ratio (total)	1.14 (0.94 ,1.39)			
Effect size (total), i.e. incremental retention rate (total)	0.07 (-0.04,0.18)			
Incremental cost per additional participant retained (ICER)	**$11.55 ($4.61,N/A)**			
GRADE	Moderate			
I^2 Statistic	0%			

Figure 4.S5: Meta-analysis; retention strategy (Nudge intervention versus usual retention procedure)

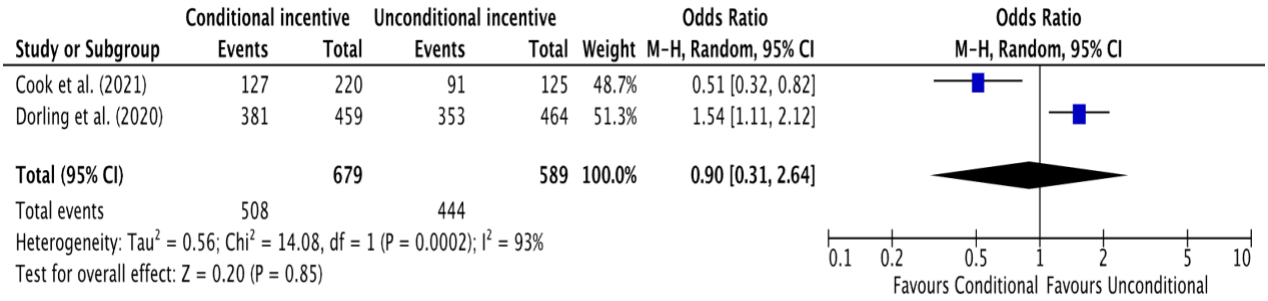


Strategy: Unconditional monetary incentive versus conditional monetary incentive

Table 4.S7: Meta-analysis; retention strategy (Unconditional monetary incentive versus conditional monetary incentive)

Study or SWAT	**Cook et al. (2021)**	**Dorling et al. (2020)**	**Overall**
Number of participants retained in the host trial in the intervention group	127/220	381/459	508/679 (74.82%)
Number of participants retained in the host trial in the control group	91/125	353/464	444/589 (75.38%)
Incremental unit cost of the strategy (2019 $ PPP)	$33.67	$4.31	$18.61
Total sample size (n=)	345	923	1268
Odds ratio (total)	0.90(0.31,2.64)		
Effect size (total), i.e. incremental retention rate (total)	-0.06 (-0.65 , 0.54)		
Incremental cost per additional participant retained (ICER)	**N/A ($10.59,N/A)**		
GRADE certainty of evidence	Low		
I^2 Statistic	93%		

Figure 4.S6: Meta-analysis; retention strategy (Unconditional monetary incentive versus conditional monetary incentive)



Retention strategies consisting of a single SWAT

Table 4.S8: Meta-analyses of retention strategies consisting of a single SWAT

Study or SWAT	**Treweek et al. (2021)**	**Clark et al. (2015)**	**Whiteside et al. (2019)**
Strategy	**Pre-notification card versus no pre-notification card**	**Electronic prompts versus no electronic prompts**	**Trial-branded pen versus no trial-branded pen (before recruitment)**
Number of participants retained in the host trial in the intervention group	231/274	157/226	27/28
Number of participants retained in the host trial in the control group	230/284	128/211	49/64
Incremental unit cost of the strategy (2019 $ PPP)	$1.02	$0.12	$0.47
Total sample size (n=)	558	437	92
Odds ratio	1.26 (0.81,1.96)	1.48 (0.99, 2.19)	8.27 (1.04, 66.00)
Effect size (total), i.e. incremental retention rate (total)	0.21 (-0.12,0.37)	0.22 (-0.01, 0.43)	1.17 (0.02,2.31)
Incremental cost per additional participant retained (ICER)	**$4.86 ($2.76, N/A)**	**$0.55 ($0.28, N/A)**	**$0.40 ($0.20, $23.50)**
GRADE certainty of evidence	Moderate	Low	Moderate

Supplemental Material 4.5. GRADE certainty of evidence of the included studies

Recruitment strategy: Financial incentive versus no financial incentive

GRADE domain	Decision	Reason
Risk of bias	Do not downgrade	All studies have a low risk of bias.
Imprecision	Do not downgrade	Sufficient sample size.
Inconsistency	Downgrade	No similarity of point estimates across studies. I^2 high
Indirectness	Do not downgrade	This recruitment strategy was properly directed towards potential trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Moderate	

Recruitment strategy: Nudge intervention vs usual recruitment

GRADE domain	Decision	Reason
Risk of bias	Downgrade	Two studies have an unclear risk of bias.
Imprecision	Do not downgrade	Sufficient sample size.
Inconsistency	Downgrade	No similarity of point estimates across studies. I^2 high
Indirectness	Downgrade	The included nudge interventions are not identical.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Very low	

Recruitment strategy: Screening for the host trial undertaken by a senior investigator vs screening undertaken by a research assistant

GRADE domain	Decision	Reason
Risk of bias	Downgrade	The included study has a high risk of bias.
Imprecision	Downgrade	Insufficient sample size. Single study only
Inconsistency	Do not downgrade	No similarity of point estimates across studies.
Indirectness	Do not downgrade	This recruitment strategy was properly directed towards potential trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Low	

Recruitment strategy: Primary text message vs primary call and no text message

GRADE domain	Decision	Reason
Risk of bias	Do not downgrade	The included study has a low risk of bias.
Imprecision	Downgrade	One study included only.
Inconsistency	Do not downgrade	No similarity of point estimates across studies.
Indirectness	Do not downgrade	This recruitment strategy was properly directed towards potential trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Moderate	

Recruitment strategy: Pre-notification leaflet versus no leaflet

GRADE domain	Decision	Reason
Risk of bias	Do not downgrade	The included study has a low risk of bias.
Imprecision	Downgrade	One study included only.
Inconsistency	Do not downgrade	No similarity of point estimates across studies.
Indirectness	Do not downgrade	This recruitment strategy was properly directed towards potential trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Moderate	

Recruitment strategy: Telephone reminder versus SMS reminder

GRADE domain	Decision	Reason
Risk of bias	Do not downgrade	The included study has a low risk of bias.
Imprecision	Downgrade	One study included only.
Inconsistency	Do not downgrade	No similarity of point estimates across studies.
Indirectness	Do not downgrade	This recruitment strategy was properly directed towards potential trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Moderate	

Recruitment strategy: Invitation pack from a surgeon vs text message

GRADE domain	Decision	Reason
Risk of bias	Downgrade	Unknown.
Imprecision	Downgrade	One study included only.
Inconsistency	Do not downgrade	No similarity of point estimates across studies.
Indirectness	Downgrade	Unknown
Publication bias	Downgrade	Unknown
Overall GRADE	Very low	

Recruitment strategy: Trial-branded pen vs no trial-branded pen

GRADE domain	Decision	Reason
Risk of bias	Do not downgrade	The included study has a low risk of bias.
Imprecision	Downgrade	One study included only.
Inconsistency	Do not downgrade	No similarity of point estimates across studies.
Indirectness	Do not downgrade	This retention strategy was properly directed towards recruited trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Moderate	

Retention strategy: Trial-branded pen vs no trial-branded pen

GRADE domain	Decision	Reason
Risk of bias	Do not downgrade	One study has a low risk of bias, and two studies have an unclear risk of bias.
Imprecision	Do not downgrade	Sufficient sample size. The effect is not at the upper versus the low end of the confidence interval.
Inconsistency	Downgrade	Differences of point estimates across studies. I^2 low
Indirectness	Do not downgrade	This retention strategy was properly directed towards recruited trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Moderate	

Retention strategy: Financial incentive versus no financial incentive

GRADE domain	Decision	Reason
Risk of bias	Downgrade	One study has a low risk of bias, one study has an unclear risk of bias and one study has a high risk of bias.
Imprecision	Do not downgrade	Sufficient sample size. The effect is not at the upper versus the low end of the confidence interval.
Inconsistency	Do not downgrade	Similarity of point estimates across studies. I^2 low
Indirectness	Do not downgrade	This retention strategy was properly directed towards recruited trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Moderate	

Retention strategy: Nudge intervention versus usual retention

GRADE domain	Decision	Reason
Risk of bias	Do not downgrade	One study has a low risk of bias, and two studies have an unclear risk of bias.
Imprecision	Do not downgrade	Sufficient sample size.
Inconsistency	Downgrade	No similarity of point estimates across studies. I^2 low
Indirectness	Do not downgrade	This retention strategy was properly directed towards recruited trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Moderate	

Retention strategy: Unconditional monetary incentive versus conditional monetary incentive

GRADE domain	Decision	Reason
Risk of bias	Downgrade	Two studies have an unclear risk of bias.
Imprecision	Do not downgrade	Sufficient sample size.
Inconsistency	Downgrade	No similarity of point estimates across studies. I^2 very high
Indirectness	Do not downgrade	This retention strategy was properly directed towards recruited trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Low	

Retention strategy: Pre-notification card versus no pre-notification card

GRADE domain	Decision	Reason
Risk of bias	Do not downgrade	The included study has a low risk of bias.
Imprecision	Downgrade	One study included only.
Inconsistency	Do not downgrade	No similarity of point estimates across studies.
Indirectness	Do not downgrade	This retention strategy was properly directed towards recruited trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Moderate	

Retention strategy: Electronic prompts versus no electronic prompts

GRADE domain	Decision	Reason
Risk of bias	Downgrade	Unclear risk of bias.
Imprecision	Downgrade	One study included only.
Inconsistency	Do not downgrade	No similarity of point estimates across studies.
Indirectness	Do not downgrade	This retention strategy was properly directed towards recruited trial participants.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Low	

Retention strategy: Trial-branded pen vs no trial-branded pen (before recruitment)

GRADE domain	Decision	Reason
Risk of bias	Do not downgrade	The included study has a low risk of bias
Imprecision	Downgrade	One study included only.
Inconsistency	Do not downgrade	No similarity of point estimates across studies.
Indirectness	Do not downgrade	This retention strategy was properly directed towards patients potentially eligible for recruitment to the original host trial.
Publication bias	Do not downgrade	No inference about missing evidence
Overall GRADE	Moderate	

Supplemental Material 4.6. Trial Forge Guidance 2

Recruitment strategy: Financial incentive versus no financial incentive

Criterion	Comments	Criterion Met (Yes or No or Partially)
GRADE	Moderate certainty of evidence. Lower than high.	Yes
Cumulated evidence	The OR has not converged.	Yes
PICOT	P: Criterion not met. The populations include adults of both sexes sufficiently. I: Criterion met. The interventions are health-based, but some of them are treatment-related and some of them are prevention-related. Also, the health conditions are different among the host trials. C: Criterion met. The comparators are different among the host trials. O: Criterion not met. The recruitment rate was reported as an outcome in all included SWATs T: Criterion not met. Sending financial payments to participants remains a popular operation during an RCT.	Partially
Balance of benefit and disadvantage to participants	Providing a financial incentive is likely to be beneficial to trial participants, as their disposable income increases. Since most SWATs included a relatively high incentive of £100, there is low chance the participants would find such an incentive to be insulting for their contribution to trials.	No
Balance of benefit and disadvantage to host trial	The benefit to host trials would be a moderate increase in recruitment rates. The disadvantage to host trials would be the direct financial costs of incentives. Workload is not expected to increase.	No

Recruitment strategy: Nudge intervention vs usual recruitment

Criterion	Comments	Criterion Met (Yes or No or Partially)
GRADE	Very low certainty of evidence. Lower than high.	Yes
Cumulated evidence	The OR has not converged.	Yes
PICOT	P: Criterion not met. The populations include adults of both sexes sufficiently. I: Criterion met. Some interventions are delivered in primary care, whereas some others are offered at the community level. Also, the nudge interventions offered are not identical between studies. C: Criterion met. There is heterogeneity in the comparators of the five host trials. O: Criterion not met. The recruitment rate was reported as an outcome in all included SWATs T: Criterion not met. The offered nudge interventions are modern and hence relatable to RCTs.	Partially
Balance of benefit and disadvantage to participants	Nudge interventions are intended not to cause any harm to patients invited for screening to an RCT.	No
Balance of benefit and disadvantage to host trial	There is a clear benefit to the host trials, as identifying effective nudge interventions could explain the underlying factors behind poor and/or slow recruitment to RCTs. However, staff workload is expected to increase wrt nudge interventions.	Partially

Retention strategy: Trial-branded pen vs no trial-branded pen

Criterion	**Comments**	**Criterion Met (Yes or No or Partially)**
GRADE	Moderate certainty of evidence. Lower than high.	Yes
Cumulated evidence	The OR has converged.	No
PICOT	P: Criterion met as no SWAT achieved sufficient sample size from young adult men. I: Criterion met. The interventions are health-based, but some of them are treatment-related and some of them are prevention-related. Also, the health conditions are differential among the host trials. C: Criterion met. The comparators are different among the host trials. O: Criterion not met. The questionnaire response rate was reported as an outcome in all included SWATs T: Criterion not met. Sending questionnaires via post remains a popular operation during an RCT.	Partially
Balance of benefit and disadvantage to participants	Providing a pen may be perceived as useful to some participants, but not useful to some others. The mechanisms on which a participant would be more likely to feel benefited from a pen are unknown.	Yes
Balance of benefit and disadvantage to host trial	There is a clear benefit to the host trials, as adding pens for follow-up may improve the retention rate at a minimal incremental cost	No

Retention strategy: Financial incentive versus no financial incentive

Criterion	**Comments**	**Criterion Met (Yes or No or Partially)**
GRADE	Moderate certainty of evidence. Lower than high.	Yes
Cumulated evidence	The OR has converged.	No
PICOT	P: Criterion not met. The populations include adults of both sexes sufficiently. I: Criterion partially met. The interventions are all related to secondary care, but some of them are education-oriented and some are psychology-oriented. C: Criterion met. The comparators are different among the host trials. O: Criterion not met. The questionnaire return rate was reported as an outcome in all included SWATs T: Criterion not met. Sending financial payments to participants remains a popular operation during an RCT	Partially
Balance of benefit and disadvantage to participants	Providing a financial incentive is likely to be beneficial to trial participants, as their disposable income increases. However, since all SWATs included a low incentive of up to £10, it could be possible that participants find such a payment insulting.	Partially
Balance of benefit and disadvantage to host trial	The benefit to host trials would be a moderate increase in retention rates. The disadvantage to host trials would be the direct financial costs of incentives. Workload is not expected to increase.	No

Retention strategy: Nudge intervention vs usual recruitment

Criterion	**Comments**	**Criterion Met (Yes or No or Partially)**
GRADE	Moderate certainty of evidence. Lower than high.	Yes
Cumulated evidence	The OR has not converged.	Yes
PICOT	P: Criterion met as participants aged less than 65 are not equally represented in the sample. I: Criterion met. The interventions are health-based, but some of them are treatment-related and some of them are prevention-related. Also, the health conditions are differential among the host trials. C: Criterion met. The comparators are different among the host trials. O: Criterion not met. The questionnaire return rate was reported as an outcome in all included SWATs T: Criterion not met. Sending questionnaires via post remains a popular operation during an RCT.	Partially
Balance of benefit and disadvantage to participants	Nudge interventions are intended not to cause any harm to recruited RCT participants.	No
Balance of benefit and disadvantage to host trial	There is a clear benefit to the host trials, as identifying effective nudge interventions could explain the underlying factors behind poor participant withdrawal from RCTs. However, staff workload is expected to increase wrt nudge interventions, and the effectiveness of such a strategy is uncertain.	Partially

Retention strategy: Unconditional monetary incentive versus conditional monetary incentive

Criterion	**Comments**	**Criterion Met (Yes or No or Partially)**
GRADE	Low certainty of evidence. Lower than high.	Yes
Cumulated evidence	The OR has not converged.	Yes
PICOT	P: Criterion met as in one of the two included SWATs only included parents of Infants born at <32 weeks' gestation or a weight of <1500g, who were receiving <30 ml/kg/day of milk at trial enrolment. I: Criterion met. One intervention is delivered in primary care, and another is delivered in secondary care. C: Criterion met. The comparators are different among the host trials. O: Criterion not met. The questionnaire response rate was reported as an outcome in all included SWATs T: Criterion not met. Sending financial payments to participants remains a popular operation during an RCT	Partially
Balance of benefit and disadvantage to participants	Providing a financial incentive is likely to be beneficial to trial participants, as their disposable income increases. However, the provision of conditional monetary incentive could instinctively signal poor appreciation for the efforts of the recruited participants.	Yes
Balance of benefit and disadvantage to host trial	There is no evidence of increase in retention rates as a result of providing unconditional incentives vs conditional ones. Also, the provision of incentives would increase the cost of the conduct of the trials. Workload is not expected to increase.	Yes

Supplemental Material 4.7. Exclusion of three included studies from meta-analysis

In one study (Gates et al., 2009), the participants were allocated either to receive a diary through postal administration, or telephone administration, or clinic visits; then, they were randomised either to receive a £2 voucher or no voucher to complete the diary. The way the study was run and analysed did not allow the reviewers to include its results in the meta-analysis, since the features of an appropriate 3x2 factorial SWAT were not addressed. The findings of the study suggest that providing a £2 voucher was a cost-effective strategy that eventually reduced the average costs per returned diary, compared to participants who did not receive the voucher. These results are in line with the corresponding meta-analysis of providing financial incentives as a strategy to improve participant retention in RCTs.

Another study included a costing figure which is hypothetical and for which no clear stratification of incremental effects and costs has been made (Kenyon et al., 2005). However, its reported cost per additional questionnaire returned (£67, or $94.63 in PPP rates) does not significantly diverge from the quantitative findings of the meta-analysis of financial incentives as a retention strategy.

Finally, another SWAT (Khadjesari et al., 2011) explored several interventions associated with financial incentives, i.e. different types of £5 vouchers and a £250 draw, however its reported figure of cost per additional questionnaire returned, i.e. £110.15, or $156.02 in PPP rates, does cumulatively correspond to all interventions, with no explanation behind the stratification of the figure. As a result, it was impossible to include this study (SWAT) in the meta-analysis of the financial incentives as a strategy to improve participant retention in RCTs, since the incremental costs associated with each intervention could not be computed by the reviewers.