Allocation concealment in randomised controlled trials: improving the implementation, reporting and developing methods to detect and prevent selection bias

Laura Clark MSc

PhD

University of York

Health Sciences

July, 2022

Abstract

Randomised Controlled Trials (RCTs) improve the robustness of evidence and the knowledge base but only if they are conducted well. A key methodological component of an RCT is allocation concealment. Inadequate allocation concealment can lead to selection bias. Selection bias needs to be prevented, and, where possible detected when undertaking quality assessments. This thesis makes an original contribution to the knowledge base by exploring a method to detect selection bias in systematic reviews and improve allocation concealment implementation and reporting in RCTs.

Paper 1 presents an exemplar of a novel technique to detect selection bias within systematic reviews. I found that systematic reviews exhibited baseline age heterogeneity, which can only be explained by the inclusion of RCTs with baseline selection bias. Paper 2 explores this technique further, demonstrating that more powerful baseline covariates in a baseline meta-analysis found greater heterogeneity.

In paper 3 I show this baseline heterogeneity is not a function of the reviews including 'old' trials because poor allocation concealment, which is the most likely explanation of selection bias, remains prevalent in recently published RCTs. In these recent RCTs, envelope and blocking implementation and reporting was suboptimal. In response, I focused my methodological research on improving the implementation and reporting of these methods.

In paper 4 I show envelopes are widely used as an allocation concealment method and their use will likely continue. Therefore, I created an original reference guide for researchers to utilise to aid the robust implementation and reporting of envelopes. In paper 5 I published the first report using empirical evidence, rather than modelling studies, that RCTs using small variable block sizes that include a block of two are at risk of subversion

Throughout my research I make recommendations to ensure allocation concealment is implemented and reported robustly to improve evidence generation and synthesis.

List of contents

List of tables

List of figures

Preface

Methodologists have worked for decades to improve the quality of the design, implementation and reporting of health research to benefit patients. The global research community needs to continue this whilst evolving and adapting to new techniques and ways to engage and disseminate findings. The following quote encapsulates the motivation behind this thesis.

> To realise the full value of his [Doug Altman's] legacy, research funders, research regulators, research organisations, journals, and the many people Doug taught and inspired must act together to design, conduct, and report better research done for the right reasons. The continuing ethical, scientific, and economic deficiencies of medical research remain scandalous'
>
> (Glasziou and Chalmers, 2018)

Acknowledgements

I have many people to thank for supporting me throughout my PhD journey. I would firstly like to thank my supervisors Alison Booth and Kate Flemming for their support, expertise and kindness. Thank you both for making this experience enjoyable and for understanding what I was trying to achieve when telling the story of my research, for inspiring me and instilling confidence to go back to my original plans.

I would like to thank David Torgerson and Catherine Hewitt, for always championing PhD by Publication and for giving me the opportunity and encouragement to pursue this.

Thank you to my colleagues, friends and co-authors. I appreciate the time you gave to help with my research and enhance my critical thinking, I have learnt a lot from you all. Thank you in particular to Catherine Arundel for your constant encouragement.

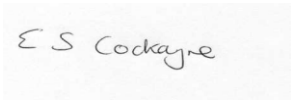
Finally, to my family, thank you to my Mum and my husband Richard for your support, help and patience whilst I wrote this thesis. To my children Oliver, Lily and Sam. Thank you, Oliver, for your technical support, Lily for your proof-reading and feedback and Sam for reminding me of my timeline and making sure I adhere to it.

Author 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 an award at this, or any other, University. All sources are acknowledged as References.

<table>
<tr><th>Paper</th><th>Publication</th></tr>
<tr><td rowspan="2">1</td><td>A methodological review of recent meta-analysis has found significant heterogeneity in age between randomised groups

Clark, L., Fairhurst, C., Hewitt, C.E., Birks, Y., Brabyn, S., Cockayne, S., Rodgers, S., Hicks, K., Hodgson, R., Littlewood, E. and Torgerson, D.J., 2014. A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups. Journal of clinical epidemiology, 67(9), pp.1016-1024.</td></tr>
<tr><td>LC contribution:
Joint research conception
Led on the research design and identification of component RCTs
Joint data collection (data extraction)
Joint analysis – received training and mentorship whilst undertaking this
Led on drafting initial publication and subsequent drafts
 
Signed: Laura Clark 15.07.2021

Confirmed by co-author (signed)  15.07.2021
Sarah Cockayne</td></tr>
<tr><td rowspan="2">2</td><td>Important outcome predictors showed greater baseline heterogeneity than age in two systematic reviews

Clark, L., Fairhurst, C., Cook, E. and Torgerson, D.J., 2015. Important outcome predictors showed greater baseline heterogeneity than age in two systematic reviews. Journal of clinical epidemiology, 68(2), pp.175-181.</td></tr>
<tr><td>LC contribution:
Led on research conception
Led on the research design and identification of component RCTs
Joint data collection (data extraction)
Undertook analysis, this was checked by CF
Led on drafting initial publication and subsequent drafts
 
Signed: Laura Clark 15.07.2021

Confirmed by co-author (signed)  15/07/2021
Elizabeth Cook</td></tr>
</table>

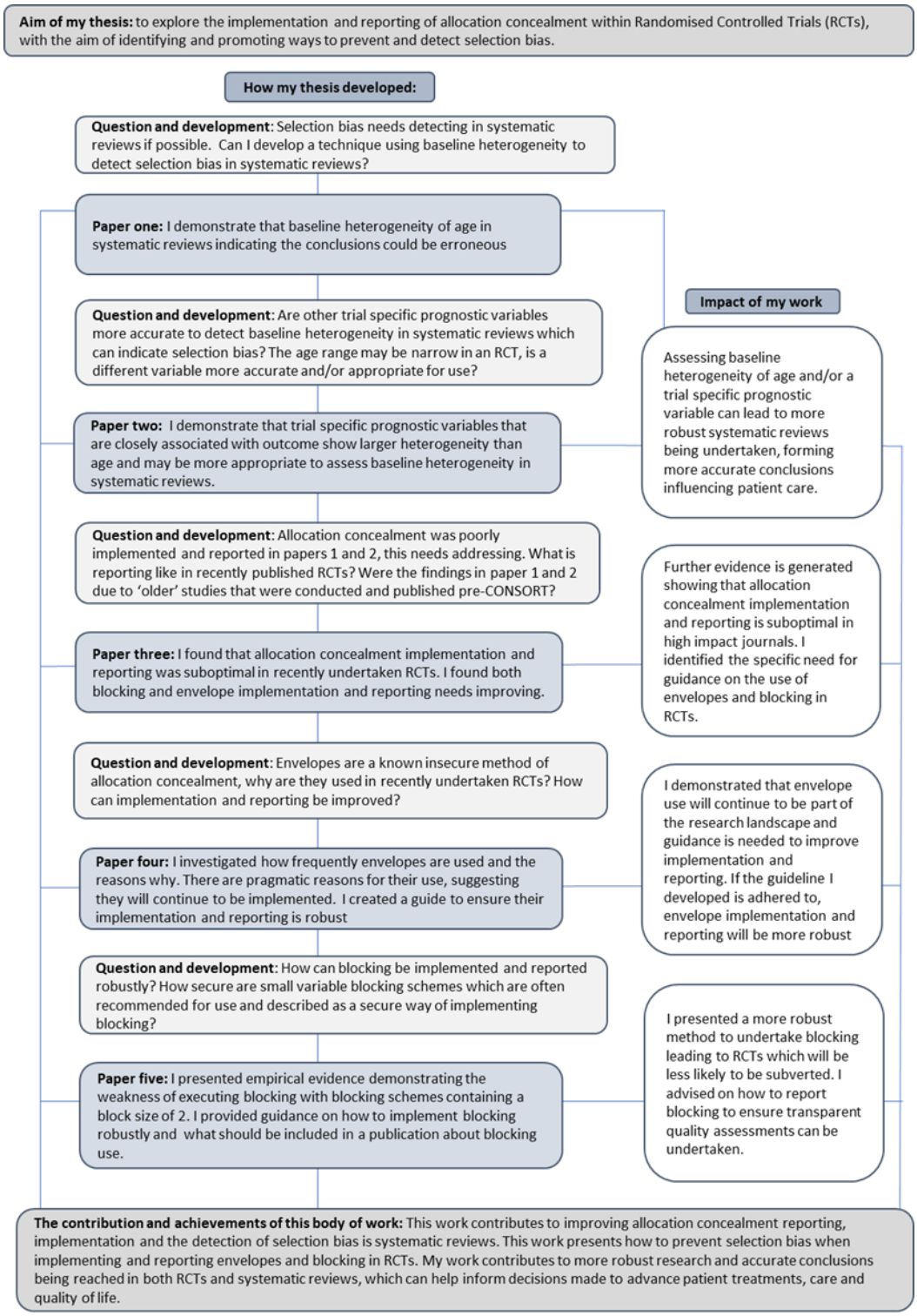
3	Allocation concealment in randomised controlled trials: are we getting better? Clark, L., Fairhurst, C. and Torgerson, D.J., 2016. Allocation concealment in randomised controlled trials: are we getting better? *BMJ*, *355*, p.i5663.
	LC contribution: Led on research conception Led on the research design and identification of component RCTs Joint data collection Led on drafting initial publication and subsequent drafts  Signed: Laura Clark 15.07.2021  Confirmed by co-author (signed) Caroline Fairhurst 15.07.2021
4	Envelope use and reporting in randomised controlled trials: a guide for researchers Clark, L., Dean, A., Mitchell, A. and Torgerson, D.J., 2021. Envelope use and reporting in randomised controlled trials: A guide for researchers. *Research Methods in Medicine & Health Sciences*, *2*(1), pp.2-11.
	LC contribution: Led on research conception Led on the research design and identification of component RCTs Joint data collection Led on the analysis Led on the development of the tool Led on drafting initial publication and subsequent drafts  Signed: Laura Clark 16.07.2021  Confirmed by co-author (signed) David Torgerson 16/07/21

5	A review found small variable blocking schemes may not protect against selection bias in randomized controlled trials Clark, L., Burke, L., Carr, R.M., Coleman, E., Roberts, G. and Torgerson, D.J., 2022. review found small variable blocking schemes may not protect against selection bias in randomized controlled trials. *Journal of Clinical Epidemiology*, *141*, pp.90-98.
	LC contribution: Led on the research conception Led on the research design and identification of component RCTs Joint data collection Led on the analysis Led on drafting initial publication and subsequent drafts  Signed: Laura Clark 19.10.2021  Confirmed by co-author (signed) Elizabeth Coleman 19.10.2021

Section 1: Introduction and context

My research developed over several years, figure 1 demonstrates the progression and impact of my research.

Figure 1: infographic demonstrating the thesis development and impact



‘To maximise the benefit to society you need to not just do research but to do it well’

Doug Altman (CONSORT, 2018)

Research has been conducted for centuries, is complex, multifaceted and depends on a set of processes working together. Although there is a multiplicity of research methods they share the same structure; the research is designed, undertaken and reported. Health research, where my work focuses, aims to advance medical understanding for the benefit of patients. Advancements include developing new treatments and diagnoses, or identifying treatments used in practice causing harm or having no effect.

The first Randomised Controlled Trial (RCT) in medicine was published in the British Medical Journal (BMJ) in 1948 (Marshall et al., 1948) and is a design frequently employed in health research. Prior to the RCT design there were before and after observational studies, which included controlled trials, but randomisation was not used to form the comparison groups. RCTs revolutionised research by providing a design that minimises bias. Randomisation ensures that participants in the intervention and control groups are similar in terms of prognostic factors – both known and unknown, so observed differences can be attributed to the causal effect of the intervention (Higgins et al., 2011, Kennedy et al., 2017, Kunz and Oxman, 1998, Martyn, 1996). Consequently, in primary research, RCTs are considered the gold standard in assessing the effectiveness of an intervention.

RCTs improve the robustness of evidence and knowledge base but only if they are conducted well. As importantly, RCTs need to be reported adequately, with methodology clearly stated so an assessment on the robustness of the RCT can be made. However, the reporting of RCTs lagged behind the improvement they brought to research robustness, and has been suboptimal for decades as reported in ‘The scandal of poor medical research’ paper (Altman, 1994).

The EQUATOR (Enhancing the QUAlity and Transparency Of health Research) network was developed to increase the quality and transparency of health research reporting (EQUATOR, 2021), by providing a repository of reporting

guidelines for authors. The Consolidated Standards of Reporting Trials (CONSORT) statement (Altman et al., 2001a, Altman et al., 2001b, Begg et al., 1996, Schulz et al., 2010) is the most well-known reporting guideline for RCTs. The Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statement (Chan et al., 2013) provides guidelines for reporting trial protocols. Twenty years after Altman's landmark paper (Altman, 1994) 'The Lancet' published a series on increasing value and reducing research waste (Ioannidis et al., 2014), highlighting reporting is still not robust despite the development of reporting guidelines and promotion of the importance of robust reporting.

Although not inevitable, RCTs may lead to changes in policies and practices (Hariton and Locascio, 2018), their results can be 'stand-alone' and/or accumulated into systematic reviews and meta-analyses (Gopalakrishnan and Ganeshkumar, 2013). Within the hierarchy of evidence, large RCTs, meta-analyses and systematic reviews of RCTs are considered to provide the highest and most reliable quality evidence with the lowest risk of bias (ROB) (Murad et al., 2016, Higgins et al., 2011). The following is one example of the impact that robust RCTs and well conducted meta-analyses can have. An observational study concluded that antioxidant supplements (including vitamin E) protects against heart disease (Khaw et al., 2001). Subsequent RCTs have shown vitamin E had no benefit (MRC/BHF, 2002) and increased colds/infections in the elderly (Gratt et al., 2002). Meta-analyses have shown that vitamin E supplementation of >800IU/day increases mortality (Bjelakovic et al., 2014, Miller III et al., 2005). Consequently, recommendations for patient care were changed.

It is essential that those conducting an RCT understand why all aspects need to be robust. Historically, the value of RCTs has been undermined by the actions of those involved in their conduct. Research investigating whether supplemental oxygen was associated with retrolental fibroplasia in premature babies was undermined by research nurses giving the control group babies oxygen overnight. The nurses believed not giving oxygen was unethical, due to their lack understanding of the purpose of randomisation in the context of the unknown harms and benefits. This delayed the finding that supplemental oxygen was positively correlated with retrolental fibroplasia (Silverman, 1977).

The opening quote of this thesis underpins my work; research is conducted to progress knowledge within a research area for the benefit of society, research undertaken needs to be designed, implemented and reported robustly (Clarke et al., 2010, Glasziou et al., 2006). Globally we have all experienced why this is essential. Throughout the COVID-19 pandemic we depended on the international research community to design, implement and report robust research leading to vaccines and treatments.

The wider picture can be lost when only considering the methodological factors constituting research: research starts and ends with patients. To patients, research is more than a set of processes. There are many reasons and agendas as to why research is funded, but for patients and their families research can contribute to changing lives; research advances the prevention and treatment of illnesses and maximises quality of life. Participants give their time to take part in research, putting their trust into the research team that it will be undertaken and reported robustly. The research community has an ethical and moral responsibility to ensure this happens.

My methodological research focuses on a key aspect of the design and implementation of an RCT: adequate concealment of the allocation sequence. Throughout the rest of this section I define allocation concealment, and, by drawing on published work on poor allocation concealment, explain why adequate allocation concealment is important in RCTs.

1.1: Allocation concealment and selection bias

The term allocation concealment was coined in 1994 and involves two separate processes (Schulz et al., 1994). The first, to generate an unbiased and truly random allocation sequence (randomisation), the second to conceal the allocation sequence from all those involved in recruiting participants into the RCT (Higgins et al., 2019). These processes ensure trial personnel do not have fore knowledge of the sequence and cannot select participants into one group or another, which would introduce selection bias (Schulz et al., 2018, Schulz et al., 1994). CONSORT defines allocation concealment as:

'A technique used to prevent selection bias by concealing the allocation sequence from those assigning participants to intervention groups, until the moment of assignment. Allocation concealment prevents researchers from (unconsciously or otherwise) influencing which participants are assigned to a given intervention group' (CONSORT, 2022b).

Allocation concealment can be implemented in every RCT (Forder et al., 2005, Piper et al., 1996), and can be confused with the term blinding (Schulz et al., 1995, Schulz and Grimes, 2002b) which cannot be achieved in all RCTs (Forder et al., 2005). Selection bias is a major problem in research as it can lead to biased conclusions, defined as:

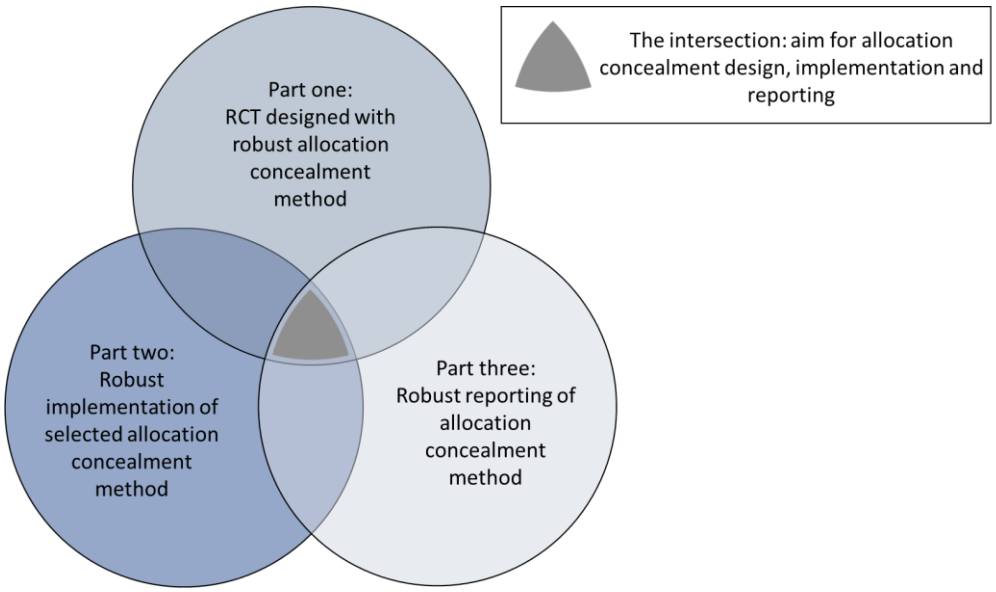
'Systematic error in creating intervention groups, such that they differ with respect to prognosis. That is, the groups differ in measured or unmeasured baseline characteristics because of the way participants were selected or assigned. Also used to mean that the participants are not representative of the population of all possible participants' (CONSORT, 2022c).

Selection bias is common within observational studies, whereas RCTs are designed to remove this (Berger, 2005). If the allocation sequence is deciphered, the trial then becomes a non-randomised trial, and is not robust evidence (Viera and Bangdiwala, 2007, Higgins et al., 2011, Murad et al., 2016). The interpretation of a poorly implemented RCT is more concerning than an observational study, as the latter is acknowledged as being at high ROB whereas the former may not be. Selection bias in RCTs needs controlling by prevention, detection or correction (Mickenautsch et al., 2014). My research focuses on prevention and detection.

1.1.1: Introducing the trifecta of allocation concealment

Allocation concealment is part of a trifecta, the components are; designing/selecting, implementing, and reporting allocation concealment. Each part of the trifecta needs to be fulfilled to achieve a robust RCT. Figure 2 demonstrates the trifecta concept.

Figure 2: the allocation concealment trifecta in RCTs: robust design, implementation and reporting



There are many stages to the robust design, implementation and reporting of allocation concealment and many stages where selection bias can be introduced. Figures 3a and 3b present the research process in relation to allocation concealment in detail, from the research question posed to the assessment of quality. Figure 3a demonstrates the process when an insecure allocation concealment method has been selected, and, 3b when a robust method is selected. When interpreting these figures I have made the assumptions that any ROB assessments undertaken are well conducted and readers understand the implications of a high risk of selection bias (Higgins et al., 2019). This thesis refers to ROB assessments in relation to randomisation and allocation concealment and not any of the other many and varied potential sources of bias. Sections 1.1.2-1.1.4 explain the allocation concealment trifecta further.

Figure 3a: Potential allocation concealment design, implementation and reporting pathways in RCTs when an insecure allocation concealment method has been selected

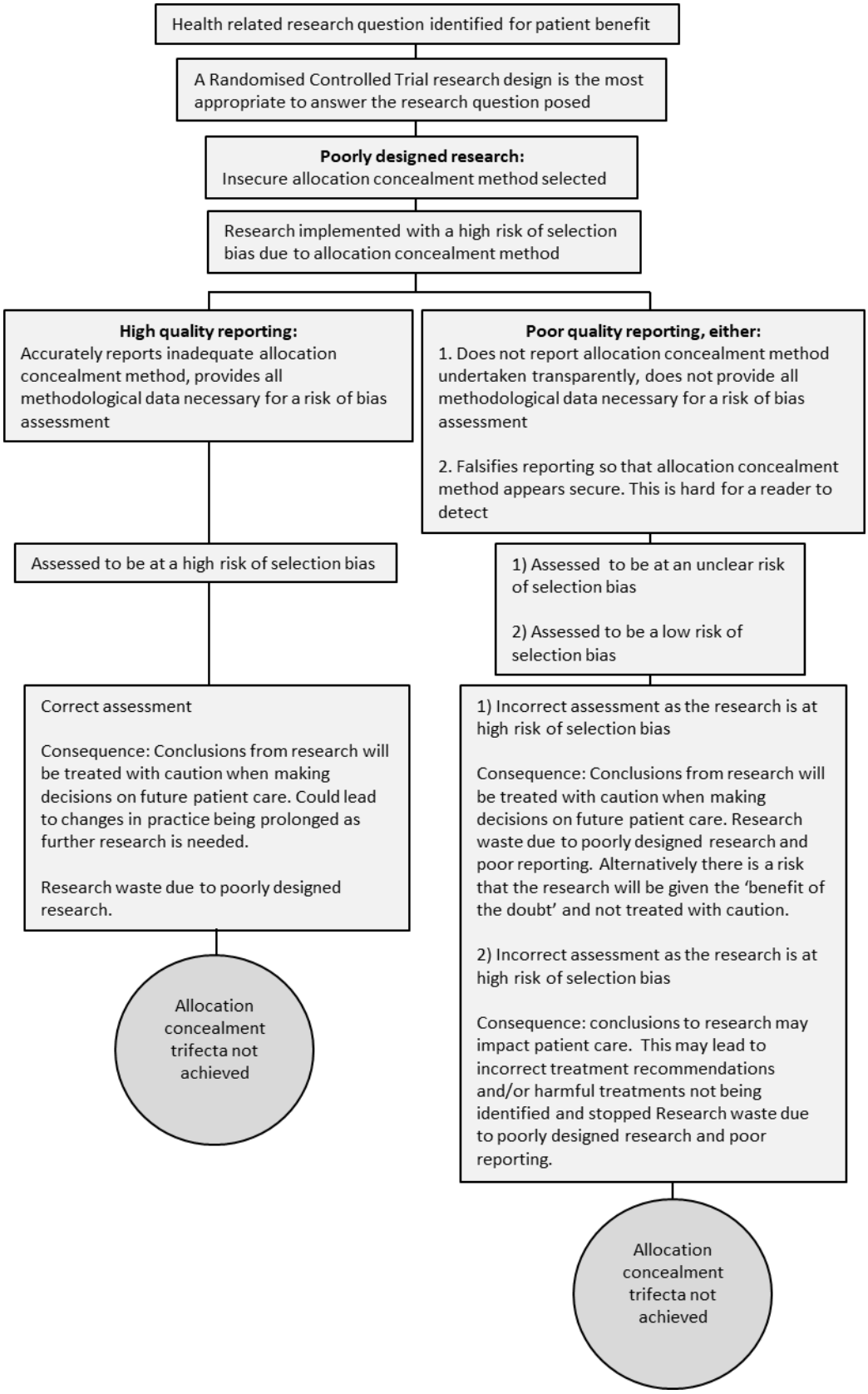
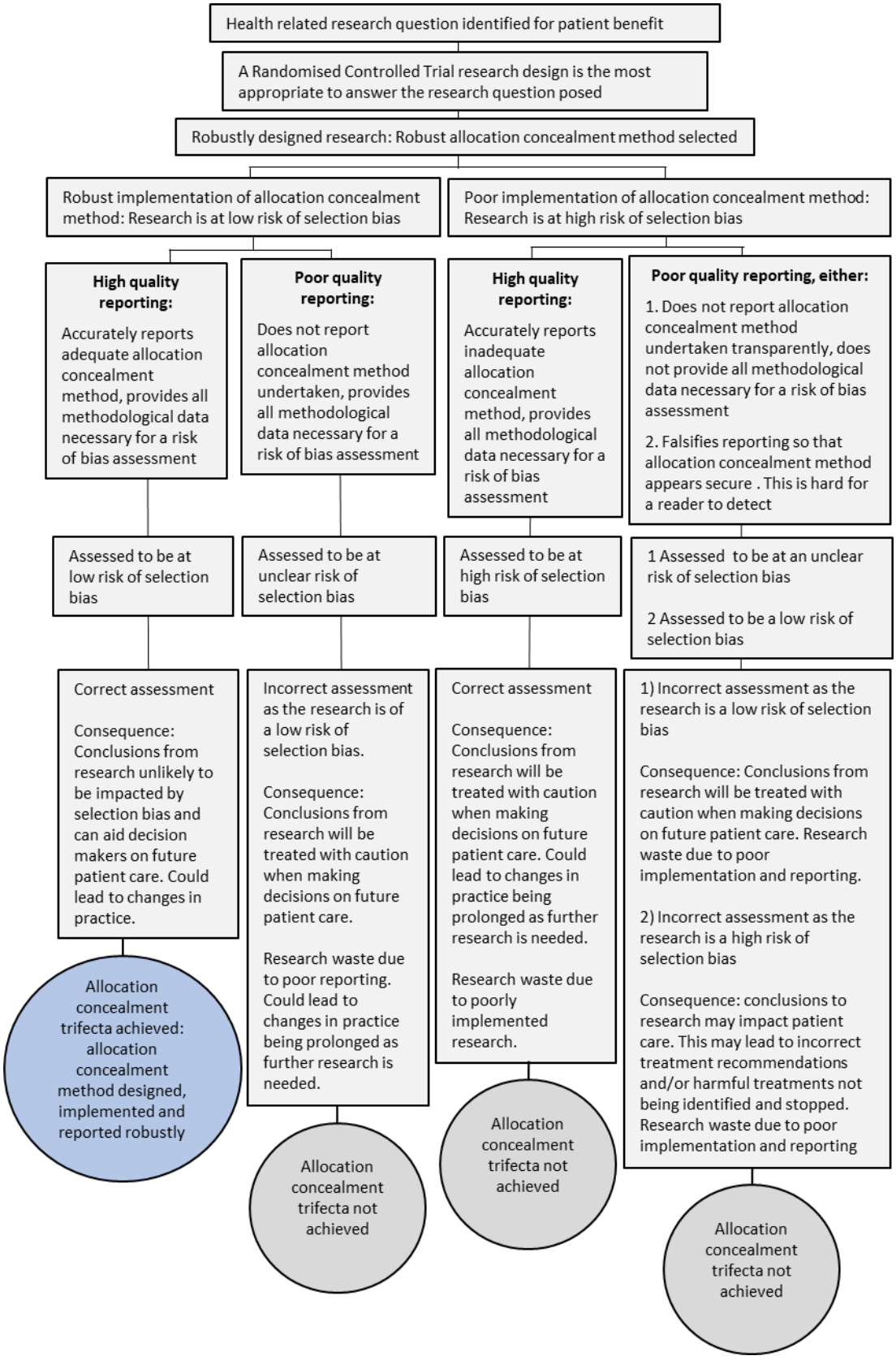


Figure 3b: Potential allocation concealment design, implementation and reporting pathways in RCTs when a robust allocation concealment method has been selected



1.1.2: Trifecta component part one: designing RCTs and selecting allocation concealment methods to prevent selection bias

When designing an RCT, robust allocation concealment methods must be selected. Inadequate (or insecure) allocation concealment methods do not enable the allocation sequence to be hidden from those recruiting participants into an RCT, putting the RCT at risk of subversion. Subversion is the term given to the deliberate tampering of allocating participants to one group or another (Schulz, 1995, Kennedy et al., 2017) thus introducing selection bias and potentially impacting conclusions. For example Odd Ratios were found to be exaggerated by 41% in trials with inadequate concealment and 30% where concealment was unclear compared to those with well reported methods (where transparent reporting is provided for both robust and poor allocation concealment methods) (Schulz et al., 1995), this could lead to the conclusion that the intervention is more effective than it is.

Allocation sequence generation methods can be categorised as either simple or restricted. CONSORT define simple randomisation as: '*Randomisation without restriction. In a two-group trial, it is analogous to the toss of a coin' (CONSORT, 2022b),* and restricted as*: 'Any procedure used with random assignment to achieve balance between study groups in size or baseline characteristics…'* (CONSORT, 2022c).

There is greater predictability and poorer allocation concealment observed in allocation sequences using any form of restriction, simple allocation methods are considered more secure (Schulz et al., 1994, Schulz et al., 1995, Schulz and Grimes, 2002a, Berger, 2009). Simple randomisation can result in groups that are unbalanced at baseline in important prognostic factors by chance, particularly in sample sizes of less than 100 (Altman and Bland, 1999).

Blocking is a common type of restricted randomisation, associated with increased predictability of the allocation sequence, small fixed block sizes increase the chance of prediction further (Berger, 2005). Central randomisation is considered the most robust method as it separates the recruiter from the allocation sequence, making subversion more difficult (Higgins et al., 2019). Methodologically, envelopes are a challenging allocation concealment method as they are associated with more manipulations to the allocation sequence and

increased effect sizes when compared to more secure allocation concealment methods, such as central randomisation (Hewitt et al., 2005, Kennedy et al., 2017, Peto, 1999, Mitchell et al., 2019, Berger, 2007, Brown et al., 2005, Viera and Bangdiwala, 2007, Hewitt et al., 2009, Paludan-Müller et al., 2016). Envelope use is discussed in Section 4. Details of randomisation and allocation concealment methods are provided in appendix 1, table 1.

1.1.3: Trifecta component part two: Implementing allocation concealment methods robustly to prevent selection bias

Although RCTs are considered the most reliable of all study designs to mitigate against selection bias, research shows that this reliability is often not enacted in practice (Berger, 2005). One reason is poor implementation of allocation concealment. A robust allocation concealment method that is not implemented securely can enable recruiters to select participants into a treatment group or deliberately not recruit a patient, based on a prognostic variable detectable at baseline (Schulz, 1995, Jüni et al., 2001). A common prognostic variable on which a participant can be selected into one group or another is age, as generally older age is associated with poorer health outcomes. Implementing allocation concealment methods poorly is likened to pinning a randomisation list up for all those involved in a trial to see ahead of recruitment therefore rendering the RCT at risk of selection bias (Altman and Dore, 1990, Chalmers et al., 1983, Kennedy et al., 2017, Schulz et al., 1995). Researchers can subvert allocation sequences in a variety of ways, table 1 summarises these.

Subversion can occur for seemingly good reasons. The intervention treatment may be seen as a ‘last chance’ for some patients. Recruiters may believe a treatment should be rolled out on a wider scale due to their pre-conceived belief the treatment is effective, so select participants to support its effect, or recruiters may receive pressure from participants (Paludan-Müller et al., 2016, Brown et al., 2005). The intellectual challenge may be too tempting to try and ‘solve’ the allocation sequence or perhaps there is not enough knowledge as to the importance of concealing the allocation (Schulz and Grimes, 2002a). Once

personally involved in a trial, recruiters may not maintain their impartiality (Schulz and Grimes, 2002a).

Deliberate violation of the allocation sequence results in a biased trial that could have far reaching consequences on future patient care. Returning to the example of nurses giving control babies supplemental oxygen – many patients are blind and living with the consequence of deliberate tampering of the allocation sequence. If this research had been implemented robustly, the practice of supplementary oxygen would likely have ceased earlier, preventing lifelong blindness for those patients (Silverman, 1977).

Table 1: methods of subversion

How the allocation sequence is subverted	**References**
• Keeping a log of previous allocations, particularly if restricted randomisation has been used such as blocking to detect a randomisation schedule that can be predicted • Requesting more than one allocation at once from third party randomisation services and selecting participants into a group accordingly • Assessing the appearance of allocation container and labels to determine which treatment group they allocate participant to • Collusion between the person who created the randomisation sequence and the recruiter (the person implementing the randomisation sequence)	(Jüni et al., 2001, Kjaergard et al., 1999, Moher et al., 1998, Schulz et al., 1995, Kennedy et al., 2017).
How envelopes are subverted	
• They can be opened in advance of recruitment • Transilluminating the envelopes to determine the allocation written inside • Assessing the weight and size of the envelopes to detect differences to determine which treatment group they allocate participants to • The recruiter can also create the envelope so they are aware of the allocation sequence	(Schulz, 1995, Jüni et al., 2001, Schulz and Grimes, 2002a, Viera and Bangdiwala, 2007).

1.1.4: Trifecta component part three: Robust reporting of allocation concealment

When reporting quality is discussed in this thesis, I am referring specifically to the reporting of allocation concealment methods. Poor reporting is reporting that lacks details to enable an assessment of the methodological quality of the implementation of allocation concealment. If an inadequate method had been selected for use or a method is implemented poorly, this should be stated. Any assessments made on the methodological quality of allocation concealment of an RCT depends on the completeness and accuracy of the reporting (Jüni et al., 2001). There are multiple ways that allocation sequences can be created and concealed, an issue when assessing the quality of their implementation is due to the plethora of ways that they can be described within publications. I have provided examples to illustrate this point (appendix 1, table 2).

It is imperative that the risk of selection bias is identified by those assessing the quality of an RCT as these should not contribute to the knowledge base without this bias being taken into account (Higgins et al., 2019). Only with transparent robust reporting can an accurate ROB assessment be undertaken, if methodological details are not reported they cannot be assessed.

1.1.5: Bringing the trifecta together

If selection bias is not identified there are a number of potentially negative outcomes. Although the research is seen to contribute to the knowledge base it could be doing so in a dangerous/negative way. At worst the results may be biased, conclusions reached may be inaccurate, but with no consequence other than wasted resource. Or, they may be misleading and result in unsafe conclusions on healthcare decisions or slow treatment progression (Jackson and Kuriyama, 2018). Systematic reviews combine results of several RCTs, if the trifecta is not fulfilled 'unclear' ROB assessments will be made, which do not reflect what occurred in the RCT. Appendix 2 details a worked example using a published RCT where the trifecta is not fulfilled.

1.2: Aims and objectives of this thesis

Aim:

This thesis explores the implementation and reporting of allocation concealment within RCTs, with the aim of identifying and promoting ways to detect and prevent selection bias in RCTs. I examine envelopes and blocking, which are common approaches used to conceal and/or randomise allocation sequences.

Objectives:

- Develop and pilot techniques for use when undertaking a systematic review, to detect selection bias caused by poor allocation concealment (Papers 1 and 2, appendix 3 and 4)
- Examine reporting of allocation concealment and develop a tool to improve the implementation and reporting of envelopes in RCTs (Papers 3 and 4, appendix 5 and 6)
- Explore the use of blocking in RCTs, develop empirical evidence of the insecurity of the frequently recommended small variable blocking schemes and detail how to perform and report blocking robustly (Paper 5, appendix 7)

In this chapter I discuss my research. The papers included were published between 2014-2021 and are presented in chronological order, the context, development and progression of each study to the next is highlighted. Each section starts with a screenshot of the relevant publication abstract or equivalent for on-line reader accessibility.

Section 2: Exploring the use of pooled baseline imbalances and heterogeneity in systematic reviews to detect selection bias

Paper 1: A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups. The full publication is presented in appendix 3.

ELSEVIER

CrossMark

Journal of Clinical Epidemiology 67 (2014) 1016–1024

Journal of Clinical Epidemiology

A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups

Laura Clark[a,*], Caroline Fairhurst[a], Catherine E. Hewitt[a], Yvonne Birks[b], Sally Brabyn[c], Sarah Cockayne[a], Sara Rodgers[a], Katherine Hicks[a], Robert Hodgson[a], Elizabeth Littlewood[c], David J. Torgerson[a]

[a]*Department of Health Sciences, York Trials Unit, University of York, York YO10 5DD, United Kingdom*
[b]*Department of Social Policy, University of York, York YO10 5DD, United Kingdom*
[c]*Department of Health Sciences, Mental Health and Addictions Research Group, University of York, York YO10 5DD, United Kingdom*

Accepted 23 April 2014; Published online 6 June 2014

Abstract

Background: There is evidence to suggest that component randomized controlled trials (RCTs) within systematic reviews may be biased. It is important that these reviews are identified to prevent erroneous conclusions influencing health care policies and decisions.

Purpose: To assess the likelihood of bias in trials in 12 meta-analyses.

Design: A review of 12 systematic reviews.

Data Sources: Twelve recently published systematic reviews with 503 component randomized trials, published in the *British Medical Journal*, *The Lancet*, *Journal of the American Medical Association*, and *The Annals of Internal Medicine* before May 2012.

Study Selection and Data Extraction: Systematic reviews were eligible for inclusion if they included only RCTs. We obtained the full text for the component RCTs of the 12 systematic reviews (in English only). We extracted summary data on age, number of participants in each treatment group, and the method of allocation concealment for each RCT.

Data Synthesis: Five of the 12 meta-analyses exhibited heterogeneity in age differences ($I^2 > 0.30$), when there should have been none. In two meta-analyses, the age of the intervention group was significantly greater than that of the control group. Inadequate allocation concealment was a statistically significant predictor of heterogeneity in one trial as observed by a metaregression.

Conclusions: Most of the sample of recent meta-analyses showed that there were signs of imbalance and/or heterogeneity in ages between treatment groups, when there should have been none. Systematic reviewers might consider using the techniques described here to assess the validity of their findings.

Keywords: Systematic review; Selection bias; Randomized controlled trials; Methods; Meta-analysis; Heterogeneity

2.1: Beginning this body of work

I began my research by considering the power of systematic reviews and meta-analyses in influencing policy and practice, the large resources required to conduct them, and the known importance of detecting whether the component RCTs are impacted by selection bias. I focussed on selection bias as I had previously undertaken research around allocation concealment and wanted to further progress this work (Clark et al., 2013c, Clark et al., 2013b). As discussed in section 1, poor allocation concealment can lead to selection bias impacting the results of single RCTs. I wanted to extend this and explore methods to detect selection bias in systematic reviews.

Systematic reviews and meta-analyses inform clinical practice, policy decisions and future research (Clarke et al., 2010, Glasziou et al., 2006). A systematic reviewer undertakes ROB assessment to determine if component RCTs are biased using a ROB tool with prespecified eligibility criteria (Page et al., 2016, Wood et al., 2008, Berger, 2005, Roberts and Ker, 2016, Cochrane consumer network, 2020, Propadalo et al., 2019). A range of ROB tools are available (Olivo et al., 2008), the risk of bias 2 (RoB2) tool is now used in Cochrane reviews (Sterne et al., 2019).

The risk of selection bias is assessed by determining how rigorously allocation concealment has been implemented (Propadalo et al., 2019, Higgins et al., 2019). Although the use of insecure allocation concealment methods may not result in subversion, the RCT will still be judged as at high risk of selection bias, as shown in Figures 3a and 3b.

To assess the methodological quality of an RCT, systematic reviewers usually only have the publication of an RCT to gain methodological data, demonstrating the challenge of distinguishing between reporting and conduct. The protocol for the RCT may not be available and information recorded in trial registrations may not be sufficiently detailed. Direct contact with authors may be attempted, however this is not always a successful strategy which I have experienced previously (Clark et al., 2013c).

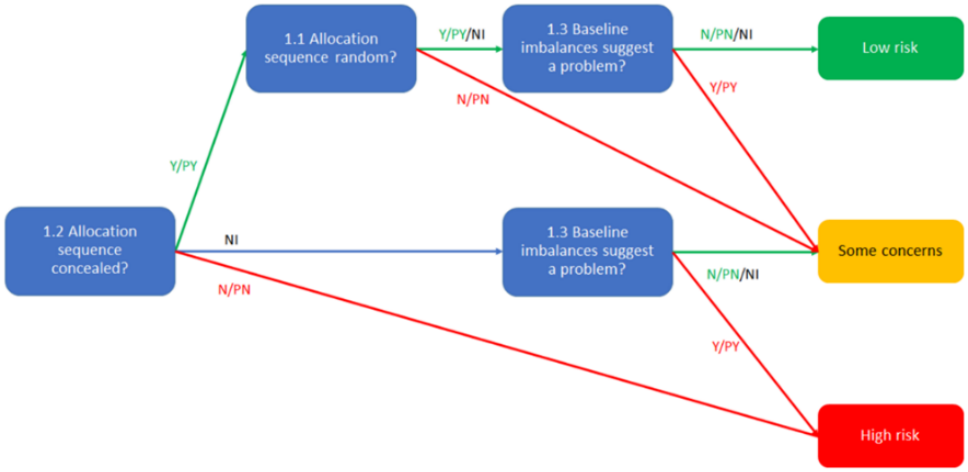
As allocation concealment methods are often not reported for an accurate ROB assessment to be undertaken, I wanted to explore an objective, complementary and low resource-intensive method that could be applied when conducting a systematic review which would detect whether component RCTs were impacted by selection bias. I started by examining the use of baseline imbalances and heterogeneity to detect selection bias in systematic reviews.

2.1.1: Systematic reviews: pooled baseline imbalances and heterogeneity to detect selection bias

Paper 1 developed as I considered the assessment of selection bias by examining baseline imbalances that occur in demographic variables between the treatment arms. There should be no difference at baseline if secure allocation concealment

has been implemented, except those occurring by random chance (Berger, 2009), therefore baseline testing is often not implemented at an individual RCT level (Senn, 1989, Senn, 1994). It is widely agreed that sample sizes and covariate values do not have to be equal across arms, but, large imbalances may be a symptom of poor randomisation and warrant investigation (Schulz and Grimes, 2002d). It has been suggested that baseline imbalances of important prognostic variables of component RCTs in systematic reviews should be assessed (Corbett et al., 2014). The Cochrane RoB2 now supports this; figure 4 demonstrates the algorithm the ROB tool suggests (Sterne et al., 2019).

Figure 4: Algorithm for suggested judgement of risk of bias arising from the randomisation process in the Cochrane ROB2 tool. Where Y = yes, N = no, PY = probably yes, PN = probably no, NI= not indicated (Sterne et al., 2019)



Imbalances would not be expected at baseline within a systematic review when all component RCTs are pooled together. However, pooled baseline imbalances are not currently assessed in systematic reviews. A case study observed the presence of pooled baseline imbalances in a prognostic variable in a systematic review (Trowman et al., 2006, Trowman et al., 2007), suggesting the component RCTs could be impacted by selection bias. I wanted to further build on Trowman's work by assessing baseline imbalances and additionally examining whether pooled heterogeneity could indicate the presence of selection bias, measured by the I^2 statistic. This would describe the variability between studies in systematic reviews

that is not due to random chance (Fletcher, 2007, Higgins and Thompson, 2002, Higgins et al., 2003).

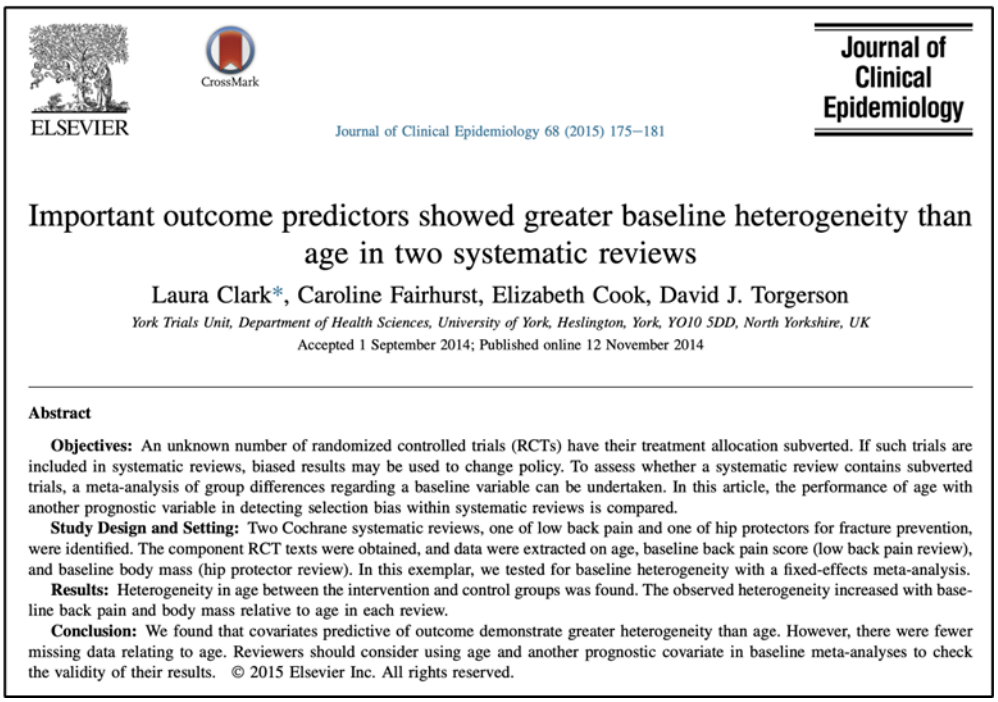
My interest in assessing heterogeneity is due to work undertaken by Trowman observing only a proportion of component trials in a systematic review were imbalanced favouring one group or another (Trowman et al., 2006). However, subversion can go in either direction; a proportion of component trials could have selection bias favouring both groups, so an imbalance may not necessarily be observed. This may cancel out the differences in the observed variable making the selection effects on an unknown variable unobservable. Therefore, a better measure of the problem is a measure of heterogeneity; if true secure allocation and concealment has been implemented there should be no heterogeneity.

2.2: Assessment of pooled heterogeneity and an extension

To pilot the technique of assessing pooled baseline heterogeneity I identified the key prognostic variable of age, frequently reported in 'table one' of an RCT publication (Furler et al., 2012). I identified 12 recently published systematic reviews in the Journal of the American Medical Association (JAMA), New England Journal of Medicine (NEJM), BMJ and the Lancet. These are high impact journals, therefore these systematic reviews are more likely (but not guaranteed) to have been well conducted. I found five systematic reviews exhibited heterogeneity of age (the I^2 statistic was >30%) (Higgins and Thompson, 2002), a meta-regression demonstrated that inadequate allocation concealment was a statistically significant predictor of heterogeneity in one systematic review indicating the systematic review conclusion may be incorrect.

Age may not always be an appropriate prognostic variable; it may not be reported or the component RCTs in a systematic review could be of a narrow range. The success of piloting the technique, questions I had regarding age being the most suitable prognostic, and reviewer comments supporting the development of assessing baseline heterogeneity, led to the development of paper 2.

Paper 2: Important outcome predictors showed greater baseline heterogeneity than age in two systematic reviews. The full publication is presented in appendix 4.



ELSEVIER

CrossMark

Journal of Clinical Epidemiology 68 (2015) 175–181

Journal of Clinical Epidemiology

Important outcome predictors showed greater baseline heterogeneity than age in two systematic reviews

Laura Clark*, Caroline Fairhurst, Elizabeth Cook, David J. Torgerson

York Trials Unit, Department of Health Sciences, University of York, Heslington, York, YO10 5DD, North Yorkshire, UK

Accepted 1 September 2014; Published online 12 November 2014

Abstract

Objectives: An unknown number of randomized controlled trials (RCTs) have their treatment allocation subverted. If such trials are included in systematic reviews, biased results may be used to change policy. To assess whether a systematic review contains subverted trials, a meta-analysis of group differences regarding a baseline variable can be undertaken. In this article, the performance of age with another prognostic variable in detecting selection bias within systematic reviews is compared.

Study Design and Setting: Two Cochrane systematic reviews, one of low back pain and one of hip protectors for fracture prevention, were identified. The component RCT texts were obtained, and data were extracted on age, baseline back pain score (low back pain review), and baseline body mass (hip protector review). In this exemplar, we tested for baseline heterogeneity with a fixed-effects meta-analysis.

Results: Heterogeneity in age between the intervention and control groups was found. The observed heterogeneity increased with baseline back pain and body mass relative to age in each review.

Conclusion: We found that covariates predictive of outcome demonstrate greater heterogeneity than age. However, there were fewer missing data relating to age. Reviewers should consider using age and another prognostic covariate in baseline meta-analyses to check the validity of their results.

2.3: Trial specific prognostic variables

In paper 2 I assessed whether a different prognostic variable would be more accurate and accessible to different trial types in two systematic reviews; baseline back pain score for a systematic review assessing low back pain and Body Mass Index (BMI) for a review assessing hip protectors. Having extracted data on age and the prognostic variable I assessed baseline heterogeneity with a fixed effect meta-analysis. I did not assess baseline P value distribution in paper 2 as a paper was published advising against this to assess the validity of randomisation (Bland, 2013). My research demonstrated a prognostic variable closely related to treatment outcome can show larger heterogeneity than age and may be more appropriate. For example, in the back pain review the I^2 was 22.0 for baseline age and 55.8 for backpain score.

However, a common trial-specific prognostic variable may not be reported in all component RCTs whereas age most likely will. A case-by-case decision must be made to determine the most appropriate pooled baseline variable to assess.

2.4: Why assessing baseline heterogeneity is desirable

Allocation concealment is considered so important that it was originally the only ROB item incorporated into RevMan used for Cochrane reviews (Hróbjartsson et al., 2013) and remains important today.

Poorly implemented allocation concealment enables subversion to occur on a baseline variable, the subversion can be subtle and not detected within a single RCT. In papers 1 and 2 I demonstrate that within an accumulated set of RCTs this subversion, and therefore selection bias can be detected, potentially leading to more rigorous conclusions.

Some researchers consider the ROB assessment suboptimal (Hartling et al., 2009). Stating the ROB tools do not go into enough detail or ask the correct questions, so insufficient information is extracted to enable reviewers to adequately or consistently judge allocation concealment methodologies (Propadalo et al., 2019, Jordan et al., 2017, Barcot et al., 2018). Additionally, systematic reviewers may not be able to ascertain the methodological information needed for an accurate ROB assessment or a systematic review may be poorly conducted and/or reported. Therefore, assessing the baseline pooled heterogeneity of a prognostic variable as an additional step enables an un-biased and objective way to assess the ROB of component RCTs in a systematic review.

Some trial designs, such as double-blind placebo controlled are considered more secure and harder to subvert. When quality assessed they are more likely to be classed as 'low risk of selection bias'. However it is known that subversion can still occur in these RCTs (Schulz, 1995). The technique described in papers 1 and 2 can be used to identify subversion where it is not apparent from the written description of trial methodologies. As found in a meta-epidemiological study that contained many placebo controlled trials - heterogeneity and imbalances were observed (Hulshof et al., 2019).

2.5: Impact of these exemplars: for the wider research community and this body of work

The exemplar presented in paper 1 was disseminated at the 2nd Clinical Trials Methodology Conference: Methodology Matters and was well received (Clark et al., 2013a). This work has been cited in other publications and prompted further work and explorations of the technique (Mitchell et al., 2019, Hicks et al., 2018, Hulshof et al., 2019).

The component RCTs included in papers 1 and 2 lacked methodological information needed to make judgements on allocation concealment and age was not always reported. This is basic information that should be available in RCT publications. These observations led to my line of research enquiry for paper 3.

2.6: Conclusion

Papers 1 and 2 highlighted that systematic reviews may contain RCTs which are impacted by selection bias, these could impact the reviews' conclusions. ROB assessments should continue to evolve building on the original ROB methodologies by taking into account empirical evidence, experiences of systemic reviewers and authors (Higgins et al., 2011). Systematic reviews are resource intensive, costly and change patient care; the results need to be as accurate as possible. This complementary technique has the advantage of not relying on subjective assessments of bias, which could be incorrect. It is recommended that assessments should be made on age as this is usually reported and has fewer missing data, but, where possible an appropriate and important prognostic variable should also be investigated.

Section 3: Allocation concealment methods and reporting

Paper 3: Allocation concealment in randomised controlled trials: are we getting better? The full publication is presented in appendix 5.

thebmj

BMJ 2016;355:i5663 doi: 10.1136/bmj.i5663 (Published 17 November 2016)

ANALYSIS

Allocation concealment in randomised controlled trials: are we getting better?

Laura Clark and colleagues assess the allocation concealment methods in a sample of randomised controlled trial publications

Laura Clark *research fellow*, Caroline Fairhurst *research fellow*, David J Torgerson *director*

York Trials Unit, University of York, York YO10 5DD, UK

A robust randomised controlled trial (RCT) must use allocation concealment—that is, separate the act of randomisation from the person recruiting participants. Poor randomisation methods cause exaggerated treatment effects, are open to subversion by researchers or clinicians, and have a knock-on effect on systematic reviews.[1-3]

The CONSORT statement, which leading medical journals endorse, states that the method of allocation (comprising sequence generation, allocation concealment mechanism, and implementation) should be clearly described.[4] Allocation concealment is dependent on the method of sequence generation as well as the concealment mechanism.

Almost a fifth of trials published in major medical journals in 2002 used inadequate concealment, and a quarter failed to describe how the allocation was concealed.[2] Here we examine a sample of RCTs published in 2015 to see whether the situation has improved.

Defining inadequate allocation concealment

We searched four high impact medical journals (*The BMJ*, *Journal of the American Medical Association (JAMA)*; the *Lancet*, and the *New England Journal of Medicine (NEJM)*) and found 79 RCTs published between June and August 2015. We extracted and judged their mechanism for allocation concealment, taking into consideration the study design, sequence generation method, and allocation concealment mechanism. We defined an inadequate process as one that used envelopes as the method of allocation concealment (box 1) or used stratified block randomisation by site with small block sizes as the sequence generation method (box 2), except in double blind trials. If insufficient detail was provided in the paper, we checked the protocol or emailed the authors.

Fifteen trials were poorly randomised

Twenty seven (34%) of the RCTs were placebo controlled double blind trials, in which allocation is generally well concealed; participants are assigned a number corresponding to a packet of drugs, and only the pharmacist has access to the unblinding codes. One of these trials used envelopes, but as the pharmacist opened the envelopes after the clinician enrolled the participant it was deemed adequate. We judged these trials, and 22 (28%) of the remaining trials, as adequate. We initially found that 13 trials (16%) had used a randomisation method that put them at risk of bias and 17 trials (22%) contained insufficient detail to determine whether the method of allocation concealment was adequate. We received more information from the authors of nine (53%) of these trials, two of which were found to be inadequate, giving a total of 15 (19%) trials with inadequate concealment (table 1⇓). Seven trials used envelopes to allocate participants, seven trials used small block sizes and/or stratified by site, and one used both small blocks and envelopes.

We noted two inconsistencies with the use of block randomisation. The trial by Senn et al had an imbalance of 12 participants between the randomised groups; the largest possible imbalance for a block size of two stratified by three centres is three. Correspondence with the authors confirmed that 10 cases were misallocated to the intervention group.[26] The authors said that a combination of technical and human errors accounted for the imbalance—"the laptop froze during randomisation, the server went down temporarily, research assistants inadvertently practiced on a live site, and participants went to the wrong session." They reassigned the 10 participants to their originally assigned groups and found no change in benefit for the intervention.

The trial by Cox and colleagues[27] said that patients were allocated 2:1 with blocks of four, but with a 2:1 ratio the block size should be divisible by three. Correspondence with the author confirmed that the statement in the paper was a mistake, and an erratum has been published in the *Lancet*.

More rigour is needed

Our findings—that 19% of trials described inadequate methods of allocation concealment and 22% failed to report their

Correspondence to: laura.clark@york.ac.ukAccepted 18 October 2016

3.1: Allocation concealment implementation and reporting

During the research conducted for papers 1 and 2, I found reporting of the allocation concealment methods to be poor, thus the allocation concealment trifecta was not fulfilled. In paper 1 I found 20% of the included RCTs reported an adequate allocation concealment method and 4% an insecure method. In 76% of the RCTs the reporting was so poor the allocation concealment method employed was unclear. As a result it was uncertain what method was used and/or how robustly the allocation concealment was implemented as the RCT lacked the necessary methodological information. Many component RCTs were published before or shortly after the CONSORT statement was published, therefore reporting quality may have been poorer than in more recently published RCTs.

The aim of paper 3 was to ascertain how adequately allocation concealment was implemented in more recent RCTs and determine whether the poor implementation and reporting of allocation concealment observed in papers 1 and 2 was still an issue in more recent publications.

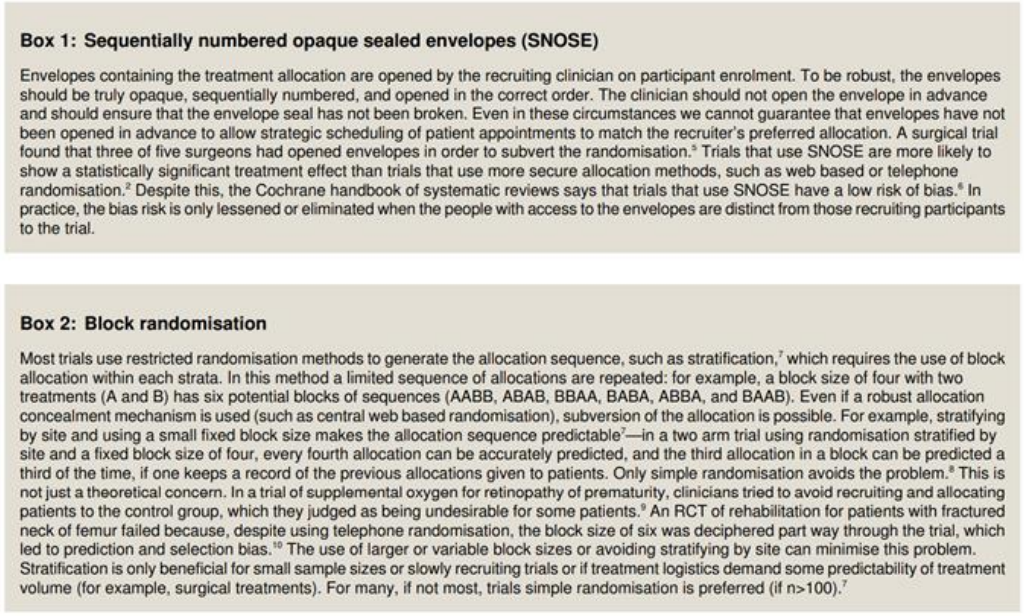
I identified RCTs published over a three-month period in 2015 in JAMA, BMJ, NEJM and the Lancet. These journals have the highest impact factors and would enable an assessment on the 'best practice' of reporting completeness as they promote the use of reporting guidelines in their instructions to authors. I extracted methodological data and contacted authors when clarifications were needed. Finally, I suggested ways to improve allocation concealment method selection, implementation and reporting.

3.2: Allocation concealment quality

Stringent assessment criteria of allocation concealment as applied in paper 3 (figure 5) is arguably necessary. Allocation concealment implementation needs to be robust to ensure the allocation can be deemed random, the defining methodological feature of an RCT. Readers, reviewers and journals should be demanding to know how envelopes were prepared and who opened them, not just accepting the statement that 'sequentially numbered opaque sealed envelopes' (SNOSE) were used. Similarly, with blocking – it is not acceptable to be

performing research where it is possible to accurately predict the allocation sequence as an insecure block size has been used. The idea of designing research that has a predictable allocation sequence is preposterous yet in paper 3 I demonstrate that this is still happening.

Figure 5: Assessment criteria for trials using envelopes and block randomisation taken from paper three



Box 1: Sequentially numbered opaque sealed envelopes (SNOSE)

Envelopes containing the treatment allocation are opened by the recruiting clinician on participant enrolment. To be robust, the envelopes should be truly opaque, sequentially numbered, and opened in the correct order. The clinician should not open the envelope in advance and should ensure that the envelope seal has not been broken. Even in these circumstances we cannot guarantee that envelopes have not been opened in advance to allow strategic scheduling of patient appointments to match the recruiter's preferred allocation. A surgical trial found that three of five surgeons had opened envelopes in order to subvert the randomisation.[5] Trials that use SNOSE are more likely to show a statistically significant treatment effect than trials that use more secure allocation methods, such as web based or telephone randomisation.[2] Despite this, the Cochrane handbook of systematic reviews says that trials that use SNOSE have a low risk of bias.[6] In practice, the bias risk is only lessened or eliminated when the people with access to the envelopes are distinct from those recruiting participants to the trial.

Box 2: Block randomisation

Most trials use restricted randomisation methods to generate the allocation sequence, such as stratification,[7] which requires the use of block allocation within each strata. In this method a limited sequence of allocations are repeated: for example, a block size of four with two treatments (A and B) has six potential blocks of sequences (AABB, ABAB, BBAA, BABA, ABBA, and BAAB). Even if a robust allocation concealment mechanism is used (such as central web based randomisation), subversion of the allocation is possible. For example, stratifying by site and using a small fixed block size makes the allocation sequence predictable[7]—in a two arm trial using randomisation stratified by site and a fixed block size of four, every fourth allocation can be accurately predicted, and the third allocation in a block can be predicted a third of the time, if one keeps a record of the previous allocations given to patients. Only simple randomisation avoids the problem.[8] This is not just a theoretical concern. In a trial of supplemental oxygen for retinopathy of prematurity, clinicians tried to avoid recruiting and allocating patients to the control group, which they judged as being undesirable for some patients.[9] An RCT of rehabilitation for patients with fractured neck of femur failed because, despite using telephone randomisation, the block size of six was deciphered part way through the trial, which led to prediction and selection bias.[10] The use of larger or variable block sizes or avoiding stratifying by site can minimise this problem. Stratification is only beneficial for small sample sizes or slowly recruiting trials or if treatment logistics demand some predictability of treatment volume (for example, surgical treatments). For many, if not most, trials simple randomisation is preferred (if n>100).[7]

3.3: Envelopes and blocking

I found that 19% (n=15) of the RCTs I assessed in paper 3 used inadequate allocation concealment methods. Seven used envelopes, seven blocking and one used both. The issue observed with blocking was frequently implementation – small block sizes were selected, increasing allocation sequence prediction (Berger, 2005). Here, part one of the allocation concealment trifecta is not fulfilled.

With envelopes, the issues observed were both poor implementation (such as the recruiter also creating the envelope) and reporting. Frequently insufficient information was provided for a judgement to be made regarding how envelopes were implemented, therefore it was unknown which part(s) of the trifecta were/were not fulfilled. Envelopes could have been implemented robustly but the

reporting lacked detail to determine this (inadequate reporting). Or, envelopes were not implemented robustly and this information was omitted from the publication (inadequate implementation and reporting). Either scenario contributes to research waste and could slow down the increase to the knowledge base (Sterne et al., 2019).

When contacting authors for further information, only 53% responded. Some author feedback demonstrated the lack of methodological details included in their publications was not poor allocation concealment implementation but rather poor reporting. Of the remaining 47%, some RCTs may be of a low ROB but due to poor reporting this cannot be determined.

Some authors clarified observed inconsistencies with blocking (Senn et al., 2015, Cox et al., 2015) and envelope use (Salminen et al., 2015). When these perceived errors were clarified, although positive, this should not have occurred if transparent reporting, a thorough reviewing process - including review by a methodologist - had happened prior to publication. These basic methodological errors should have, and could have, been identified within the publication process (see section 6.1).

3.4: Improving allocation concealment trifecta- implementation and reporting

Paper 3 demonstrates the implementation and reporting of allocation concealment is suboptimal and the trifecta is frequently not fulfilled. The research provided data that enabled me to make suggestions on how to improve this and enabled the identification of two key areas- envelopes and blocking- which were frequently implemented and reported sub-optimally. Suggestions were based on discussion with colleagues and students during teaching sessions I undertook as well as practical experience gained from working in the area of trials.

One way to ignite change is from the top down – unacceptable methodologies should not be funded. Sponsors and Research Ethics Committees should scrutinise research proposals to check that they are methodologically robust. If envelopes or small block sizes are suggested to be used, this should be questioned and checked to ensure the appropriateness, central randomisation should be considered where

possible (Swingler and Zwareenstein, 2000). Of course, there needs to be allowances for the use of envelopes within some research designs and settings (Doig and Simpson, 2005, Swingler and Zwareenstein, 2000), considered further in section 4. Journal editors and peer reviewers need to ensure that allocation concealment is reported transparently. Section 6.1 further discusses the role authors and journals in improving reporting of allocation concealment methodology.

3.5: Impact and development

It is of note that paper 3 has been cited in key methodological texts including the international Cochrane Handbook for Systematic Reviews of Interventions Chapter 8 - 'Assessing risk of bias in a randomized trial' (Higgins et al., 2019). Following the publication of paper 3, online comments were published:

'This report [paper 3] provides yet one more very important revelation about the reality of the randomized controlled trial (RCT). Now we find the "randomized" in RCT is up for grabs, leaving only "controlled" …the RCT is the scientific "gold standard" for evidence-based medicine, implementation errors and manipulated reporting of results make it appear to be medicine's "fool's gold." Mounting evidence from every which way points to the sham that has become the RCT...' (responses, 2016).

These comments, the need for guidance and evidence to improve the implementation of allocation concealment, alongside the dire nature of the reporting led me to further develop methods around allocation concealment, reported in papers 4 and 5, as a way of improving use of both envelopes and blocking where these are the appropriate method of choice.

3.6: Conclusions and recommendations: where is change needed

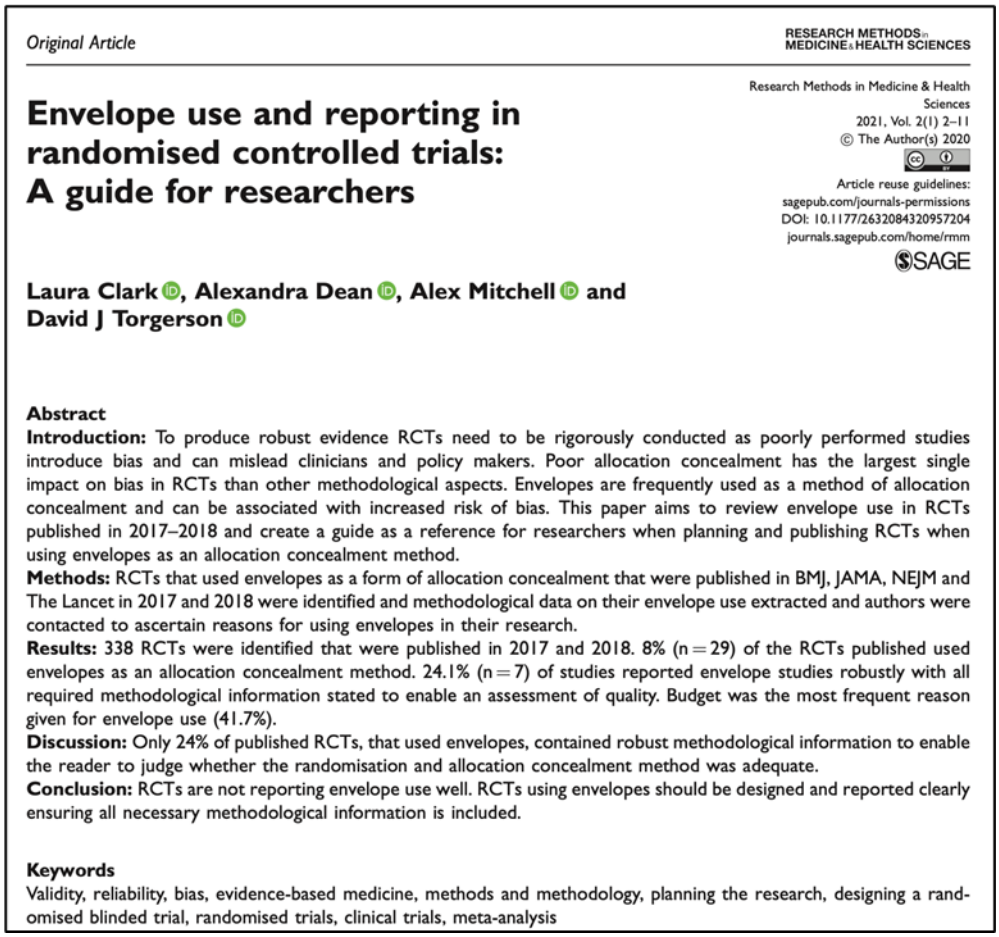
Despite the narrow selection of journals reviewed, it is clear there was poor reporting of essential allocation concealment implementation methods, even when CONSORT had become established. Component RCTs in systematic reviews

frequently come from low impact journals that have poorer reporting quality, accurate ROB assessments therefore cannot be made.

The research waste that suboptimal RCTs cause is almost minor compared to the effects that poor research could be having at the patient level, where beneficial treatments could be prevented from reaching patients or harmful treatments not stopped (Glasziou and Chalmers, 2018). In summary - allocation concealment in randomised controlled trials: we *must* do better.

Section 4: The implementation and reporting of envelopes as a form of allocation concealment

Paper 4: Envelope use and reporting in randomised controlled trials: A guide for researchers. The full publication is presented in appendix 6.



Original Article

RESEARCH METHODS IN MEDICINE & HEALTH SCIENCES

Research Methods in Medicine & Health Sciences
2021, Vol. 2(1) 2–11
© The Author(s) 2020

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2632084320957204
journals.sagepub.com/home/rmm

SAGE

Envelope use and reporting in randomised controlled trials: A guide for researchers

Laura Clark, Alexandra Dean, Alex Mitchell and David J Torgerson

Abstract

Introduction: To produce robust evidence RCTs need to be rigorously conducted as poorly performed studies introduce bias and can mislead clinicians and policy makers. Poor allocation concealment has the largest single impact on bias in RCTs than other methodological aspects. Envelopes are frequently used as a method of allocation concealment and can be associated with increased risk of bias. This paper aims to review envelope use in RCTs published in 2017–2018 and create a guide as a reference for researchers when planning and publishing RCTs when using envelopes as an allocation concealment method.

Methods: RCTs that used envelopes as a form of allocation concealment that were published in BMJ, JAMA, NEJM and The Lancet in 2017 and 2018 were identified and methodological data on their envelope use extracted and authors were contacted to ascertain reasons for using envelopes in their research.

Results: 338 RCTs were identified that were published in 2017 and 2018. 8% (n = 29) of the RCTs published used envelopes as an allocation concealment method. 24.1% (n = 7) of studies reported envelope studies robustly with all required methodological information stated to enable an assessment of quality. Budget was the most frequent reason given for envelope use (41.7%).

Discussion: Only 24% of published RCTs, that used envelopes, contained robust methodological information to enable the reader to judge whether the randomisation and allocation concealment method was adequate.

Conclusion: RCTs are not reporting envelope use well. RCTs using envelopes should be designed and reported clearly ensuring all necessary methodological information is included.

Keywords

Validity, reliability, bias, evidence-based medicine, methods and methodology, planning the research, designing a randomised blinded trial, randomised trials, clinical trials, meta-analysis

4.1: Envelopes and RCTs

Envelopes should contain a truly random allocation sequence and the person who creates them should be separate to the person who opens the envelopes and recruits participants into the trial. All of which should be reported to fulfil the trifecta. I therefore decided to undertake research to address improving the implementation and reporting of this allocation concealment method.

Envelope use can be confusing to those who do not have detailed knowledge of allocation concealment. Conflicting messages state they are acceptable if implemented robustly (Doig and Simpson, 2005, Swingler and Zwareenstein, 2000) but they are also associated with subversion and increased effect sizes

(Hewitt et al., 2005, Kennedy et al., 2017, Peto, 1999, Mitchell et al., 2019). Additionally, a prevalent online randomisation service is called 'sealed envelope', which, further implies their robustness as it normalises the use of the phrase 'sealed envelope', associating envelopes with an acceptable allocation concealment method (Envelope, 2022).

Envelope use is not new, the first RCT published in 1948 employed envelopes (Marshall et al., 1948). However, the danger of using envelopes has been known for over 20 years (Schulz, 1995). Technology and allocation concealment methods have moved on but envelopes are still a popular method despite the known problems (section 1, table 1). Throughout my research career, my views have changed over time on the appropriateness of the use of envelopes as an allocation concealment method. In paper 4 I explored why envelopes are chosen despite their known insecurities. I also utilised my practical and methodological experience on the robust implementation and reporting of envelope use to explore a pragmatic way to improve this within research.

In paper 4 my research began with surveying a group of authors who had recently published RCTs using envelopes to understand their reasons for selecting this method, and to assess the quality of the reporting and implementation in these publications. The 338 RCTs included in paper 4 were published in 2017-2018 and 8.5% used envelopes, a review conducted in 2015 observed similar use (9%) (Yelland et al., 2018). In comparison, a review conducted in 2001 found 16.6% of RCTs used envelopes (Hewitt et al., 2005). It is encouraging that envelope use is steadily decreasing. My survey determined that envelopes are still used and will continue to be used for pragmatic reasons, however assessment of implementation and reporting demonstrated their use needs improving. The survey results were used to formulate recommendations to ensure the trifecta is fulfilled.

4.2: Exploring envelopes: implementation and reporting

There was the common misconception both observed in my research in paper 4 and the wider literature that envelopes maintain optimal blinding (CONSORT, 2022a, Higgins et al., 2019), or 'the small size of the study' meant that envelopes

were the appropriate method – even small studies can employ robust central randomisation methods. One author stated the need for rapid randomisation was a reason for selection. Here, the randomisation needed to occur within hours, rapid randomisation normally refers to allocation within minutes, randomisation within hours could have accommodated a central randomisation approach.

Security measures are important as they add an additional barrier to prevent the envelope being tampered with and provide more confidence that the allocation concealment has been implemented robustly. Paper 4, found only one RCT reported outstanding security measures going beyond SNOSE to conceal the allocation sequence (Boden et al., 2018), further suggesting a lack of knowledge around envelope execution and reporting. Additional security measures will increase the costs and time taken to prepare the envelope, but only marginally. They will however be more likely to result in robust high-quality research, if they are reported well.

There was a lack of knowledge regarding the necessary factors to report when reporting envelope use, shown with only 7% of trials adequately reporting envelopes. It could be that the guidance for reporting envelopes may not be explicit enough (Schulz and Grimes, 2002a). There needs to be more awareness among researchers to understand the impact of not reporting envelopes adequately; it impacts ROB assessments. As seen in paper 4, where two RCTs described the same overarching trial but one did so adequately the other omitted important information needed to make a full judgement (Nadkarni et al., 2017, Patel et al., 2017).

Budget was found to be a significant driver for selecting envelopes as an allocation concealment method. Rather than budgeting for a more secure method such as central randomisation, researchers are risking implementing an insecure method that could introduce selection bias into their research rendering the research at a high ROB. This could be due to a lack of understanding around the insecurity of using envelopes. Although only 8.5% of studies used envelopes in paper 4, this represents a significant amount of money being invested within these research studies, highlighting how essential improving the robust implementation and reporting of envelopes is to fulfil the trifecta.

Paper 3 received the following online comment when published:

'Having read this article [paper 3] I have to confess yet again disappointment with the UK research thinking. Of course we should constantly strive to improve our research methodology and using technology to change from 'sealed envelopes' to 'independent electronic randomisation' is to be encouraged. But why do the UK research organisation not provide such a simple service for free to researchers…'

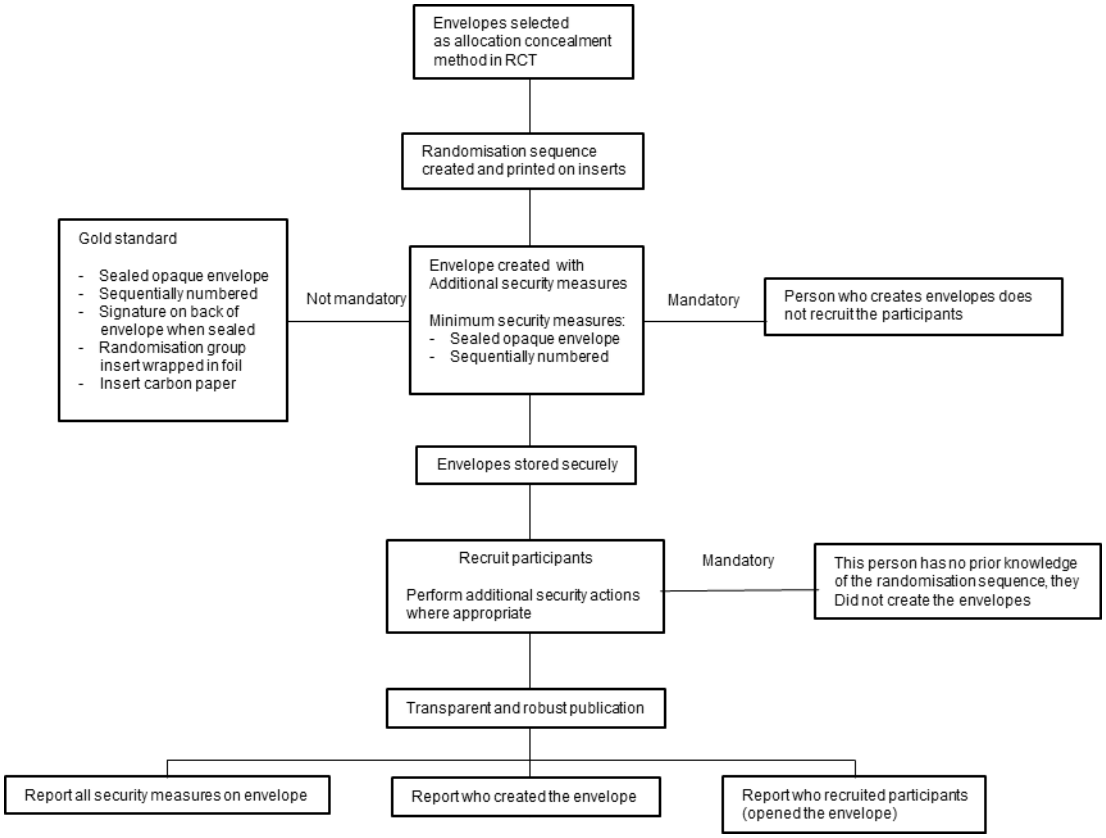
Although a free randomisation service would be a welcome addition to researchers, it would not be the panacea. There are justified reasons why envelopes are implemented, and in some situations most likely the only feasible option, such as in emergency medicine. Applying a blanket approach that central randomisation only should be undertaken would result in RCTs not being performed within some settings, as this method cannot always be accommodated (Swingler and Zwareenstein, 2000). It would be unethical to not support an RCT to be performed using envelopes in these situations, as the alternative would be research not being undertaken, or a less robust research design being performed (Murad et al., 2016). Encouragingly Swingler et al found that robust envelope use is possible with scrupulous record keeping and monitoring (Swingler and Zwareenstein, 2000). The key is that if envelopes are selected they are implemented and reported robustly.

4.3: Envelopes: time to change or time for acceptance and guidance?

Having established that there are practical reasons why envelopes will continue to be used as an allocation concealment method, this situation should be accepted and the shortcomings addressed.

I analysed the survey results to produce recommendations in the form of a tool to aid researchers to implement and report envelope use robustly (figure 6). Journal editors and peer reviewers could also use these criteria to ensure that envelopes are reported well.

Figure 6: tool developed stating methodological steps to creating a robust envelope as an allocation concealment method in an RCT and essential methodological information to be reported in the publication of an RCT

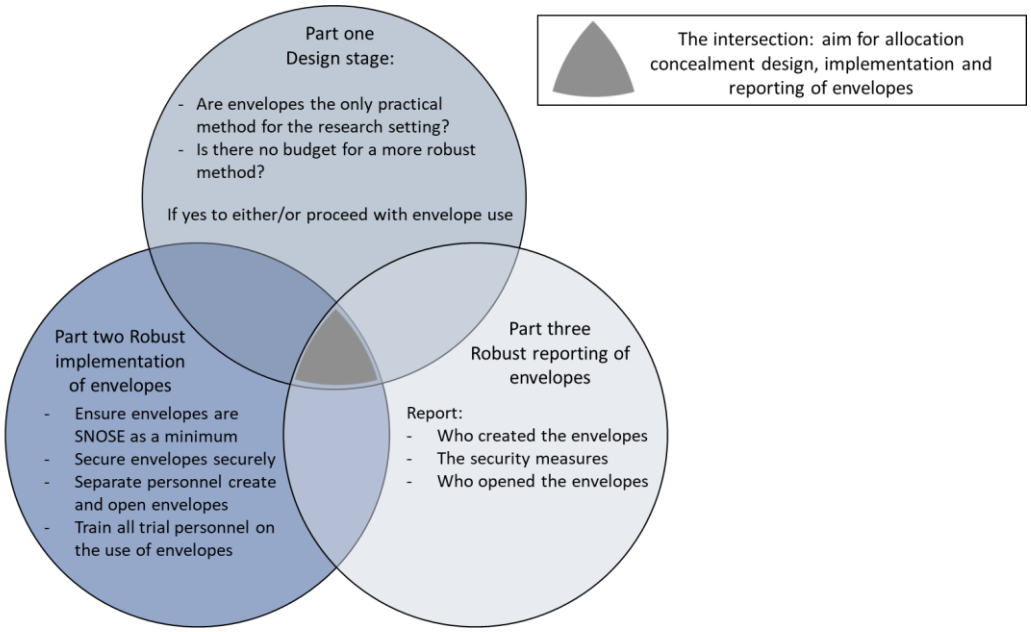


If envelopes are used within placebo-controlled double-blind trials they may be harder to subvert but subversion can occur (CONSORT, 2022a, Higgins et al., 2019). In a systematic review, the placebo-controlled double-blind RCT is likely to be judged as a low ROB; an example of why assessing pooled baseline heterogeneity as detailed in Section 1 is useful.

4.4: Envelopes and the trifecta challenge

Envelopes pose a challenge, they are an allocation concealment method more likely to be subverted (section 1.1.3) and may not be considered a robust method. However, they can be implemented and reported well. Figure 7 poses the allocation concealment trifecta specific to envelope use.

Figure 7: Allocation concealment trifecta for envelopes



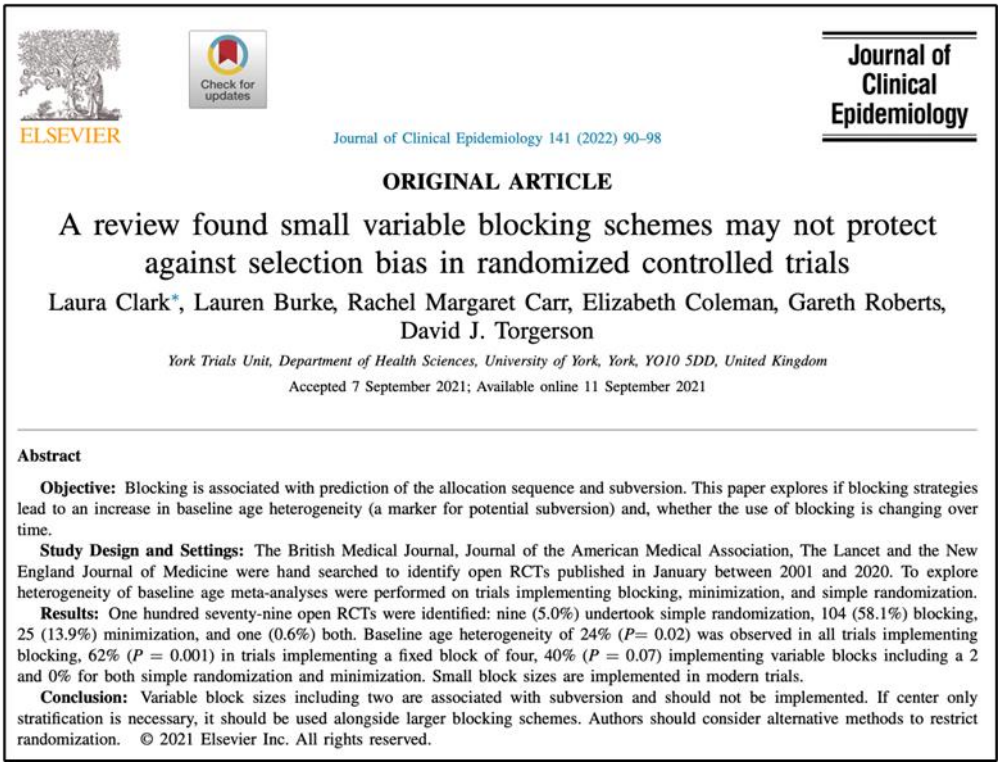
4.5: Conclusion

This section summarises and contextualises the findings of paper 4; envelopes are going to continue to be implemented as a form of allocation concealment. Pragmatically each research proposal needs considering on a case-by-case basis by funders and stakeholders and the appropriateness of envelope use scrutinised.

If selected, envelopes need to be implemented and reported robustly with all members of the research team educated regarding their use to ensure the allocation concealment trifecta is fulfilled. As there is extraordinary human curiosity and ingenuity shown in deciphering the allocation sequences, it is recommended that envelopes should not be used if it is feasible to implement a more secure allocation concealment method.

Section 5: Blocking and prediction

Paper 5: A review found small variable blocking schemes may not protect against selection bias in randomized controlled trial. The full publication is presented in appendix 7.



ELSEVIER

Journal of Clinical Epidemiology 141 (2022) 90–98

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

A review found small variable blocking schemes may not protect against selection bias in randomized controlled trials

Laura Clark*, Lauren Burke, Rachel Margaret Carr, Elizabeth Coleman, Gareth Roberts, David J. Torgerson

York Trials Unit, Department of Health Sciences, University of York, York, YO10 5DD, United Kingdom

Accepted 7 September 2021; Available online 11 September 2021

Abstract

Objective: Blocking is associated with prediction of the allocation sequence and subversion. This paper explores if blocking strategies lead to an increase in baseline age heterogeneity (a marker for potential subversion) and, whether the use of blocking is changing over time.

Study Design and Settings: The British Medical Journal, Journal of the American Medical Association, The Lancet and the New England Journal of Medicine were hand searched to identify open RCTs published in January between 2001 and 2020. To explore heterogeneity of baseline age meta-analyses were performed on trials implementing blocking, minimization, and simple randomization.

Results: One hundred seventy-nine open RCTs were identified: nine (5.0%) undertook simple randomization, 104 (58.1%) blocking, 25 (13.9%) minimization, and one (0.6%) both. Baseline age heterogeneity of 24% (P= 0.02) was observed in all trials implementing blocking, 62% (P = 0.001) in trials implementing a fixed block of four, 40% (P = 0.07) implementing variable blocks including a 2 and 0% for both simple randomization and minimization. Small block sizes are implemented in modern trials.

Conclusion: Variable block sizes including two are associated with subversion and should not be implemented. If center only stratification is necessary, it should be used alongside larger blocking schemes. Authors should consider alternative methods to restrict randomization.

5.1: Restricted randomisation, blocking and RCTs

In this section I report how I continued with the strategy of developing a way to improve allocation concealment implementation and reporting in areas of need. I considered restricted randomisation as this is used in 90% of studies (Hewitt and Torgerson, 2006). Blocking is the main way that this is achieved (Kahan et al., 2015), and observed to be implemented and reported sub-optimally throughout papers 1-3. Blocked randomisation schemes can also be executed with envelopes. Therefore, developing a strategy to ensure blocking is implemented and reported robustly was a logical progression in my work. It had been proposed that random block sizes decrease rather than abolish the risk of predicting allocation sequences (Matts and Lachin, 1988, Schulz and Grimes, 2002d). The purpose of paper 5 was to explore this, investigate the vulnerability of some blocking schemes to

subversion and provide practical advice to implement robust blocking rather than advise against its use.

Blocking is often implemented to ensure balance between treatment groups, however 'with great balance comes great predictability' (van der Pas, 2019). There is a known risk of prediction when blocking is employed and a misconception that this prediction is associated with small fixed block sizes with stratification based on centre only, and, that variable blocking schemes mitigate this issue. Blocking is often used as it is practical and, like envelopes, will continue to be used within research.

I applied the technique of assessing pooled heterogeneity of age undertaken in papers 1 and 2 to explore whether this could be used to identify evidence of subversion amongst trials using blocking. This method was chosen due to the positive reception of these papers, the method is low cost, easy to undertake and the variable age should be reported in every RCT, so was practical. I undertook analysis on different block sizes. I demonstrated variable blocking schemes, that included a block size of two within the blocking scheme were associated with increased heterogeneity and therefore subversion of the allocation sequence. This challenges the belief that small variable blocking schemes are a safeguard against subversion. I found that larger variable blocking schedules have zero heterogeneity showing that there are robust ways to implement blocking. My research indicates that where the use of blocking is appropriate and justified, strategies to ensure robust implementation are possible and should be developed.

My research for paper 5 began by performing a scoping review that covered a 20-year period to examine the changes to the use of blocking, the types of permutations used and completeness of reporting. I found recommendations around robust implementation and reporting are relevant to modern research as insecure blocking schemes had been recently implemented. To disseminate this, I produced a methodological report advising caution to researchers using blocked randomisation when designing their research, the first report to be supported by empirical evidence.

5.2: Building blocks for the future

A relevant question to ask is why the focus has not been on advocating for the use of simple randomisation as this is considered more robust than restricted (Hewitt and Torgerson, 2006). Although simple randomisation can lead to imbalances between groups at baseline this can be accounted for within the statistical analysis, and loss of power only occurs at greater than 2:1 imbalance (Hewitt and Torgerson, 2006). However, simple randomisation is rarely used in practice (Hewitt and Torgerson, 2006, Kahan and Morris, 2012, Kahan et al., 2015). Although many alternative methods are available to restrict randomisation, some are in their infancy such as pairwise randomisation (Fairhurst et al., 2020), others are expensive and prone to technical issues (minimisation) (Brown et al., 2005) and others still at the simulation stage (merged block randomisation) (van der Pas, 2019). To create a positive change and improve allocation concealment implementation it is sensible to be pragmatic and focus on what is used in practice and is acceptable to researchers.

5.3: Challenging beliefs

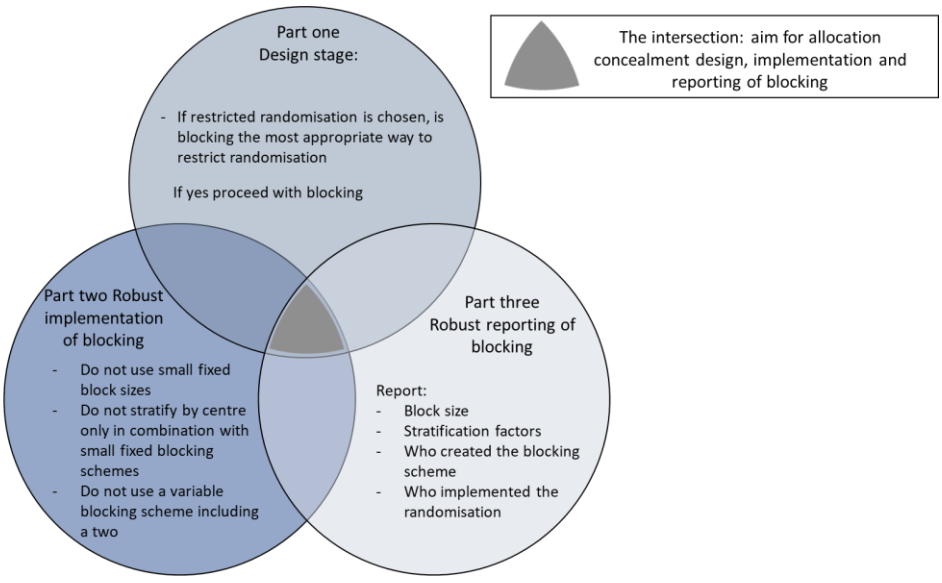
It is well reported that small fixed blocking schemes and stratification by centre should be avoided. In paper 5 a fixed block of four was found to be associated with increased significant heterogeneity (62%). This was not only expected, due to the known prediction of the allocation sequence associated with this block size, but confirmed that the analysis undertaken in paper 5 was identifying expected heterogeneity and the use was appropriate.

It was encouraging that there were limited small fixed blocks used within the data set and no fixed blocks of four were used in conjunction with centre stratification. This suggests the message of avoiding centre stratification and small fixed blocks has filtered through the research community. Variable blocking schemes that include a block of two within the scheme, were used more frequently in recently published RCTs. These schemes were associated with a heterogeneity of 40%, with centre stratification slightly increasing the heterogeneity to 42%. This finding indicates the vulnerability of a block size of two being incorporated within a scheme irrespective of central stratification and the importance that this finding is

disseminated. The European medicines agency recommend randomising each centre separately (CHMP, 2015), it is plausible to assume researchers could follow this advice. They may also implement the safeguard of using small variable blocking schemes (including a block size of two) to decrease selection bias, my research suggests the contrary could occur.

If central randomisation is implemented in conjunction with blocking, essential methodological details need to be reported. If blocking with small fixed sizes or block sizes including a two is undertaken and/or the stratification factor not reported, a full ROB assessment cannot be performed. Typically, if an RCT had undertaken central randomisation it would be considered a low ROB (Higgins et al., 2019). My research demonstrates this could be incorrect. Highlighting again that assessing baseline heterogeneity, as detailed in papers 1 and 2, is important as an objective way to detect selection bias. To achieve the trifecta, figure 8 summarises the steps needed when using blocking.

Figure 8: Allocation concealment trifecta for blocking



A concern that may arise regarding the recommendation that block sizes of two are avoided, is that larger block sizes are more likely to create a mid-block inequality. This would ensue if a treatment occurs at greater frequency earlier within the block, there is an interim analysis or the trial ceases recruiting part way through a block (Efird, 2011). This is one reason why small blocking schemes of random sizes which likely would include a block size of two, have been suggested

historically (Schulz, 1995, Schulz and Grimes, 2002c), but which my research findings caution against. Alternatively, large random blocking schemes can be implemented and the chance of initial treatment runs in a block offset by allocating participants using a biased coin approach (Efird, 2011). When using random sized blocks in a multicentre study the sample size may vary by site but on average be similar, therefore is not considered problematic (Efird, 2011).

5.4: Conclusion

My research highlights that although blocking can be a robust method when larger blocking schemes are implemented, it is vulnerable to subversion when a blocking scheme including a block size of two or small fixed blocks are selected. An accurate ROB assessment is not always possible as methodological information related to blocking is frequently missing in RCT publications. RCTs examined in paper 5 were from CONSORT endorsing journals, however, wider reporting which includes journals which do not endorse the CONSORT statement are most likely to be of poorer quality.

It is probable that restricted randomisation will continue to be used as larger multicentre trials are performed. The insecurity of variable blocking schemes using a block size of two needs to be promoted. Strategies to enable robust restricted randomisation methods to be implemented need to be developed, researched and tested.

Section 6: Discussion and future research

This section presents a discussion around my thesis, how I have met my objectives, considers the stakeholders my work impacts, reflects on effective dissemination strategies and presents future research questions.

6.1: The trifecta: why is it rarely fulfilled?

My research has shown that the allocation concealment trifecta is commonly not fulfilled. It is however challenging to ascertain whether poor reporting of allocation concealment is also indicative of poor implementation.

Journals are the primary vehicles for disseminating RCT findings and the main data source for systematic reviews. Poor reporting quality in journals was observed throughout my research. There is no shortage of support and reporting guidance (Altman et al., 2008), rather, the issue is poor adherence to the guidelines. Not all journals endorse the CONSORT statement, this is a frustrating situation when evidence demonstrates adherence to CONSORT improves completeness of reporting of RCTs (Hopewell et al., 2010, Moher et al., 2001, Hopewell et al., 2008, Altman, 2005, Turner et al., 2012, Xu et al., 2008, del Giglio and Costa, 2004).

Journals could prevent inadequate reporting of allocation concealment by adding safeguarding steps into their publication process. Editors and journals need to encourage factually correct, transparent reporting of methodologies even if they are not methodologically robust. If the journal endorses CONSORT, the author should submit a CONSORT checklist, this needs to be checked thoroughly during the reviewing process. It is reported that these checklists do not adequately reflect the information needed in published papers, including details pertaining to randomisation and allocation concealment, this needs improving (Blanco et al., 2018). CONSORT checklists may be aimed at an audience with a level of methodological knowledge and experience that not all users have.

Peer reviewers are experts within a field and are asked to critically review and appraise research; typically they are not methodologists. This lack of specialist knowledge may impact the methodological quality of the publication, due to their understanding of how to appraise the extent CONSORT has been adhered to in a

manuscript. This demonstrates the importance of involving a methodologist in the reviewing process. All parties have a role to play in ensuring robust reporting. It is my belief that authors should take responsibility for reporting their research robustly and should not rely on the publishing process alone to drive the methodological robustness of their publication.

Reporting guidelines may provide insufficient advice on how to report the use of envelopes as an allocation concealment method. When an author states 'SNOSE only', they may believe they are adhering to CONSORT. The exploration and explanatory documents published alongside reporting guidelines provide detailed methodological information and examples of how allocation concealment should be reported (Altman et al., 2001b, Moher et al., 2012). However, these are lengthy documents and may not be read in full by all those involved in writing and reviewing publications.

In 2018, Schultz et al reflected on the term 'allocation concealment' which they coined in 1994 and recognised that perhaps many authors and editors define the term incorrectly (Schulz et al., 2018). If this lack of understanding is occurring in practice it could explain the inadequate implementation and reporting by these groups.

Work is needed to ensure the allocation concealment trifecta is fulfilled, this would facilitate high quality research being designed, implemented and reported that can contribute to the evidence base with a low ROB.

6.2: Original contribution to knowledge provided by the thesis

There is a significant body of work describing the problems of poor allocation concealment in randomised controlled trials; my research explores ways to address this. This thesis has formed a coherent body of work; with each paper building on the previous. I met my objectives of identifying ways to detect selection bias in systematic reviews, and prevent selection bias by improving the implementation and reporting of the allocation concealment methods of envelope and blocking. My work makes an original contribution to the knowledge base in the following ways:

- I developed a novel technique to identify systematic reviews of RCTs that potentially suffer from selection bias by including biased component trials and are therefore at risk of erroneous conclusions. I demonstrate baseline age heterogeneity, which can only be explained by the inclusion of RCTs with baseline selection bias (paper 1).
- In paper 2 I extended this technique by demonstrating a more powerful/trial specific covariates in a baseline meta-analysis are associated with greater heterogeneity.
- I showed that this baseline heterogeneity is not a function of the reviews including 'old' trials because poor allocation concealment, which is the most likely explanation of selection bias, is still prevalent in recently published trials (paper 3).
- I identified that allocation concealment is both poorly reported and implemented; envelope use was observed to be suboptimal and insecure blocking schedules were selected (paper 3). I therefore conducted research to aid the robust application of these methods (papers 4 and 5).
- I show that envelopes are widely employed and will continue to be so for pragmatic reasons. In response I created an original reference guide for researchers to improve envelope implementation and reporting (paper 4)
- I have published the first report based on empirical evidence (rather than modelling studies) that demonstrates RCTs using small variable blocking schemes that include a two are at risk of subversion (paper 5).

6.3: Disseminating the methodological research presented in this thesis

Disseminating methodological research benefits all stakeholders involved in RCT design and conduct, providing the foundation for high-quality research outputs leading to treatment progression for patients. The key stakeholders that could benefit from my research are:

- Systematic reviewers: the techniques demonstrated in papers 1 and 2 may lead to more accurate conclusions when undertaking systematic reviews.
- Policy and decision makers: if more robust systematic reviews are undertaken by applying techniques presented in 1 and 2, and implementation and reporting of allocation concealment is more robust (papers 4 and 5) policy and decisions will be able to make firmer recommendations benefitting patient care.
- Healthcare and clinical staff: research in healthcare is often undertaken by clinical staff, not methodologists. My research can aid their understanding of the power of an RCT to influence practice and the importance of robust allocation concealment implementation and subsequent reporting when critiquing research. Robust conduct of research may also increase professional and organisation reputation.
- Medical, nursing and allied health professional (AHP) students: increased knowledge about robust implementation and reporting of allocation concealment may safeguard future research from selection bias and suboptimal reporting.
- Recruiting personnel: my research may prevent the subversion of trials as recruiters will have increased awareness of the importance of robust allocation concealment implementation.
- Research funders and reviewers: paper 3 demonstrates suboptimal allocation concealment methods are funded, papers 4 and 5 provide information to aid decisions on funding applications proposing to use envelopes or blocking.
- Journal editors and peer reviewers: paper 3 demonstrates that reporting of allocation concealment methods needs to be improved. Papers 4 and 5 detail what methodological information should be reported for envelopes and blocking respectively.
- Patients: if allocation concealment implementation and reporting is improved this would reduce selection bias, improving the accuracy of research and systematic reviews which could positively impact patient care.

6.3.1: Effective dissemination

This thesis is submitted for a PhD by Publication. Each publication has been through an editorial and peer review process which has influenced the contents as well as where the paper was published. It took several weeks to find suitable peer reviewers for papers 1 and 4 as there was such a small pool with relevant backgrounds and research interests. If methodological papers are mainly reported within specialist journals it will take longer for trial methodology and reporting to reach a wide audience and effect change, impacting the time that it will take for patients to benefit from allocation concealment trifecta fulfilment.

Effective dissemination strategies for methodological research are debated amongst methodologists, with a focus on more novel, audience-specific avenues rather than via the traditional journal route. Suggestions include training and informative webinars targeting specific audiences, short videos/animations for distribution via social media, using outlets such as Slack, Discord and scientific blogs (Better Methods, September 10th 2021). The introduction of open reporting platforms, such as F1000 Research (F1000, 2022) may change the reporting landscape and lead to more open and transparent reporting. Research on Research (RoR) (Research-on-Research, 2022) is a new initiative with the tag line 'Building and connecting global communities to enhance research through research' and may be an avenue to disseminate methodological research.

6.4: Future research

My research highlights the need for practical ways to ensure the allocation concealment trifecta is achieved, selection bias is minimised as far as possible and detected when it is not possible to prevent. This can only occur by identifying and working with all stakeholders involved in research design, conduct and reporting.

It is important that in addition to focussing on the immediate improvements that can be made, we look to the future and try to change practices to improve research design, implementation and reporting. Ultimately, we need to stop insecure allocation concealment being implemented and advocate for secure methods to be robustly implemented and reported.

During the progression of my research, I identified a number of areas that would benefit from further investigation. The areas, the methodologies I feel would be appropriate, and the stakeholders for which the research would be relevant are presented in table 2.

Table 2: Future research questions identified from research undertaken within this thesis

Dissemination methods and engagement		
Research question	**Proposed methods**	**Relevant stakeholders**
How can the findings of methodological research be disseminated more effectively to ensure take up in practice?	As has been identified, methodological research could benefit from more novel dissemination approaches, this now needs trialling. There are many different groups that would be positively impacted by methodological research, but not all would engage with, or read, traditional methodological journal publications. For example, I have had positive responses and email correspondence from systematic reviewers who had read my research in papers 1 and 2, but recruiters or health care staff may be less likely to read a journal publication, such as paper 4, around envelope use. An initial method would be to survey different stakeholders to determine how to best disseminate this information. Engagement with different platforms (such as Research on Research(Research-on-Research, 2022)) and different working groups to help identify mechanisms to disseminate research would also be useful. The aim of disseminating and engaging with groups about methodological research would be to improve the allocation concealment trifecta and ultimately lead to advances in patient care.	• Funders • Peer reviewers of publications • Researchers/trial personnel • Personnel involved in recruiting participants • Policy/decision makers • Reviewers for funding applications • Regulatory/advisory bodies • Medical, nursing and AHP clinical staff and students • Systematic reviewers

Exploring the use of baseline heterogeneity as reported in papers 1 and 2		
Research question	**Proposed methods**	**Relevant stakeholders**
How can the techniques piloted in papers 1 and 2 be explored further to promote their use within systematic reviews	Papers 1 and 2 were exemplars of the use of baseline heterogeneity, but given the amount of missing data on allocation concealment in the included studies, replication of this work is needed. Collaborating with systematic reviewers and organisations responsible for the conduct of systematic reviews (such as the Cochrane (Cochrane, 2022) and the Campbell Collaborations (Collaboration, 2022)) would enable targeted work to promote, pragmatically explore and refine further the use of the technique explored in papers 1 and 2. Protocols could be developed which include this complementary technique. Systematic reviews could be assessed using this technique within a range of different research areas. Different prognostic variables could be identified across a range of trial types to further inform systematic reviewers regarding which variables should be tested at baseline. The technique presented in papers 1 and 2 could also be applied to systematic reviews that have been previously reported. Systematic reviews that are found to be impacted by biased RCTs would then be identified and conclusions investigated to ensure biased and potentially harmful conclusions are not impacting healthcare decisions and treatments.	• Systematic reviewers • Policy/decision makers

Education and understanding to ensure that allocation concealment is implemented and reported robustly		
Research question	**Proposed methods**	**Relevant stakeholders**
What methods can be explored to increase education and understanding to increase the robust implementation and reporting of allocation concealment	I found allocation concealment was implemented and reported poorly within RCTs that were included in all my research. As stated in section 6.1, many individuals involved in healthcare research are not methodologists. It is essential that ways to engage and work with all stakeholders involved in research to improve their knowledge around the robust implementation and reporting and allocation concealment. Suggested methods to explore are targeted lectures, workshops, mentors and practical work experience tailored specifically for different stakeholder groups (Sharp and Curlewis, 2019). It has been suggested that one group that could be targeted are students in healthcare and related professions/fields as they may undertake research within their careers (Glasziou and Chalmers, 2018). If this group can be taught about allocation concealment this could stop perpetuating poor methodological knowledge and thus poor implementation and reporting. Specifically identifying ways that methodologists can engage with this group would be beneficial to avoid future research waste. Engaging with funders and funding panel members may decrease suboptimal allocation concealment methods from being funded.	• Medical, nursing and AHP clinical staff and students • Policy/decision makers • Regulatory/advisory bodies • Researchers/trial personnel • Personnel involved in recruiting participants • Funders

Motives to subvert trials		
Research question	**Proposed methods**	**Relevant stakeholders**
Why do individuals subvert trials? Can further understanding of motives lead to the development of interventions to reduce subversion?	My research observed poor implementation of allocation concealment. These observations imply that there was deliberate tampering of the allocation sequence (papers 1, 2 and 5 did demonstrate evidence of baseline heterogeneity of age). Irrespective of the motive, this is evidence that the research was subverted. Similarly, papers 3 and 4 demonstrated poor implementation of allocation concealment where subversion could have occurred. It is known that frequently intervention/experimental groups are associated with more hope and positive expectations so subversion is more likely to occur in this group (Paludan-Müller et al., 2016). There are known situations where the patient recruiters will prefer the control group, such as when the perceived benefit is negligible. This has been observed in an RCT testing homeopathy versus standard care (Hróbjartsson et al., 2012). The patient recruiter may also want the best possible care for the patient which conflicts with the desire to conduct a methodologically robust RCT. The most dominant motive will influence whether they try to subvert allocations and into which group (Paludan-Müller et al., 2016). Determining why individuals subvert trials may be helpful in determining methods to prevent subversion. This could lead to identifying training about the importance of allocation concealment and the promotion of the importance of adhering to the	• Personnel involved in recruiting participants • Medical, nursing and AHP clinical staff and students

	allocation schedule to prevent subversion. The aim of this research would be to improve the allocation concealment trifecta by improving allocation concealment implementation	
Writing and publishing robust RCT publications		
Research question	**Proposed methods**	**Relevant stakeholders**
How can reporting of allocation concealment methods be improved in publications of RCTs?	All papers included in this thesis observed poor reporting. Poor reporting of RCTs is not new and was the driver for the CONSORT statement in 1996. However, recently reported RCTs are still lacking essential methodological details. Exploring what the barriers and challenges are to robust reporting of allocation concealment would enable targeted guidance created for the needs of different audiences to support writing robust and transparent RCT publications. Work could be undertaken with editors to assess the reviewing process to identify ways to improve reporting quality that is operational within the infrastructure journals can provide, as well as a process that would be acceptable to authors. Working with funders and authors would enable targeted recommendations to be made to improve the implementation and reporting of allocation concealment. However, work so far in this area has still not improved reporting adequately. An avenue to explore would be to assess the effectiveness and acceptability of using a more prescriptive template to support less experienced authors to report methodological information.	• Researchers/trial personnel involved in publishing RCTs • Funders • Journal Editors • Medical, nursing and AHP clinical staff and students involved in publishing RCTs

	Currently the journal 'Trials' is piloting the use of a SPIRIT protocol submission template (Trials, 2022). The completeness of reporting and acceptability with authors and journal personnel could be explored. This may lead to the development of a simple template to aid the reporting of key methodological factors associated with RCTs. This template could act as a reporting prompt by breaking down the detail that is necessary to be reported in relation to allocation concealment.	
Research Waste		
Research question	**Proposed methods**	**Relevant stakeholders**
Does promoting the research waste associated with suboptimal implementation and/or reporting of allocation concealment result in fulfilment of the allocation concealment trifecta?	The implementation and reporting of allocation concealment observed throughout this thesis was suboptimal with many component publications of papers 1-5 contributing to research waste. This is not only a waste of resources; it could potentially influence progress to patient care. Work needs to be undertaken in promoting the research waste (both the financial and other resources) of trials that use poor allocation concealment methods. This would enable the consequence of choosing an inexpensive but insecure allocation concealment method against a more expensive but robust method of allocation concealment to be understood. This may motivate researchers and funders to budget and employ more expensive third-party services if appropriate to their research design as these will result in more robust research.	• Researchers/trial personnel • Personnel involved in recruiting participants • Systematic reviewers • Reviewers for funding applications • Policy/Decision makers • Funders • Reviewers of publications

	Work with funders is needed to identify mechanisms to assess the robustness of allocation concealment methods which could then result in more robust methods being funded. Additionally, some groups involved in research (such as medical staff/students) may not be familiar with the term research waste. Providing education and training on research waste highlights and puts into context the importance of robust allocation concealment implementation and reporting. Illustrating the importance of adhering to the randomisation schedule to those involved in recruiting participants may avoid future research waste by preventing subversion.	• Regulatory/advisory bodies • Medical, nursing and AHP clinical staff and students
Envelope Use		
Research question	**Proposed methods**	**Relevant stakeholders**
How can the implementation and reporting of the allocation concealment method of envelopes be improved?	Paper 4 demonstrated that envelopes are going to continue to be implemented but, crucially, their implementation and reporting needs to be improved to fulfil the allocation concealment trifecta. Paper 4 presented a reference guide detailing how to do this but did not explore how to disseminate this knowledge, or how best to engage and promote this with a range of stakeholders that are impacted by envelope use. Exploring the media avenue and creating an instructional video may be a useful tool in promoting robust secure envelopes. A video of best practice in	• Researchers/trial personnel • Personnel involved in recruiting participants • Reviewers for funding applications • Funders

	randomisation which includes envelopes, may be accessed more widely as a video on envelopes only could be too niche and result in less engagement. Promoting the appropriateness of envelope use is needed to ensure that envelopes are not used if a more robust alterative is possible.	• Reviewers of publications
Restricted randomisation		
Research question	**Proposed methods**	**Relevant stakeholders**
How can newer, more novel and robust restricted randomisation methods be promoted to increase engagement by stakeholders?	Paper 3 demonstrated that blocking was not implemented well and paper 5 demonstrated that blocking schemes including a block size of two are at risk of subversion. Blocking is frequently used to restrict randomisation, therefore recommendations to improve this could lead in the short-term to more robustly implemented research. However, research needs to be improved in the longer term, therefore I concluded my research in paper 5 by making suggestions on other methods that could be used to restrict randomisation that may be more robust (as they are less likely to be predicted). The next step would be to investigate these different methods to restrict randomisation and the acceptability of these to different user groups. This should include examining the cost effectiveness of the more robust method of minimisation compared to the less costly but more insecure method of blocking to restrict randomisation. Promoting specific details about the robust implementation and reporting of the different methods to ensure the trifecta is complete would be beneficial.	• Researchers/trial personnel • Personnel involved in recruiting participants • Funders • Advisory bodies

Section 7: Conclusion to the thesis

Research is not performed within a vacuum. It is shaped by what has come before and what will come after. Participants give their time and share confidential data when taking part in research in the belief that they are contributing to knowledge in a given area. The research community has an ethical and moral responsibility to participants and all stakeholders to ensure that research is undertaken and reported with methodological rigour. This thesis has been written during the COVID-19 pandemic, the consequences of which are likely to be scarcer resources for undertaking research for many years. It is more important than ever that research should be designed, undertaken and reported with minimal waste and maximum value to the knowledge base (Glasziou and Chalmers, 2017). Through my exploration of methods to detect selection bias, and the implementation and reporting of allocation concealment, this thesis has demonstrated that we have a long way to go to fulfil the allocation concealment trifecta.

Section 8: Abbreviations

AHP	Allied health professional
BMI	Body Mass Index
BMJ	British Medical Journal
CONSORT	Consolidated Standards of Reporting Trials
EQUATOR	Enhancing the quality and transparency of health research
JAMA	Journal of the American Medical Association
NEJM	New England Journal of Medicine
RCT(s)	Randomised controlled trial(s)
ROB	Risk of Bias
RoB2	Risk of Bias 2
RoR	Research on research
SNOSE	Sequentially numbered opaque sealed envelopes

Section 9: Glossary

Allocation concealment: A technique used to prevent selection bias by concealing the allocation sequence from those assigning participants to intervention groups, until the moment of assignment. Allocation concealment prevents researchers from (unconsciously or otherwise) influencing which participants are assigned to a given intervention group.

Allocation Concealment Trifecta: the three elements that constitute robust allocation concealment; designing an RCT and selecting a robust allocation concealment method, robust implementation and reporting.

Methodologist: researcher who has expertise in the area of robust research design, conduct, analysis and publication

Restricted Randomisation: Any procedure used with random assignment to achieve balance between study groups in size or baseline characteristics.

Selection Bias: A systematic error in creating intervention groups, such that they differ with respect to prognosis. That is, the groups differ in measured or unmeasured baseline characteristics because of the way participants were selected or assigned. Also used to mean that the participants are not representative of the population of all possible participants.

Simple Randomisation: Randomisation without restriction. In a two-group trial, it is analogous to the toss of a coin.

Section 10: References

ALTMAN, D., MOHER, D. & SCHULTZ, K. 2001a. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *JAMA,* 285**,** 1987-91.

ALTMAN, D. G. 1994. The scandal of poor medical research. 308 (6924)**,** pp.283-284.

ALTMAN, D. G. 2005. Endorsement of the CONSORT statement by high impact medical journals: survey of instructions for authors. *Bmj,* 330**,** 1056-1057.

ALTMAN, D. G. & BLAND, J. M. 1999. How to randomise. *Bmj,* 319**,** 703-704.

ALTMAN, D. G. & DORE, C. 1990. Randomisation and baseline comparisons in clinical trials. *The Lancet,* 335**,** 149-153.

ALTMAN, D. G., SCHULZ, K. F., MOHER, D., EGGER, M., DAVIDOFF, F., ELBOURNE, D., GØTZSCHE, P. C. & LANG, T. 2001b. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of internal medicine,* 134**,** 663.

ALTMAN, D. G., SIMERA, I., HOEY, J., MOHER, D. & SCHULZ, K. 2008. EQUATOR: reporting guidelines for health research. *Open Medicine,* 2**,** e49.

BARCOT, O., BORIC, M., PERICIC, T. P., CAVAR, M., DOSENOVIC, S., VUKA, I. & PULJAK, L. 2018. Judgments of risk of bias associated with random sequence generation in trials included in Cochrane systematic reviews are frequently erroneous. *BioRxiv***,** 366674.

BEGG, C., CHO, M., EASTWOOD, S., HORTON, R., MOHER, D., OLKIN, I., PITKIN, R., RENNIE, D., SCHULZ, K. F. & SIMEL, D. 1996. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *Jama,* 276**,** 637-639.

BERGER, V. 2005. Quantifying the Magnitude of Baseline Covariate Imbalances Resulting from Selection Bias in Randomized Clinical Trials. *Biometrical Journal* 47**,** 119-127.

BERGER, V. 2007. *Selection bias and covariate imbalances in randomized clinical trials*, John Wiley & Sons.

BERGER, V. W. 2009. Do not test for baseline imbalances unless they are known to be present? *Quality of Life Research,* 18**,** 399-399.

BETTER METHODS, B. R. September 10th 2021. Better Methods, Better Research Webinar, hosted by the NIHR Incubator for Methodology.

BJELAKOVIC, G., NIKOLOVA, D. & GLUUD, C. 2014. Antioxidant supplements and mortality. *Current Opinion in Clinical Nutrition & Metabolic Care,* 17**,** 40-44.

BLANCO, D., BIGGANE, A. M. & COBO, E. 2018. Are CONSORT checklists submitted by authors adequately reflecting what information is actually reported in published papers? *Trials,* 19**,** 1-4.

BLAND, M. 2013. Do baseline p-values follow a uniform distribution in randomised trials? *PloS one,* 8**,** e76010.

BODEN, I., SKINNER, E. H., BROWNING, L., REEVE, J., ANDERSON, L., HILL, C., ROBERTSON, I. K., STORY, D. & DENEHY, L. 2018. Preoperative physiotherapy for the prevention of respiratory complications after upper abdominal surgery: pragmatic, double blinded, multicentre randomised controlled trial. *bmj,* 360.

BROWN, S., THORPE, H., HAWKINS, K. & BROWN, J. 2005. Minimization—reducing predictability for multi-centre trials whilst retaining balance within centre. *Statistics in medicine,* 24**,** 3715-3727.

CHALMERS, T. C., CELANO, P., SACKS, H. S. & SMITH JR, H. 1983. Bias in treatment assignment in controlled clinical trials. *New England Journal of Medicine,* 309**,** 1358-1361.

CHAN, A.-W., TETZLAFF, J. M., ALTMAN, D. G., LAUPACIS, A., GØTZSCHE, P. C., KRLEŽA-JERIĆ, K., HRÓBJARTSSON, A., MANN, H., DICKERSIN, K. & BERLIN, J. A. 2013. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Annals of internal medicine,* 158**,** 200-207.
CHMP, C. F. M. P. F. H. U. 2015. Guideline on adjustment for baseline covariates in clinical trials. Technical Report. EMA/CHMP/295050/2013. London: European Medicines Agency
CLARK, L., FAIRHURST, C., HEWITT, C., BIRKS, Y., BRABYN, S., COCKAYNE, S., RODGERS, S., HICKS, K., HODGSON, R., LITTLEWOOD, E. & TORGERSON, D. 2013a. Assessing the presence of selection bias in meta-analyses of randomised trials using baseline heterogeneity. *Trials,* 14**,** O96.
CLARK, L., SCHMIDT, U., THARMANATHAN, P., ADAMSON, J., HEWITT, C. & TORGERSON, D. 2013b. Allocation concealment: a methodological review. *Journal of evaluation in clinical practice,* 19**,** 708-712.
CLARK, L., SCHMIDT, U., THARMANATHAN, P., ADAMSON, J., HEWITT, C. & TORGERSON, D. 2013c. Poor reporting quality of key Randomization and Allocation Concealment details is still prevalent among published RCTs in 2011: a review. *Journal of evaluation in clinical practice,* 19**,** 703-707.
CLARKE, M., HOPEWELL, S. & CHALMERS, I. 2010. Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting. *The Lancet,* 376**,** 20-21.
COCHRANE. 2022. *Cochrane* [Online]. Available: https://www.cochrane.org/ [Accessed 02.06 2022].
COCHRANE CONSUMER NETWORK. 2020. *Cochrane consumer network, levels of evidence* [Online]. Available: https://consumers.cochrane.org/levels-evidence [Accessed 30.08 2020].
COLLABORATION, C. 2022. *Campbell Collaboration* [Online]. Available: https://www.campbellcollaboration.org/ [Accessed 02.06 2022].
CONSORT. 2018. *In Memoriam, Doug Altman 12 July 1948- 3 June 2018* [Online]. Available: http://www.consort-statement.org/news/doug [Accessed 05.07 2022].
CONSORT. 2022a. *CONSORT checklist: Randomisation: allocation concealment mechanism* [Online]. Available: http://www.consort-statement.org/checklists/view/32--consort-2010/89-randomisation-allocation-concealment-mechanism [Accessed 09.07 2022].
CONSORT. 2022b. *CONSORT Glossary* [Online]. Available: http://www.consort-statement.org/resources/glossary#A [Accessed 09.07 2022].
CONSORT. 2022c. *CONSORT Glossary* [Online]. Available: http://www.consort-statement.org/resources/glossary#S [Accessed 09.07 2022].
CORBETT, M. S., HIGGINS, J. P. & WOOLACOTT, N. F. 2014. Assessing baseline imbalance in randomised trials: implications for the Cochrane risk of bias tool. *Research Synthesis Methods,* 5**,** 79-85.
COX, T. M., DRELICHMAN, G., CRAVO, R., BALWANI, M., BURROW, T. A., MARTINS, A. M., LUKINA, E., ROSENBLOOM, B., ROSS, L. & ANGELL, J. 2015. Eliglustat compared with imiglucerase in patients with Gaucher's disease type 1 stabilised on enzyme replacement therapy: a phase 3, randomised, open-label, non-inferiority trial. *The Lancet,* 385**,** 2355-2362.
DEL GIGLIO, A. & COSTA, L. J. 2004. The quality of randomised controlled trials may be better than assumed. *Bmj,* 328**,** 24-25.
DOIG, G. S. & SIMPSON, F. 2005. Randomization and allocation concealment: a practical guide for researchers. *Journal of critical care,* 20**,** 187-191.
EFIRD, J. 2011. Blocked randomization with randomly selected block sizes. *International journal of environmental research and public health,* 8**,** 15-20.

ENVELOPE, S. 2022. *Randomisation and online databases for clinical trials* [Online]. Available: https://www.sealedenvelope.com/ [Accessed 09.07 2022].
EQUATOR. 2021. *Equator network enhancing the quality and transparency of health research* [Online]. Available: https://www.equator-network.org/ [Accessed 14.11 2021].
F1000. 2022. *F1000 Research Open for Science* [Online]. Available: https://f1000research.com/ [Accessed 16.01 2022].
FAIRHURST, C., HEWITT, C. E. & TORGERSON, D. J. 2020. Using pairwise randomisation to reduce the risk of bias. *Research Methods in Medicine & Health Sciences,* 1**,** 2-6.
FLETCHER, J. 2007. What is heterogeneity and is it important? *Bmj,* 334**,** 94-96.
FORDER, P. M., GEBSKI, V. J. & KEECH, A. C. 2005. Allocation concealment and blinding: when ignorance is bliss. *Medical Journal of Australia,* 182**,** 87.
FURLER, J., MAGIN, P., PIROTTA, M. & VAN DRIEL, M. 2012. Participant demographics reported in" Table 1" of randomised controlled trials: a case of" inverse evidence"? *International Journal for Equity in Health,* 11**,** 1-4.
GLASZIOU, P. & CHALMERS, I. 2017. *Paul Glasziou and Iain Chalmers: Funders and regulators are more important than journals in fixing the waste in research* [Online]. Available: https://blogs.bmj.com/bmj/2017/09/06/paul-glasziou-and-iain-chalmers-funders-and-regulators-are-more-important-than-journals-in-fixing-the-waste-in-research/ [Accessed 30.08 2021].
GLASZIOU, P. & CHALMERS, I. 2018. Research waste is still a scandal—an essay by Paul Glasziou and Iain Chalmers. *Bmj,* 363**,** k4645.
GLASZIOU, P., DJULBEGOVIC, B. & BURLS, A. 2006. Are systematic reviews more cost-effective than randomised trials? *The Lancet,* 367**,** 2057-2058.
GOPALAKRISHNAN, S. & GANESHKUMAR, P. 2013. Systematic reviews and meta-analysis: understanding the best evidence in primary healthcare. *Journal of family medicine and primary care,* 2**,** 9.
GRATT, G., SCHOUTERN, E. & KOK, F. 2002. Effect of daily vitamin E and multivitamin supplementation on acute respiratory tract infection in elderly persons: a randomized controlled trial. *JAMA,* 288**,** 715-21.
HARITON, E. & LOCASCIO, J. J. 2018. Randomised controlled trials—The gold standard for effectiveness research. *BJOG: an international journal of obstetrics and gynaecology,* 125**,** 1716.
HARTLING, L., OSPINA, M., LIANG, Y., DRYDEN, D. M., HOOTON, N., SEIDA, J. K. & KLASSEN, T. P. 2009. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. *Bmj,* 339**,** b4012.
HEWITT, C., HAHN, S., TORGERSON, D. J., WATSON, J. & BLAND, J. M. 2005. Adequacy and reporting of allocation concealment: review of recent trials published in four general medical journals. *Bmj,* 330**,** 1057-1058.
HEWITT, C. E. & TORGERSON, D. J. 2006. Is restricted randomisation necessary? *Bmj,* 332**,** 1506-1508.
HEWITT, C. E., TORGERSON, D. J. & BERGER, V. W. 2009. Potential for technical errors and subverted allocation can be reduced if certain guidelines are followed: examples from a web-based survey. *Journal of clinical epidemiology,* 62**,** 261-269.
HICKS, A., FAIRHURST, C. & TORGERSON, D. J. 2018. A simple technique investigating baseline heterogeneity helped to eliminate potential bias in meta-analyses. *Journal of clinical epidemiology,* 95**,** 55-62.
HIGGINS, J. P., ALTMAN, D. G., GØTZSCHE, P. C., JÜNI, P., MOHER, D., OXMAN, A. D., SAVOVIĆ, 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.

HIGGINS, J. P., SAVOVIĆ, J., PAGE, M. J., ELBERS, R. G. & STERNE, J. A. 2019. Assessing risk of bias in a randomized trial. *Cochrane Handbook for Systematic Reviews of Interventions*, 205-228.
HIGGINS, J. P. & THOMPSON, S. G. 2002. Quantifying heterogeneity in a meta-analysis. *Statistics in medicine,* 21, 1539-1558.
HIGGINS, J. P., THOMPSON, S. G., DEEKS, J. J. & ALTMAN, D. G. 2003. Measuring inconsistency in meta-analyses. *Bmj,* 327, 557-560.
HOPEWELL, S., ALTMAN, D. G., MOHER, D. & SCHULZ, K. F. 2008. Endorsement of the CONSORT Statement by high impact factor medical journals: a survey of journal editors and journal'Instructions to Authors'. *Trials,* 9, 20.
HOPEWELL, S., DUTTON, S., YU, L.-M., CHAN, A.-W. & ALTMAN, D. G. 2010. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. *Bmj,* 340, c723.
HRÓBJARTSSON, A., BOUTRON, I., TURNER, L., ALTMAN, D. G. & MOHER, D. 2013. Assessing risk of bias in randomised clinical trials included in Cochrane Reviews: the why is easy, the how is a challenge. *Cochrane Database of Systematic Reviews*.
HRÓBJARTSSON, A., THOMSEN, A. S. S., EMANUELSSON, F., TENDAL, B., HILDEN, J., BOUTRON, I., RAVAUD, P. & BRORSON, S. 2012. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *Bmj,* 344.
HULSHOF, T. A., ZUIDEMA, S. U., VAN MEER, P. J., GISPEN-DE WIED, C. C. & LUIJENDIJK, H. J. 2019. Baseline imbalances and clinical outcomes of atypical antipsychotics in dementia: A meta-epidemiological study of randomized trials. *International journal of methods in psychiatric research,* 28, e1757.
IOANNIDIS, J. P., GREENLAND, S., HLATKY, M. A., KHOURY, M. J., MACLEOD, M. R., MOHER, D., SCHULZ, K. F. & TIBSHIRANI, R. 2014. Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet,* 383, 166-175.
JACKSON, J. L. & KURIYAMA, A. 2018. From the editors' desk: bias in systematic reviews—let the reader beware. Springer.
JORDAN, V. M., LENSEN, S. F. & FARQUHAR, C. M. 2017. There were large discrepancies in risk of bias tool judgments when a randomized controlled trial appeared in more than one systematic review. *Journal of Clinical Epidemiology,* 81, 72-76.
JÜNI, P., ALTMAN, D. G. & EGGER, M. 2001. Assessing the quality of controlled clinical trials. *Bmj,* 323, 42-46.
KAHAN, B. C. & MORRIS, T. P. 2012. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. *Bmj,* 345.
KAHAN, B. C., REHAL, S. & CRO, S. 2015. Risk of selection bias in randomised trials. *Trials,* 16, 1-7.
KENNEDY, A. D., TORGERSON, D. J., CAMPBELL, M. K. & GRANT, A. M. 2017. Subversion of allocation concealment in a randomised controlled trial: a historical case study. *Trials,* 18, 1-6.
KHAW, K.-T., BINGHAM, S., WELCH, A., LUBEN, R., WAREHAM, N., OAKES, S. & DAY, N. 2001. Relation between plasma ascorbic acid and mortality in men and women in EPIC-Norfolk prospective study: a prospective population study. *The lancet,* 357, 657-663.
KJAERGARD, L., VILLUMSEN, J., GLUUD, C. & COLLOQUIUM, C. Quality of randomised clinical trials affects estimates of intervention efficacy. 1999. 57.
KUNZ, R. & OXMAN, A. D. 1998. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *Bmj,* 317, 1185-1190.

MARSHALL, G., BLACKLOCK, J., CAMERON, C., CAPON, N., CRUICKSHANK, R., GADDUM, J., HEAF, F., HILL, A. B., HOUGHTON, L. & HOYLE, J. C. 1948. Streptomycin treatment of pulmonary tuberculosis: a medical research council investigation. *Br Med J,* 2**,** 769-82.
MARTYN, C. 1996. Not quite as random as I pretended. *The Lancet,* 8994**,** 70.
MATTS, J. P. & LACHIN, J. M. 1988. Properties of permuted-block randomization in clinical trials. *Controlled clinical trials,* 9**,** 327-344.
MICKENAUTSCH, S., FU, B., GUDEHITHLU, S. & BERGER, V. W. 2014. Accuracy of the Berger-Exner test for detecting third-order selection bias in randomised controlled trials: a simulation-based investigation. *BMC medical research methodology,* 14**,** 1-11.
MILLER III, E. R., PASTOR-BARRIUSO, R., DALAL, D., RIEMERSMA, R. A., APPEL, L. J. & GUALLAR, E. 2005. Meta-analysis: high-dosage vitamin E supplementation may increase all-cause mortality. *Annals of internal medicine,* 142**,** 37-46.
MITCHELL, A., MOE-BYRNE, T., CUNNINGHAM-BURLEY, R., DEAN, A., RANGAN, A., ROCHE, J. & TORGERSON, D. J. 2019. Poor allocation concealment methods are associated with heterogeneity in age and statistical significance of the primary outcome: Review of recent trials published in four general medical journals. *Journal of Evaluation in Clinical Practice*.
MOHER, D., HOPEWELL, S., SCHULZ, K. F., MONTORI, V., GØTZSCHE, P. C., DEVEREAUX, P., ELBOURNE, D., EGGER, M. & ALTMAN, D. G. 2012. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *International journal of surgery,* 10**,** 28-55.
MOHER, D., JONES, A., LEPAGE, L. & GROUP, C. 2001. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *Jama,* 285**,** 1992-1995.
MOHER, D., PHAM, B., JONES, A., COOK, D. J., JADAD, A. R., MOHER, M., TUGWELL, P. & KLASSEN, T. P. 1998. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *The Lancet,* 352**,** 609-613.
MRC/BHF, H. P. S. C. G. 2002. MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20 536 high-risk individuals: a randomised placebo-controlled trial. *The Lancet,* 360**,** 23-33.
MURAD, M. H., ASI, N., ALSAWAS, M. & ALAHDAB, F. 2016. New evidence pyramid. *BMJ Evidence-Based Medicine,* 21**,** 125-127.
NADKARNI, A., WEOBONG, B., WEISS, H. A., MCCAMBRIDGE, J., BHAT, B., KATTI, B., MURTHY, P., KING, M., MCDAID, D. & PARK, A.-L. 2017. Counselling for Alcohol Problems (CAP), a lay counsellor-delivered brief psychological treatment for harmful drinking in men, in primary care in India: a randomised controlled trial. *The Lancet,* 389**,** 186-195.
OLIVO, S. A., MACEDO, L. G., GADOTTI, I. C., FUENTES, J., STANTON, T. & MAGEE, D. J. 2008. Scales to assess the quality of randomized controlled trials: a systematic review. *Physical therapy,* 88**,** 156-175.
PAGE, M. J., SHAMSEER, L., ALTMAN, D. G., TETZLAFF, J., SAMPSON, M., TRICCO, A. C., CATALA-LOPEZ, F., LI, L., REID, E. K. & SARKIS-ONOFRE, R. 2016. Epidemiology and reporting characteristics of systematic reviews of biomedical research: a cross-sectional study. *PLoS medicine,* 13**,** e1002028.
PALUDAN-MÜLLER, A., LAURSEN, D. R. T. & HRÓBJARTSSON, A. 2016. Mechanisms and direction of allocation bias in randomised clinical trials. *BMC medical research methodology,* 16**,** 133.
PATEL, V., WEOBONG, B., WEISS, H. A., ANAND, A., BHAT, B., KATTI, B., DIMIDJIAN, S., ARAYA, R., HOLLON, S. D. & KING, M. 2017. The Healthy Activity Program

(HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: a randomised controlled trial. *The Lancet,* 389**,** 176-185.
PETO, R. 1999. Failure of randomisation by “sealed” envelope. *The Lancet,* 354**,** 73.
PIPER, R. D., COOK, D. J., BONE, R. C. & SIBBALD, W. J. 1996. Introducing critical appraisal to studies of animal models investigating novel therapies in sepsis. *Critical care medicine,* 24**,** 2059-2070.
PROPADALO, I., TRANFIC, M., VUKA, I., BARCOT, O., PERICIC, T. P. & PULJAK, L. 2019. In Cochrane reviews, risk of bias assessments for allocation concealment were frequently not in line with Cochrane's Handbook guidance. *Journal of clinical epidemiology,* 106**,** 10-17.
RESEARCH-ON-RESEARCH. 2022. *Research on Research* [Online]. Available: https://ror-hub.org/ [Accessed 08.07 2022].
RESPONSES, R. 2016. *Rapid responses to the article 'Allocation concealment in randomised controlled trials: are we getting better?'* [Online]. Available: https://www.bmj.com/content/355/bmj.i5663/rapid-responses [Accessed 01.08 2020].
ROBERTS, I. & KER, K. 2016. Cochrane: the unfinished symphony of research synthesis. *Systematic reviews,* 5**,** 1-5.
SALMINEN, P., PAAJANEN, H., RAUTIO, T., NORDSTRÖM, P., AARNIO, M., RANTANEN, T., TUOMINEN, R., HURME, S., VIRTANEN, J. & MECKLIN, J.-P. 2015. Antibiotic therapy vs appendectomy for treatment of uncomplicated acute appendicitis: the APPAC randomized clinical trial. *Jama,* 313**,** 2340-2348.
SCHULZ, K. F. 1995. Subverting randomization in controlled trials. *JAMA: the journal of the American Medical Association,* 274**,** 1456.
SCHULZ, K. F., ALTMAN, D. G. & MOHER, D. 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC medicine,* 8**,** 18.
SCHULZ, K. F., CHALMERS, I., ALTMAN, D. G., GRIMES, D. A., MOHER, D. & HAYES, R. J. 2018. ‘Allocation concealment’: the evolution and adoption of a methodological term. *Journal of the Royal Society of Medicine,* 111**,** 216-224.
SCHULZ, K. F., CHALMERS, I., GRIMES, D. A. & ALTMAN, D. G. 1994. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *Jama,* 272**,** 125-128.
SCHULZ, K. F., CHALMERS, I., HAYES, R. J. & ALTMAN, D. G. 1995. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Jama,* 273**,** 408-412.
SCHULZ, K. F. & GRIMES, D. A. 2002a. Allocation concealment in randomised trials: defending against deciphering. *The Lancet,* 359**,** 614-618.
SCHULZ, K. F. & GRIMES, D. A. 2002b. Blinding in randomised trials: hiding who got what. *The Lancet,* 359**,** 696-700.
SCHULZ, K. F. & GRIMES, D. A. 2002c. Generation of allocation sequences in randomised trials: chance, not choice. *The Lancet,* 359**,** 515-519.
SCHULZ, K. F. & GRIMES, D. A. 2002d. Unequal group sizes in randomised trials: guarding against guessing. *The Lancet,* 359**,** 966-970.
SENN, C. Y., ELIASZIW, M., BARATA, P. C., THURSTON, W. E., NEWBY-CLARK, I. R., RADTKE, H. L. & HOBDEN, K. L. 2015. Efficacy of a sexual assault resistance program for university women. *New England journal of medicine,* 372**,** 2326-2335.
SENN, S. 1989. Covariate imbalance and random allocation in clinical trials. *Statistics in Medicine,* 8**,** 467-475.

SENN, S. 1994. Testing for baseline balance in clinical trials. *Statistics in medicine,* 13**,** 1715-1726.
SHARP, E. & CURLEWIS, K. 2019. Research waste is still a scandal—especially in medical students. *Bmj,* 364.
SILVERMAN, W. A. 1977. The lesson of retrolental fibroplasia. *Scientific American,* 236**,** 100-107.
STERNE, J. A., SAVOVIĆ, J., PAGE, M. J., ELBERS, R. G., BLENCOWE, N. S., BOUTRON, I., CATES, C. J., CHENG, H.-Y., CORBETT, M. S. & ELDRIDGE, S. M. 2019. RoB 2: a revised tool for assessing risk of bias in randomised trials. *bmj,* 366.
SWINGLER, G. H. & ZWAREENSTEIN, M. 2000. An effectiveness trial of a diagnostic test in a busy outpatients department in a developing country Issues around allocation concealment and envelope randomization. *Journal of clinical epidemiology,* 53**,** 702-706.
TRIALS. 2022. *Structured Study Protocol Template* [Online]. Available: https://trialsjournal.biomedcentral.com/submission-guidelines/preparing-your-manuscript/study-protocol/structured-study-protocol-template [Accessed 02.06 2022].
TROWMAN, R., DUMVILLE, J. C., HAHN, S. & TORGERSON, D. J. 2006. A systematic review of the effects of calcium supplementation on body weight. *British Journal of Nutrition,* 95**,** 1033-1038.
TROWMAN, R., DUMVILLE, J. C., TORGERSON, D. J. & CRANNY, G. 2007. The impact of trial baseline imbalances should be considered in systematic reviews: a methodological case study. *Journal of clinical epidemiology,* 60**,** 1229-1233.
TURNER, L., SHAMSEER, L., ALTMAN, D. G., SCHULZ, K. F. & MOHER, D. 2012. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review a. *Systematic reviews,* 1**,** 60.
VAN DER PAS, S. L. 2019. Merged block randomisation: A novel randomisation procedure for small clinical trials. *Clinical Trials,* 16**,** 246-252.
VIERA, A. J. & BANGDIWALA, S. I. 2007. Eliminating bias in randomized controlled trials: importance of allocation concealment and masking. *FAMILY MEDICINE-KANSAS CITY-,* 39**,** 132-137.
WOOD, L., EGGER, M., GLUUD, L. L., SCHULZ, K. F., JÜNI, P., ALTMAN, D. G., GLUUD, C., MARTIN, R. M., WOOD, A. J. & STERNE, J. A. 2008. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *Bmj,* 336**,** 601-605.
XU, L., LI, J., ZHANG, M., AI, C. & WANG, L. 2008. Chinese authors do need CONSORT: reporting quality assessment for five leading Chinese medical journals. *Contemporary clinical trials,* 29**,** 727-731.
YELLAND, L. N., KAHAN, B. C., DENT, E., LEE, K. J., VOYSEY, M., FORBES, A. B. & COOK, J. A. 2018. Prevalence and reporting of recruitment, randomisation and treatment errors in clinical trials: a systematic review. *Clinical Trials,* 15**,** 278-285.

Appendix 1: randomisation and allocation concealment details

Table 1: different randomisation methods and additional comments related to their implementation and reporting

	Example randomisation methods	Definition	How this may be described in a publication	Comments related to implementing and reporting
Adequate methods	**Central or third-party randomisation**	The recruiter contacts a central centre by telephone or via web access to enrol the participant to the study. The centre holds the randomisation sequence and allocates the participant to the RCT	*Third party telephone to secretariat of the clinical effectiveness unit at the royal college of surgeons, England. Computer randomisation without restriction (Brown et al., 2007)*	A full description of the methodology is still needed detailing who performed the randomisation The randomisation method should be stated even when 'central randomisation' has been implemented. For example blocking could have been undertaken and it is important that the block size is also stated. These are adequate allocation methods but an explanation of how they were concealed from trial personnel is also needed.
	Random number tables	A table of numbers arranged randomly in row and columns. The table is read into one way (by column) and the participant allocated a number. The number dictates which treatment group they are in.	*Patients were randomised to three month or six month follow up groups. After the research assistant collected baseline data and obtained consent for an eligible patient, the project coordinator assigned the patient of the particular doctor to three month or six month follow up from a predetermined list generated from a random number table (Birtwhistle et al., 2004).*	
	Computer generated random numbers	Similar to the above, a computer program generates a list of random numbers	*The nurse provided antibiotic (azithromycin) treatment, explained the trial, and asked for written consent. Those who agreed were randomised individually using computer generated random numbers in permuted blocks, stratified by practice (Low et al., 2006)*	
Problematic methods		The treatment allocation is placed in an envelope. Either 1.The allocation is placed in the envelopes are then shuffled	*The sequence was generated by one of us, who enclosed the numbers indicating the assignments in sealed, numbered envelopes. (Borghi et al., 2002)*	Envelopes are easy to subvert and need to be implemented robustly and reported transparently

	Envelopes	2.An allocation sequence has been created and the allocations are placed in the envelopes in accordance to the allocation sequence	*Randomization was performed with the use of sealed envelopes (Gæde et al., 2003).*	The details of who created the envelopes is needed as well as who opened them. A description of the envelope security measures is needed. Envelopes are frequently reported poorly and their description varies widely. Paper 4 and section 4 of this thesis discuss envelopes in further detail
	Restricted methods such as blocking	Randomisation occurs in blocks and prevents groups becoming numerically imbalanced. Eg a block size of 6. A =control group B = intervention group Examples blocks include: AAABBB. AABBAB. ABABAB	*Randomization was computer derived, with blocking into 3 groups to allow for orderly recruitment into both study groups and to reduce the risk of uneven recruitment late in the series (Dixon et al., 2008)* *Randomisation was done via a central telephone service using a computer-generated randomisation schedule prepared by a researcher not involved in treatment allocation, with balanced, variable blocks (sizes 2, 4, and 6), stratified by centre (Morris et al., 2016)*	Restriction of the allocation sequence can lead to prediction of the allocation sequence Stratification factor should be stated in publication The block size needs to be stated. The method of generating the allocation sequence still needs to be reported Paper 5 and section 5 discuss blocking in further details
	Minimisation and adaptative designs such as biased coin or urn randomisation	The first participant enters the trial and the treatment allocated to the next and subsequent participants enrolled on the characteristics of those participants already enrolled. This results in trial groups which are not imbalanced (Altman and Bland, 2005)	*A deterministic minimisation algorithm, based on side of haematoma and minimum depth from cortical surface, was initially used to ensure balance between the two groups, within every country where patients were recruited.(Mendelow et al., 2005)*	If Minimisation is used it should be fully reported in the publication including the variables used to allocate participants and how the random element was applied if necessary. It is accepted as a robust method even when there is no random element (Treasure and MacRae, 1998, Brown et al., 2005).

		Adaptative designs: Adaptative randomisation methods can be used to randomise participants to groups on several covariates so that the groups are balanced at baseline on potentially prognostic variables. It is particularly useful in small studies where simple randomisation or blocking with one or two stratification factors could result in imbalance (Frane, 1998). Covariate adaptive randomisation uses minimisation as it assesses the imbalance of sample size amongst several covariates (Suresh, 2011).	*Adaptative designs:* *The assignment to treatment groups was balanced according to age at entry (≤49 or ≥50 years), clinical tumor size (2.0 to 4 cm or ≥4.1 cm), hormone-receptor status (estrogen-receptor–positive, progesterone-receptor–positive, or both vs. estrogenreceptor–negative and progesterone-receptor–negative), and clinical nodal status (negative vs. positive). Randomization was performed within these strata, with the use of a biased-coin approach to ensure balanced treatment assignments within an institution.(Bear et al., 2012)* *Patients were randomly assigned in a 1:1 ratio to either interventional management of the pneumothorax (intervention group) or conservative management (conservative-management group), with stratification according to trial site; randomization was performed with the use of an adaptive biased-coin (urn) technique (Brown et al., 2020)*	*Adaptative designs:* Recommended that consultation with a statistician may be necessary (Schulz and Grimes, 2002b, Schulz and Grimes, 2002a, Berger et al., 2003) when assessing the robustness of the technique reported in a publication. Methodological details should be reported in full in the publication.
Inadequate methods	**Pseudo or quasi randomisation**	Participants were randomised using a known sequence or pattern, therefore is not truly random. Examples include alternation or month of birth	Consecutively enrolled patients were assigned in alternating order to receive daily or conventional hemodialysis. (Schiffl et al., 2002).	These are unacceptable as the allocation sequence is not concealed – the allocation sequence is open and easy for participants to be selected into one group or another as there is foreknowledge of the allocation assignment.

Table 2: Randomisation and allocation reporting: examples and an explanation of good and poor reporting of methodological details related to randomisation and allocation concealment

Reporting judgement	Publication description	Methodological details described in full in the publication		Explanation of reporting judgement
		Randomisation	Allocation concealment	
Good	*Randomization was conducted in block sizes of 4 and included equal numbers of women and men in each treatment group. Participants and investigators were not blinded to the randomization due to significant differences between formulations (eTable 1 in Supplement 2). Allocation concealment was not performed. (Matta et al., 2020)*	Yes	Yes	Randomisation details were provided, including the block size. Allocation concealment implementation was poor, however the reporting quality was good as the authors stated that allocation concealment was not undertaken.
	WHO prepared lists for randomisation using permuted blocks of variable length (6-8-10), with block sizes presented in random order. Separate randomisation lists were prepared for each site according to nutrition status of the children (severely malnourished, defined by WHO as oedema or severe wasting=weight for height <70% (−3 z score) or severe stunting=height for age <85% (−3 z score) versus not severely malnourished, as assessed during the baseline examination), and individual patient assignments were placed in opaque sealed envelopes. After each patient was selected for study, the next envelope in order of study numbers (that is, in numerical sequence) was opened to determine the treatment assignment: thus the investigator could not know the order of randomisation and was unable to predict the next assignment. *Before opening each envelope the doctor in charge signed and dated the opening flap of the envelope. The card inside, with the patient's treatment assignment,*	Yes	Yes	Full details of the randomisation and allocation concealment methods were reported

	and the signed envelope were attached to the patient's study file. To prevent tampering with the randomisation process, envelopes were checked to ensure that the assignment could not be seen before the envelope was opened. During site visits the presence of the signature, date, and time notification was evaluated and compared with the date and time of randomisation recorded in the medical record.(Asghar et al., 2008)			
Poor	*After random assignment to transplantation with peripheral blood cells or bone marrow, the patients were stratified according to treatment center, age («30 or >30 years), and stage of cancer (less advanced or more advanced). Within these strata, assignments were balanced in blocks of random size. (Bensinger et al., 2001)*	No	No	Blocking was undertaken, the stratification factors were reported but the block sizes were not stated, only 'blocks of random size'. No allocation concealment details were provided
	The nurse provided antibiotic (azithromycin) treatment, explained the trial, and asked for written consent. Those who agreed were randomised individually using computer generated random numbers in permuted blocks, stratified by practice. Allocation was concealed until assignment by using a central computerised telephone system, which the nurse had to telephone to obtain the assignment.(Low et al., 2006)	No	Yes	Allocation concealment described. The block size was not stated when describing the randomisation. If an insecure block size was used, the allocation sequence could have been predicted leading to potential selection bias
	Patients were randomly assigned, in a 1:1 ratio, to receive posaconazole or either fluconazole or itraconazole (Cornely et al., 2007)	No	No	No details provided on either randomisation or allocation concealment methods

References

ALTMAN, D. G. & BLAND, J. M. 2005. Treatment allocation by minimisation. *Bmj,* 330**,** 843.

ASGHAR, R., BANAJEH, S., EGAS, J., HIBBERD, P., IQBAL, I., KATEP-BWALYA, M., KUNDI, Z., LAW, P., MACLEOD, W. & MAULEN-RADOVAN, I. 2008. Chloramphenicol versus ampicillin plus gentamicin for community acquired very severe pneumonia among children aged 2-59 months in low resource settings: multicentre randomised controlled trial (SPEAR study). *Bmj,* 336**,** 80-84.

BEAR, H. D., TANG, G., RASTOGI, P., GEYER JR, C. E., ROBIDOUX, A., ATKINS, J. N., BAEZ-DIAZ, L., BRUFSKY, A. M., MEHTA, R. S. & FEHRENBACHER, L. 2012. Bevacizumab added to neoadjuvant chemotherapy for breast cancer. *New England Journal of Medicine,* 366**,** 310-320.

BENSINGER, W. I., MARTIN, P. J., STORER, B., CLIFT, R., FORMAN, S. J., NEGRIN, R., KASHYAP, A., FLOWERS, M. E., LILLEBY, K. & CHAUNCEY, T. R. 2001. Transplantation of bone marrow as compared with peripheral-blood cells from HLA-identical relatives in patients with hematologic cancers. *New England Journal of Medicine,* 344**,** 175-181.

BERGER, V. W., IVANOVA, A. & DELORIA KNOLL, M. 2003. Minimizing predictability while retaining balance through the use of less restrictive randomization procedures. *Statistics in Medicine,* 22**,** 3017-3028.

BIRTWHISTLE, R. V., GODWIN, M. S., DELVA, M. D., CASSON, R. I., LAM, M., MACDONALD, S. E., SEGUIN, R. & RÜHLAND, L. 2004. Randomised equivalence trial comparing three month and six month follow up of patients with hypertension by family practitioners. *Bmj,* 328**,** 204.

BORGHI, L., SCHIANCHI, T., MESCHI, T., GUERRA, A., ALLEGRI, F., MAGGIORE, U. & NOVARINI, A. 2002. Comparison of two diets for the prevention of recurrent stones in idiopathic hypercalciuria. *New England Journal of Medicine,* 346**,** 77-84.

BROWN, C. T., YAP, T., CROMWELL, D. A., RIXON, L., STEED, L., MULLIGAN, K., MUNDY, A., NEWMAN, S. P., VAN DER MEULEN, J. & EMBERTON, M. 2007. Self management for men with lower urinary tract symptoms: randomised controlled trial. *Bmj,* 334**,** 25.

BROWN, S., THORPE, H., HAWKINS, K. & BROWN, J. 2005. Minimization—reducing predictability for multi-centre trials whilst retaining balance within centre. *Statistics in medicine,* 24**,** 3715-3727.

BROWN, S. G., BALL, E. L., PERRIN, K., ASHA, S. E., BRAITHWAITE, I., EGERTON-WARBURTON, D., JONES, P. G., KEIJZERS, G., KINNEAR, F. B. & KWAN, B. C. 2020. Conservative versus interventional treatment for spontaneous pneumothorax. *New England Journal of Medicine*.

CORNELY, O. A., MAERTENS, J., WINSTON, D. J., PERFECT, J., ULLMANN, A. J., WALSH, T. J., HELFGOTT, D., HOLOWIECKI, J., STOCKELBERG, D. & GOH, Y.-T. 2007. Posaconazole vs. fluconazole or itraconazole prophylaxis in patients with neutropenia. *New England Journal of Medicine,* 356**,** 348-359.

DIXON, J. B., O'BRIEN, P. E., PLAYFAIR, J., CHAPMAN, L., SCHACHTER, L. M., SKINNER, S., PROIETTO, J., BAILEY, M. & ANDERSON, M. 2008. Adjustable gastric banding and conventional therapy for type 2 diabetes: a randomized controlled trial. *Jama,* 299**,** 316-323.

FRANE, J. W. 1998. A method of biased coin randomization, its implementation, and its validation. *Drug information journal: DIJ/Drug Information Association,* 32**,** 423-432.

GÆDE, P., VEDEL, P., LARSEN, N., JENSEN, G. V., PARVING, H.-H. & PEDERSEN, O. 2003. Multifactorial intervention and cardiovascular disease in patients with type 2 diabetes. *New England Journal of Medicine,* 348**,** 383-393.

LOW, N., MCCARTHY, A., ROBERTS, T. E., HUENGSBERG, M., SANFORD, E., STERNE, J. A., MACLEOD, J., SALISBURY, C., PYE, K. & HOLLOWAY, A. 2006. Partner notification of chlamydia infection in primary care: randomised controlled trial and analysis of resource use. *Bmj,* 332**,** 14-19.

MATTA, M. K., FLORIAN, J., ZUSTERZEEL, R., PILLI, N. R., PATEL, V., VOLPE, D. A., YANG, Y., OH, L., BASHAW, E. & ZINEH, I. 2020. Effect of sunscreen application on plasma concentration of sunscreen active ingredients: a randomized clinical trial. *Jama,* 323**,** 256-267.

MENDELOW, A. D., GREGSON, B. A., FERNANDES, H. M., MURRAY, G. D., TEASDALE, G. M., HOPE, D. T., KARIMI, A., SHAW, M. D. M., BARER, D. H. & INVESTIGATORS, S. 2005. Early surgery versus initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas in the International Surgical Trial in Intracerebral Haemorrhage (STICH): a randomised trial. *The Lancet,* 365**,** 387-397.

MORRIS, J. M., ROBERTS, C. L., BOWEN, J. R., PATTERSON, J. A., BOND, D. M., ALGERT, C. S., THORNTON, J. G., CROWTHER, C. A. & COLLABORATION, P. 2016. Immediate delivery compared with expectant management after preterm pre-labour rupture of the membranes close to term (PPROMT trial): a randomised controlled trial. *The Lancet,* 387**,** 444-452.

SCHIFFL, H., LANG, S. M. & FISCHER, R. 2002. Daily hemodialysis and the outcome of acute renal failure. *New England Journal of Medicine,* 346**,** 305-310.

SCHULZ, K. F. & GRIMES, D. A. 2002a. Generation of allocation sequences in randomised trials: chance, not choice. *The Lancet,* 359**,** 515-519.

SCHULZ, K. F. & GRIMES, D. A. 2002b. Unequal group sizes in randomised trials: guarding against guessing. *The Lancet,* 359**,** 966-970.

SURESH, K. 2011. An overview of randomization techniques: an unbiased assessment of outcome in clinical research. *Journal of human reproductive sciences,* 4**,** 8.

TREASURE, T. & MACRAE, K. D. 1998. Minimisation: the platinum standard for trials?: randomisation doesn't guarantee similarity of groups; minimisation does. British Medical Journal Publishing Group.

Appendix 2: worked example of an RCT explaining why it could be impacted by selection bias

Title of RCT: *Efficacy of a Sexual Assault Resistance Program for University Women (Senn et al., 2015)*

Aim: To assess whether a new, four-unit, small-group sexual assault resistance program as compared with access to brochures on sexual assault, could reduce the 1-year incidence of completed rape among first-year female students at three universities.

Methodological information on randomisation and allocation concealment included in publication: Randomization was performed in permuted blocks of two with the use of the online tool Randomize.net, with stratification according to site

Number of sites: three

Number and allocation of participants: 452 into the control, 464 into the intervention group

Conclusion: 'Rigorously designed and executed sexual assault resistance program was successful in decreasing the occurrence of rape, attempted rape, and other forms of victimization among first-year university women.'

Methodological analysis and commentary

A block size of two was used which is considered a block size that should be avoided due to the chance of predicting the allocation sequence. Additionally, the authors state that a block of two stratified by site was undertaken then the allocation will be perfectly balanced at each site if an even number of women were recruited at each site. If recruitment finished mid-way through a block at each site, then with three sites the biggest imbalance across the trial should be three. In this trial the imbalance between groups was 12 which is not possible with the block size employed, this suggests that an insecure allocation system was used and the results could be impacted by selection bias.

Without further understanding regarding this imbalance the conclusions made in the paper should be treated with caution.

Considering the wider picture:

Participants: Participants have given their time and confidential data to take part in a research study with methodological flaws. The conclusions may be biased and impacted by selection bias, this strategy may not decrease rape, attempted rape and victimisation in this participant group.

Researchers: There are research teams designing research with insecure allocation concealment (due to a block size of two being implemented) and not transparently explaining the numerically impossible group assignments when reporting the RCT. Without further information it suggests that recruiters could have selected participants into one of the two treatment groups.

Journals: The reviewing process did not identify the methodological flaws of this RCT, it resulted in the RCT being accepted for publication in a high impact medical journal.

References

SENN, C. Y., ELIASZIW, M., BARATA, P. C., THURSTON, W. E., NEWBY-CLARK, I. R., RADTKE, H. L. & HOBDEN, K. L. 2015. Efficacy of a sexual assault resistance program for university women. *New England journal of medicine,* 372**,** 2326-2335.

Appendix 3: paper 1

A methodological review of recent meta-analysis has found significant heterogeneity in age between randomised groups

Appendices for paper 1:

Appendix 1. Reference list of RCTs included in our review from each of the 12 Systematic Reviews.

Appendix 2: Forest plots for the meta-analysis of baseline age for each of the 12 Systematic Reviews

Appendix 3. Bubble plots and meta-regression statistics for each of the 12 Systematic review.

Reference

Clark, L., Fairhurst, C., Hewitt, C.E., Birks, Y., Brabyn, S., Cockayne, S., Rodgers, S., Hicks, K., Hodgson, R., Littlewood, E. and Torgerson, D.J., 2014. A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups. *Journal of clinical epidemiology*, *67*(9), pp.1016-1024.

Journal of Clinical Epidemiology 67 (2014) 1016−1024

Journal of Clinical Epidemiology

A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups

Laura Clark[a,*], Caroline Fairhurst[a], Catherine E. Hewitt[a], Yvonne Birks[b], Sally Brabyn[c], Sarah Cockayne[a], Sara Rodgers[a], Katherine Hicks[a], Robert Hodgson[a], Elizabeth Littlewood[c], David J. Torgerson[a]

[a]*Department of Health Sciences, York Trials Unit, University of York, York YO10 5DD, United Kingdom*
[b]*Department of Social Policy, University of York, York YO10 5DD, United Kingdom*
[c]*Department of Health Sciences, Mental Health and Addictions Research Group, University of York, York YO10 5DD, United Kingdom*

Accepted 23 April 2014; Published online 6 June 2014

Abstract

Background: There is evidence to suggest that component randomized controlled trials (RCTs) within systematic reviews may be biased. It is important that these reviews are identified to prevent erroneous conclusions influencing health care policies and decisions.

Purpose: To assess the likelihood of bias in trials in 12 meta-analyses.

Design: A review of 12 systematic reviews.

Data Sources: Twelve recently published systematic reviews with 503 component randomized trials, published in the *British Medical Journal*, *The Lancet*, *Journal of the American Medical Association*, and *The Annals of Internal Medicine* before May 2012.

Study Selection and Data Extraction: Systematic reviews were eligible for inclusion if they included only RCTs. We obtained the full text for the component RCTs of the 12 systematic reviews (in English only). We extracted summary data on age, number of participants in each treatment group, and the method of allocation concealment for each RCT.

Data Synthesis: Five of the 12 meta-analyses exhibited heterogeneity in age differences ($I^2 > 0.30$), when there should have been none. In two meta-analyses, the age of the intervention group was significantly greater than that of the control group. Inadequate allocation concealment was a statistically significant predictor of heterogeneity in one trial as observed by a metaregression.

Conclusions: Most of the sample of recent meta-analyses showed that there were signs of imbalance and/or heterogeneity in ages between treatment groups, when there should have been none. Systematic reviewers might consider using the techniques described here to assess the validity of their findings.

Keywords: Systematic review; Selection bias; Randomized controlled trials; Methods; Meta-analysis; Heterogeneity

1. Introduction

Ideally, a systematic review and meta-analysis of randomized controlled trials (RCTs) should be used to inform clinical practice [1]. However, there is good evidence to suggest that some RCTs in systematic reviews may be biased because of insecure allocation concealment [2−8]. The assessment of whether a meta-analysis contains such biased trials is important in relation to the validity of the findings. One approach outlined by Trowman et al. is to undertake a meta-analysis of an important baseline variable, in their case body weight [9], incorporating all the trials in the review. They found that there was significant imbalance, which explained virtually all the difference between groups at follow-up. This approach of using a baseline variable, which randomization should ensure would differ only by chance, allows us to assess whether a meta-analysis of RCTs is reliable. To date, techniques using baseline information are not routinely used to assess the quality of meta-analyses. Trowman et al. used baseline body weight because this was the most powerful predictor of the outcome (ie, final body weight). However, in many meta-analyses, the outcome variable either is not reported at randomization or is not possible to collect. Another potentially powerful predictor of outcome is age. Age is usually a good predictor of outcome; older people tend to have worse outcomes than younger people. Furthermore, if one wishes to subvert the allocation of a trial then consciously or unconsciously misallocating according to a person's age is

Conflicts of interest: The authors confirm that they have no conflict or financial interest in the publication of this manuscript.

* Corresponding author. Tel.: +44-1904-321115 fax: +44-1904-321387;.

E-mail address: Laura.clark@york.ac.uk (L. Clark).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2014.04.007

What is new?

What this adds to what was known?

- A previous meta-analysis of RCTs has found a statistically significant difference in a baseline variable. This study adds to this single finding by looking at a sample of 12 recent meta-analyses of RCTs and shows that a majority of studies show either heterogeneity or baseline imbalance in age when there should be none.

What this adds to what was known?

- This study suggests that there is a widespread problem among meta-analyses of RCTs.

What is the implication and what should change now?

- All meta-analysis of RCTs should undertake a meta-analysis of key baseline variables to check for imbalance and heterogeneity.

relatively simple to do. Finally, we expect that most trials should report the mean age with standard deviations (SDs) of participants by group allocation. This study, therefore, is an advance on the article by Trowman et al. Furthermore, Trowman et al. were mainly interested in baseline imbalance rather than heterogeneity (where they found none). Whether there is baseline imbalance, whilst of concern, is probably not as important as whether baseline heterogeneity exists. Significant heterogeneity of baseline age within a systematic review of randomized trials would indicate that the whole review is unsound, because it includes trials with biased randomization. Consequently, results from such a review must be treated with a great deal of caution. In this article, therefore, we assess heterogeneity in age between randomized groups in a sample of recently published meta-analyses to assess whether these are at risk of bias. This review is an exemplar to investigate the extent of the issue in a set of systematic reviews published in high-impact, prestigious journals.

2. Methods

In editions dating back from May 2012, we identified the first three systematic reviews containing only RCTs published in each of *The Annals of Internal Medicine*, The *British Medical Journal*, The *Journal of the American Medical Association*, and *The Lancet*. These journals were selected, because they are the highest impact medical journals that frequently publish meta-analyses of RCTs. Other high-impact journals, such as the *New England Journal of Medicine* do not publish high numbers of meta-analyses.

2.1. Data extraction

Where available, the full-text articles of the component RCTs from the 12 reviews available in English were retrieved. In addition to the method of allocation concealment, the following information was extracted for each trial arm where possible: summary of age (mean or median), its measure of dispersion (SD, standard error [SE], range, and interquartile range), and number of participants. Double, independent data extraction was performed by two researchers (ie, authors of this article); any disagreements were resolved by discussion. No other aspect of trial design was extracted from the articles, and we did not contact the trial authors of randomized trials where data were not available in the published articles.

2.1.1. Allocation concealment

Adequacy of allocation concealment was judged using the Cochrane handbook criteria [10]. Trials were classified as adequate, inadequate, or unclear. The reviewers made a judgment as to the quality of allocation concealment without knowing whether there was an age imbalance for that appraised trial.

2.1.2. Age

If age was not summarized using the mean and SD, for example, if median and range were presented, measures were converted using standard approximation formulas [11]. A fixed effect meta-analysis of age was performed for each review on the assumption that there was a common treatment estimate (ie, zero) across the randomized trials. The *P*-value for the difference in age between the control and intervention groups for each systematic review was calculated. The I^2 value of heterogeneity from the meta-analysis was interpreted in line with the Cochrane handbook guidelines: 0%−40% might not be important; 30%−60% may represent moderate heterogeneity; 50%−90% may represent substantial heterogeneity; and 75%−100% considerable heterogeneity [12]. A metaregression was performed for each review to assess whether allocation concealment adequacy was a predictor of heterogeneity. Meta-analyses and regression were performed in Comprehensive Meta-Analysis 2 (Biostat, Englewood, NJ, USA).

If the null hypothesis is true, the *P*-values from independent two-sample *t*-tests should follow a uniform distribution reasonably well provided the data on which the *t*-tests are carried out are normally distributed and the tests are independent [13]. We applied this theory here. A systematic reviewer can assess for differences in a continuous baseline covariate between the control and intervention groups using a *t*-test for each RCT in their review. For each test, the null hypothesis should be true; therefore, statistically testing differences in the control and intervention groups at baseline across trials included in a systematic review should yield *P*-values with a uniform distribution if the randomizations were all faithful. That is, ~10% of the

P-values should lie within each interval of size 0.1. For each systematic review, we compared the distribution of *P*-values obtained from a *t*-test for the age of the control and intervention groups from each component trial to that expected using the chi-square goodness-of-fit test. A further analysis for continuous variables is based on the principle of the central limit theorem (CLT), and its use has previously been described in a similar context [14]. The CLT states that the mean value of samples taken repeatedly from the same population will be approximately normally distributed around the mean of the population. The SD of this curve can be estimated by the standard error of the mean (SEM) of any of the samples used to construct the curve, in the relation SEM $= SD_s/\sqrt{n_s}$, where SD_s is the standard deviation of the sample s of size n_s. We calculated the trial population mean age and pooled SD from the summary data for each trial and standardized the mean of the control group$\overline{X}$ $[(\overline{X} - \mu)/\text{SEM})] - \mu)/\text{SEM})]$. The standardized means should follow a curve with mean 0 and SD 1. We plotted histograms of these standardized sample means for each review overlaid by a standard normal curve and tested whether their distribution differed from the expected distribution using a test for the equality of variance (sdtest, Stata v12; StataCorp LP, College Station, Texas, USA). Deviations from this expected distribution could suggest that participants randomized to the control groups in the component RCTs of a particular review were statistically significantly older or younger than the population mean. Histograms were created in the statistical package R.

Sensitivity analyses were performed as described previously but omitting trials where the mean and/or SD of age had been derived from approximation formulas.

3. Results

Fig. 1 summarizes the flow of systematic reviews and component randomized trials in this review. The 12 systematic reviews included a total of 503 trials [15–26]; 184 trials were not available, so 319 have been included in this analysis. Reasonable attempts were made to retrieve all texts through the University of York library service, but in certain circumstances, this was not possible and it was not deemed necessary to expend excessive monetary or time resources on sourcing texts, because this review is just an exemplar of the techniques described. Of the trials included, 20% ($n = 64$) had adequate allocation concealment, 4% ($n = 14$) inadequate, and 76% ($n = 243$) failed to provide sufficient information for a judgment to be made (see Appendix A at www.jclinepi.com for list of included trials). A further 110 trials were excluded from the meta-analyses because of baseline age data being unavailable for extraction.

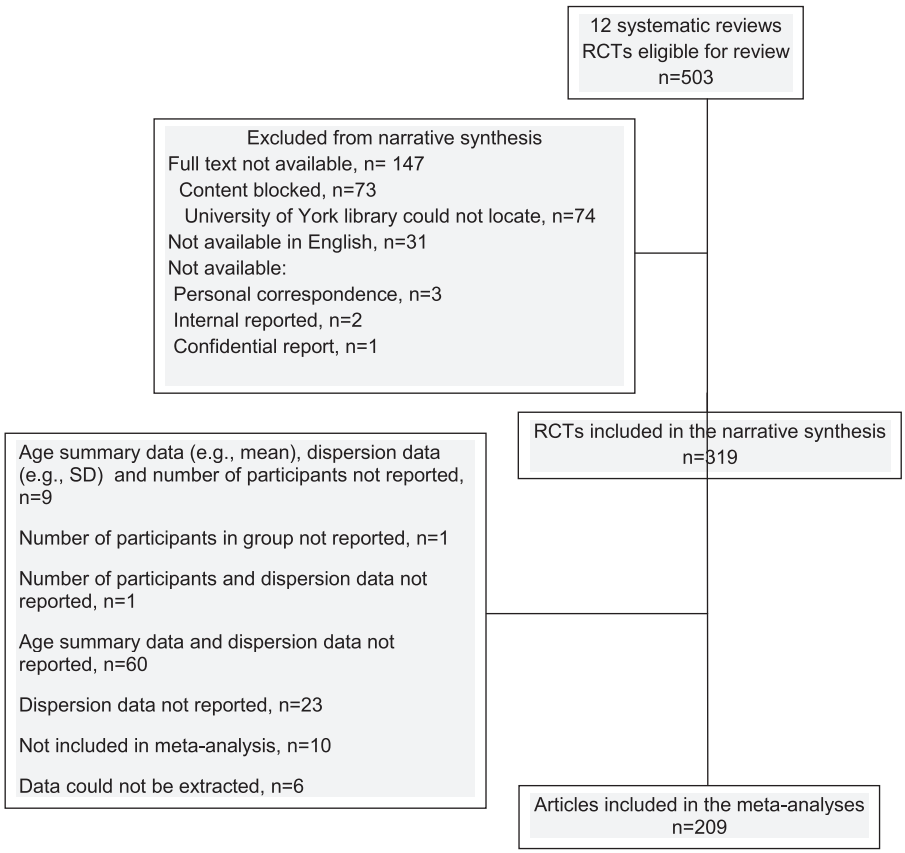


Fig. 1. Flow diagram of included systematic reviews.

Table 1. Table showing the mean (standard deviation [SD]) age in years for the intervention and control groups in each included systematic reviews with the I^2 value from the meta-analysis of baseline age and the *P*-value for the difference in baseline age ranked in order of heterogeneity

Systematic review	Number of studies in meta-analysis	Research area	Intervention	Control	Difference in age (*P*-value)	I^2 value
			Age, yr, mean (SD)			
Anothaisintawee et al. 2012	10	Drug	44.85 (5.56)	42.84 (5.67)	0.001	84.42
Thangaratinam et al. 2012	20	Pregnancy and childbirth	28.15 (2.27)	27.95 (2.05)	0.113	50.11
Umpierre et al. 2011	26	Lifestyle	58.29 (4.27)	58.79 (4.44)	0.098	45.46
Heneghan et al. 2011	7	Other	64.17 (7.61)	63.97 (9.06)	0.223	40.13
Neumann et al. 2012	9	Drug	64.18 (2.45)	63.94 (2.94)	0.821	33.46
Palmer et al. 2012	11	Drug	51.99 (8.35)	52.86 (8.95)	0.173	29.03
Rutjes et al. 2012	38	Drug	62.17 (4.34)	62.44 (3.82)	0.616	20.39
Orrow et al. 2012	10	Lifestyle	62.57 (10.29)	62.82 (9.72)	0.736	16.18
Hemmingsen et al. 2012	13	Drug	57.65 (3.97)	58.54 (4.14)	0.347	0.00
Coombes et al. 2010	18	Drug	48.08 (6.9)	48.08 (7.25)	0.362	0.00
Leucht et al. 2012	21	Drug	40.31 (9.24)	39.92 (9.78)	0.008	0.00
Hempel et al. 2012	26	Drug	41.84 (24.43)	42.19 (25.24)	0.818	0.00

3.1. Meta-analysis

Of the 12 included systematic reviews, we found that five had heterogeneity ($I^2 > 0.30$) [15,19,21,25,26], including one with substantial heterogeneity ($0.50 < I^2 \leq 0.75$) [25] and one with considerable heterogeneity ($I^2 > 0.75$) [15]. Two systematic reviews [15,20] had a statistically significant difference in age between the control and intervention groups overall. In each instance, the mean age in the intervention group was greater than that of the control. Table 1 presents the results for age of the meta-analysis ranked in order of heterogeneity and forest plots for each systematic review can be found in Appendix B at www.jclinepi.com. A forest plot for one particular review can be found in Fig. 2 as an example. When sensitivity analyses were performed omitting the trials where approximation formulas had been used to convert median age to mean age, there was no difference in results observed; that is, the results remained the same whether we included trials where we had to use approximations to the mean and SD.

3.2. P-value comparison and histograms of standardized control group means of baseline age

There was evidence of a statistically significant discrepancy between the distributions of *P*-values observed compared with that expected in two of the reviews [24,25].

In six of the reviews, the distribution exhibited a statistically significant deviation from that expected ($P < 0.01$) [15,21,23–26] when histograms of standardized control groups means at baseline were investigated. Fig. 2 shows an example of these two plots for one particular review.

3.3. Metaregression results

Metaregression analysis suggested that allocation concealment explained the heterogeneity observed in one review ($P = 0.03$; 15). See Appendix C at www.jclinepi.com for the metaregression statistics and bubble plots for each review.

In Table 2, we summarize the conclusions made in the included 12 systematic reviews and the findings from our analysis. The dots indicate the presence of evidence for the marker of potential bias in the RCT, of which 8 of the 12 reviews had at least one suggesting the conclusions drawn may not be reliable.

4. Discussion

Systematic reviews of randomized controlled trials are considered the highest form of evidence that underpins evidence-based medicine. There is a wealth of research that shows that some randomized trials have had their random allocation subverted. Much of this evidence examines the relationship between allocation description and effect sizes. Broadly, such evidence shows that descriptions of rigorous methods of allocation tend to be associated with smaller effect sizes. In this article, we have examined the problem using a different approach. We hypothesized that a systematic review that contains a significant proportion of trials with biased randomization will exhibit significant heterogeneity in a baseline variable, in this instance age. Such an approach does not rely on detailed descriptions of the randomization process, which is often missing or may be false. We argue that our approach is easier and more sensitive than qualitatively making a judgment about the randomization procedure from descriptions in the article and should be undertaken routinely as a marker for quality of a systematic review of randomized trials. In this article, we have found a surprising and disturbing result. In 12 recent systematic reviews, we found that only four reviews demonstrated the expected zero heterogeneity [16,18,22,24]. All other reviews had at least one ‘‘symptom’’ of the following: significant heterogeneity in baseline age; significant differences in baseline age; or an unlikely distribution of *P*-values or standardized means of baseline

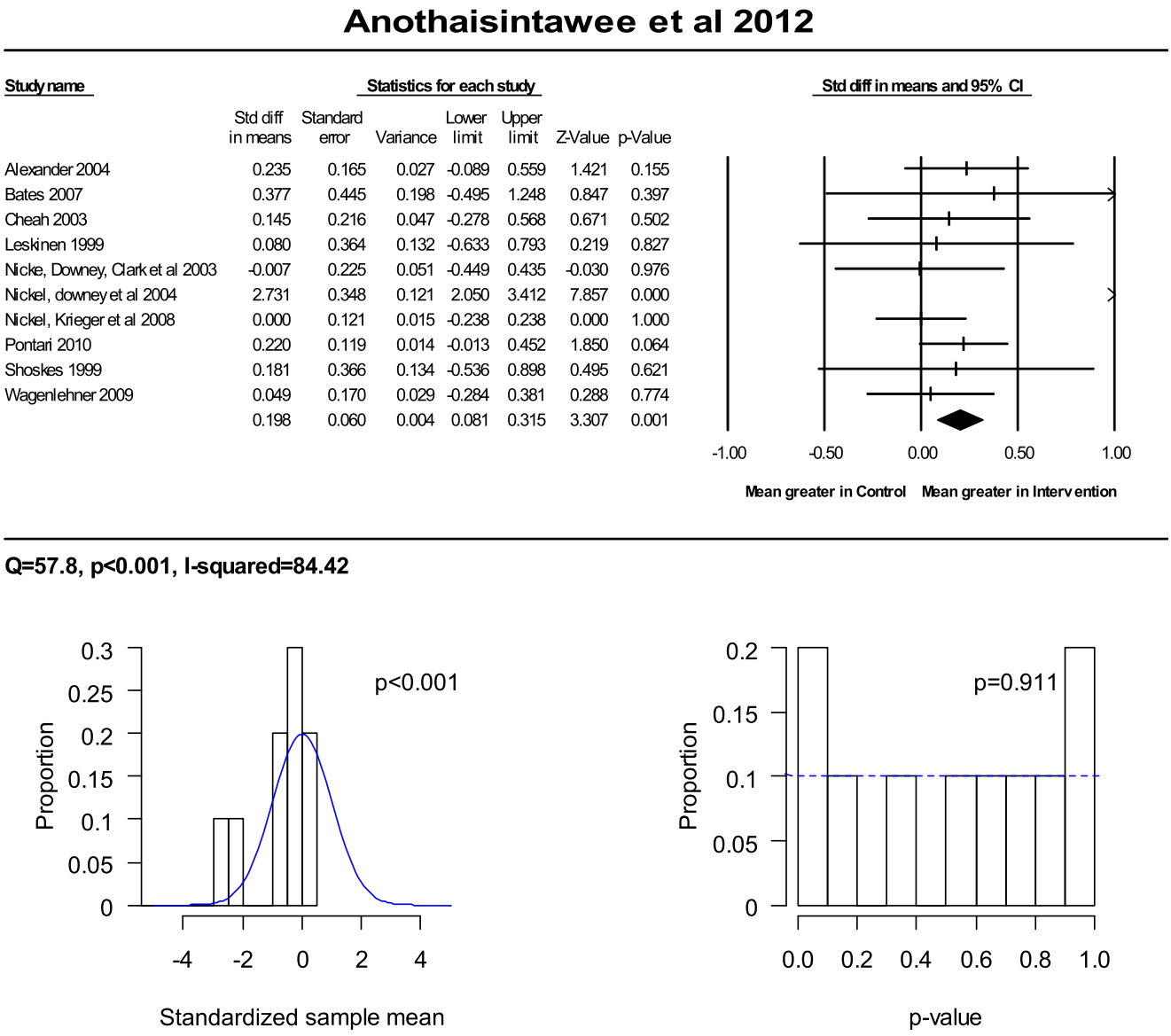


Fig. 2. Example analysis for one systematic review.

age. Two of the reviews [15,25] had three factors and three had two factors [21,24,26] that suggest that some of their component RCTs were unreliable. Because all the trials in these reviews were reported to have been randomized, we would expect that the null hypothesis of no age differences between groups except by chance would be true. Thus, we should have observed no heterogeneity in ages between groups; a meta-analysis of age should show equivalence and the distribution of *P*-values of baseline ages should be uniform. As we had used a baseline variable common to the trials and not an outcome variable, explanations for heterogeneity often found in meta-analyses, such as different populations or slightly different interventions leading to differing effect sizes, do not apply in this instance. The most plausible explanation for heterogeneity is therefore poor randomization practice. Although we did explore this in a metaregression, it was only statistically significant as an explanation in one study. This may be because allocation concealment was so poorly reported in most randomized trials that it is masked as a source of the heterogeneity. Although some of the "statistically significant" differences in age are relatively minor, the fact that they exist is important because they act as a marker for poor allocation practice. Because a systematic review contains some trials that have misallocation, the entire review is weak and should not be used to drive major changes in clinical practice.

4.1. Limitations

We were unable to examine all the RCTs that were included in the original systematic reviews, as a number

Table 2. Conclusions made in each systematic review with the results of each test applied in this current review

Systematic review	Evidence of heterogeneity ($I^2 > 0.3$)	Statistically significant difference in age between control and intervention groups	Distribution of *P*-values deviated from uniform distribution	Distribution of standardized control means deviates from standard normal	Conclusion of systematic review taken directly from the abstract
Anothaisintawee et al. 2012	Yes	Yes, $P < 0.01$		Yes, $P < 0.001$	α-Blockers, antibiotics and combinations of these therapies appear to achieve the greatest improvement in clinical symptom scores compared with placebo. Anti-inflammatory therapies have a lesser but measurable benefit on selected outcomes. However, beneficial effects of α-blockers may be overestimated because of publication bias.
Rutjes et al. 2012			Yes, $P < 0.05$	Yes, $P < 0.01$	In patients with knee osteoarthritis, viscosupplementation is associated with a small and clinically irrelevant benefit and an increased risk for serious adverse events.
Hemmingsen et al. 2012					There was no evidence or even a trend toward improved all-cause mortality or cardiovascular mortality with metformin and insulin, compared with insulin alone in type 2 diabetes. Data were limited by the severe lack of data reported by trials for patient relevant outcomes and by poor bias control.
Thangaratinam et al. 2012	Yes		Yes, $P < 0.05$	Yes, $P < 0.01$	Dietary and lifestyle interventions in pregnancy can reduce maternal gestational weight gain and improve outcomes for both mother and baby. Among the interventions, those based on diet are the most effective and are associated with reductions in maternal gestational weight gain and improved obstetric outcomes.
Umpierre et al. 2011	Yes			Yes, $P < 0.001$	Structured exercise training that consists of aerobic exercise, resistance training, or both combined is associated with HbA_{1c} reduction in patients with type 2 diabetes. Structured exercise training of more than 150 minutes per week is associated with greater HbA_{1c} declines than that of 150 minutes or less per week. Physical activity advice is associated with lower HbA_{1c}, but only when combined with dietary advice.
Neumann et al. 2012	Yes			Yes, $P < 0.001$	Compared with LMWH, lower doses of oral factor Xa inhibitors can achieve a small absolute risk reduction in symptomatic deep venous thrombosis without increasing bleeding.
Heneghan et al. 2011	Yes				Our analysis showed that self-monitoring and self-management of oral anticoagulation is a safe option for suitable patients of all ages. Patients should also be offered the option to self-manage their disease with suitable health-care support as back-up.
Palmer et al. 2012				Yes, $P < 0.01$	Benefits for antiplatelet therapy among persons with CKD are uncertain and are potentially outweighed by bleeding hazards

(*Continued*)

Table 2. Continued

Systematic review	Evidence of heterogeneity ($I^2 > 0.3$)	Statistically significant difference in age between control and intervention groups	Distribution of *P*-values deviated from uniform distribution	Distribution of standardized control means deviates from standard normal	Conclusion of systematic review taken directly from the abstract
Orrow et al. 2012					Promotion of physical activity to sedentary adults recruited in primary care significantly increases physical activity levels at 12 months, as measured by self report. We found insufficient evidence to recommend exercise referral schemes over advice or counseling interventions. Primary care commissioners should consider these findings while awaiting further trial evaluation of exercise referral schemes and other primary care interventions, with longer follow-up and use of objective measures of outcome.
Coombes et al. 2010					Despite the effectiveness of corticosteroid injections in the short term, non-corticosteroid injections might be of benefit for long-term treatment of lateral epicondylalgia. However, response to injection should not be generalised because of variation in effect between sites of tendinopathy.
Leucht et al. 2012		Yes, $P < 0.01$			Maintenance treatment with antipsychotic drugs benefits patients with schizophrenia. The advantages of these drugs must be weighed against their side-effects. Future studies should focus on outcomes of social participation and clarify the long-term morbidity and mortality of these drugs.
Hempel et al. 2012					The pooled evidence suggests that probiotics are associated with a reduction in AAD. More research is needed to determine which probiotics are associated with the greatest efficacy and for which patients receiving which specific antibiotics.

Abbreviations: AAD, antibiotic-associated diarrhea; HbA_{1c}, hemoglobin A1c; LMWH, low-molecular-weight heparin.

could not be obtained for inclusion in the current review. Some were confidential reports or were reported in another language than English. Research suggests that methodological quality is similar or higher in English language publications compared with those that are published in other languages [27–30]; therefore, any findings here are likely to at least replicate what may be found in other language articles.

We had relatively small sample sizes for some of our meta-analyses because of the exclusion of a number of RCTs where age data were not reported ($n = 186$). However, randomized trials that did not report the mean ages with their associated SDs and group numbers are likely to be poorer in quality than well-reported RCTs, so it is possible the heterogeneity we observed may have increased had we been able to include them. To perform the metaregression analysis, 10 or more trials are required for the test to be reliable. In some of our metaregression, we did not have 10 trials, and so the results should be interpreted with caution.

Although we found some evidence that poor allocation concealment was likely to be a driver for the heterogeneity, this finding was only statistically significant in one review. There are difficulties in the judgment of allocation concealment as different criteria and scales are used to judge the adequacy of methods [31]. We used only the Cochrane criteria, so it was possible that had we chosen to use a different criteria, we may have obtained different results. However as previously emphasized, a large proportion of the trial publications inadequately reported the method of allocation concealment so that its adequacy could not be judged by any criteria.

Randomization of participants to a trial arm should ensure that comparisons between randomized groups for

independent variables at baseline produce a uniform distribution of *P*-values. Deviations from this expected distribution could hint at a failure of randomization. We suggest that this method can be used to assess the reliability of randomization; however, caution must be taken, because the *P*-values for baseline tests will follow a uniform distribution only if the tests used to compute them are valid and independent [13]. The distribution of continuous variables being compared between the groups must be approximately normal for the *t*-test to be valid, which age seldom is within trial participant populations because for one, upper and/or lower age restrictions are often imposed by the inclusion criteria. This method could more reliably be applied to other prognostic factors known to be approximately normally distributed within trial groups. It is pertinent to acknowledge, however, that this review observed difficulties in accessing all necessary age data, and it may be the case that prognostic factors are not reported well or at all. It would be possible to check these assumptions explicitly if you were conducting a review and had access to individual participant data from each component trial, as opposed to just the reported aggregate summaries.

4.2. Recommendations

We recommend that meta-analysts use the techniques described in this article to check the validity of their analyses. We suggest that all four techniques should be used because a single approach may not be sufficient. For example, in the review by Trowman et al., the heterogeneity in baseline body weight was 0%; similarly, the review by Leucht et al. showed 0% heterogeneity yet both these reviews found statistically significant imbalances in age [9,20]. A reviewer should expect to see no heterogeneity and no statistically significant differences in the outcome between the groups. If a review showed a significant difference in baseline measure of outcome but no heterogeneity, the conclusion from this is that a proportion of the included trials have allocation subversion all favoring the same treatment arm. However, if there were simply significant heterogeneity but no overall difference in the baseline covariate, this does not imply that because subversion is operating in both directions that the review's results are believable. It may mean that the trials are in significant imbalance in a particular direction that favors an unknown covariate. In truly randomized trials, this unknown or unmeasured covariate will be balanced across a group of trials, but when heterogeneity is present, we cannot be confident that this holds true. If a meta-analysis has significant numbers of randomized trials where the allocation is subverted, there may be no heterogeneity because the ''true'' difference is a difference in age so no heterogeneity is observed. Nevertheless, such trials are biased and the meta-analysis is unreliable. Current practice for estimating bias in meta-analyses is to grade component trials as being at high or low risk of bias using measures such as the Jadad scale [32] or the Cochrane guidance. Unfortunately, it is likely these scales will misclassify at least some trials as either being good, when they are poor, or vice versa. The approach we recommend here should be used as a complementary method to assess the rigor of a meta-analysis as it gives more information to the reader. We have chosen the prognostic factor of age; however, there are other potentially more important variables on which to do baseline testing. We would recommend that reviewers perform baseline testing on an important, prespecified, prognostic factor relative to their review (eg, if the review was in the field of obesity, body weight could be the baseline characteristic tested) where possible. Finally, we recommend that authors of systematic reviews should prespecify which baseline variables they choose to include in their analysis in their systematic review protocol. This is to avoid undertaking multiple baseline tests and only presenting those that have little or no heterogeneity.

In summary, there is significant unexplained heterogeneity of age in most of the sample of systematic reviews published in high-impact journals. Reviewers should adopt techniques to identify potential baseline imbalances in their trials and use this to drive sensitivity analyses.

Appendix

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2014.04.007.

References

[1] Trowman R, Dumville JC, Hahn S, Torgerson DJ. A systematic review of the effects of calcium supplementation on body weight. Br J Nutr 2006;95(6):1033−8.
[2] Altman DG, Schulz KF. Statistics notes: concealing treatment allocation in randomised trials. BMJ 2001;323:446−7.
[3] Attia A. Bias in RCTs: confounders, selection bias and allocation concealment. Middle East Fertil Soc J 2005;10(3):258−61.
[4] Doig GS, Simpson F. Randomization and allocation concealment: a practical guide for researchers. J Crit Care 2005;20(2):187−91.
[5] Hewitt C, Hahn S, Torgerson DJ, Watson J, Bland JM. Adequacy and reporting of allocation concealment: review of recent trials published in four general medical journals. BMJ 2005;330:1057−8.
[6] Hewitt CE, Torgerson DJ, Berger VW. Potential for technical errors and subverted allocation can be reduced if certain guidelines are followed: examples from a web-based survey. J Clin Epidemiol 2009; 62:261−9.
[7] Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet 2002;359:614−8.
[8] Viera AJ, Bangdiwala SI. Eliminating bias in randomized controlled trials: importance of allocation concealment and masking. Fam Med 2007;39(2):132−7. KANSAS CITY.
[9] Trowman R, Dumville JC, Torgerson DJ, Cranny G. The impact of trial baseline imbalances should be considered in systematic reviews: a methodological case study. J Clin Epidemiol 2007;60:1229−33.
[10] Higgins JPT, Altman DG, Sterne J. Chapter 8: assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0

(updated March 2011). The Cochrane Collaboration; 2011. Available at www.cochrane-handbook.org. 2011.

[11] Hozo S, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol 2005;5:13.

[12] Deeks JJ, Higgins J, Altman DG. Chapter 9: analysing data and undertaking meta-analyses. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration; 2011. Available at www.cochrane-handbook.org. 2011.

[13] Bland M. Do baseline p-values follow a uniform distribution in randomised trials? PLoS One 2013;8(10):e76010.

[14] Carlisle JB. The analysis of 168 randomised controlled trials to test data integrity. Anaesthesia 2012;67(5):521−37.

[15] Anothaisintawee T, Attia J, Nickel JC, Thammakraisorn S, Numthavaj P, McEvoy M, et al. Management of chronic prostatitis/-chronic pelvic pain syndrome. JAMA 2011;305:78−86.

[16] Coombes BK, Bisset L, Vicenzino B. Efficacy and safety of corticosteroid injections and other injections for management of tendinopathy: a systematic review of randomised controlled trials. Lancet 2010;376:1751−67.

[17] Hemmingsen B, Christensen LL, Wetterslev J, Vaag A, Gluud C, Lund SS, et al. Comparison of metformin and insulin versus insulin alone for type 2 diabetes: systematic review of randomised clinical trials with meta-analyses and trial sequential analyses. BMJ 2012;344:e1771.

[18] Hempel S, Newberry SJ, Maher AR, Wang Z, Miles JNV, Shanman R, et al. Probiotics for the prevention and treatment of antibiotic-associated diarrhea: a systematic review and meta-analysis. JAMA 2012;307:1959−69.

[19] Heneghan C, Ward A, Perera R, Bankhead C, Fuller A, Stevens R, et al. Self-monitoring of oral anticoagulation: systematic review and meta-analysis of individual patient data. Lancet 2011;379:322−34.

[20] Leucht S, Tardy M, Komossa K, Heres S, Kissling W, Salanti G, et al. Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. Lancet 2012;379:2063−71.

[21] Neumann I, Rada G, Claro JC, Carrasco-Labra A, Thorlund K, Akl EA, et al. Oral direct factor Xa inhibitors versus low-molecular-weight heparin to prevent venous thromboembolism in patients undergoing total hip or knee replacement: a systematic review and meta-analysis. Ann Intern Med 2012;156:710−9.

[22] Orrow G, Kinmonth AL, Sanderson S, Sutton S. Effectiveness of physical activity promotion based in primary care: systematic review and meta-analysis of randomised controlled trials. BMJ 2012;344:e1389.

[23] Palmer SC, Lucia Di Micco M, Razavian M, Craig JC, Perkovic V, Pellegrini F, et al. Effects of antiplatelet therapy on mortality and cardiovascular and bleeding outcomes in persons with chronic kidney disease. Ann Intern Med 2012;156:445−59.

[24] Rutjes AWS, Jüni P, da Costa BR, Trelle S, Nüesch E, Reichenbach S. Viscosupplementation for osteoarthritis of the Knee: a systematic review and meta-analysis. Ann Intern Med 2012;157:180−91.

[25] Thangaratinam S, Rogozińska E, Jolly K, Glinkowski S, Roseboom T, Tomlinson J, et al. Effects of interventions in pregnancy on maternal weight and obstetric outcomes: meta-analysis of randomised evidence. BMJ 2012;344:e2088.

[26] Umpierre D, Ribeiro PAB, Kramer CK, Leitão CB, Zucatti ATN, Azevedo MJ, et al. Physical activity advice only or structured exercise training and association with HbA1c levels in type 2 diabetes: a systematic review and meta-analysis. JAMA 2011;305:1790−9.

[27] Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet 1997;350:326.

[28] Jüni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. Int J Epidemiol 2002;31:115−23.

[29] Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. JAMA 1996;276:637−9.

[30] Shiwa SR, Moseley AM, Maher CG, Pena Costa LO. Language of publication has a small influence on the quality of reports of controlled trials of physiotherapy interventions. J Clin Epidemiol 2013;66:78−84.

[31] Dechartres A, Charles P, Hopewell S, Ravaud P, Altman DG. Reviews assessing the quality or the reporting of randomized controlled trials are increasing over time but raised questions about how quality is assessed. J Clin Epidemiol 2011;64:136−44.

[32] Clark HD, Wells GA, Huët C, McAlister FA, Salmi LR, Fergusson D, et al. Assessing the quality of randomized trials: reliability of the Jadad scale. Control Clin Trials 1999;20:448−52.

Appendix A: Reference list of RCTs included in our review from each of the 12 Systematic Reviews.

Those with * are included in the narrative only, those with ** are included in our meta analysis also.

Hempel, S., S. J. Newberry, et al. (2012). "Probiotics for the Prevention and Treatment of Antibiotic-Associated Diarrhea A Systematic Review and Meta-analysis." JAMA: the journal of the American Medical Association 307(18): 1959-196

Adam J, Barret A, Barret-Bellet C. Essais cliniques controles en double insu de l'ultra-levure lyphilisee: etude multicentrique par 25 medicins de 388 cas. Gaz Med Fr. 1977;84:2072-2078.

Armuzzi A, Cremonini F, Bartolozzi F, et al. The effect of oral administration of Lactobacillus GG on antibiotic-associated gastrointestinal side-effects during Helicobacter pylori eradication therapy. Aliment Pharmacol Ther. Feb 2001;15(2):163-169.

Armuzzi A, Cremonini F, Ojetti V, et al. Effect of Lactobacillus GG supplementation on antibiotic-associated gastrointestinal side effects during Helicobacter pylori eradication therapy: a pilot study. Digestion. 2001;63(1):1-7.

*Arvola T, Laiho K, Torkkeli S, et al. Prophylactic Lactobacillus GG reduces antibiotic-associated diarrhea in children with respiratory infections: a randomized study. Pediatrics. Nov 1999;104(5):e64.

**Beausoleil M, Fortier N, Guenette S, et al. Effect of a fermented milk combining Lactobacillus acidophilus Cl1285 and Lactobacillus casei in the prevention of antibiotic-associated diarrhea: a randomized, double-blind, placebo-controlled trial. Can J Gastroenterol. Nov 2007;21(11):732-736.

Benhamou. Antibiotic-associated diarrhoea in children: A computer monitored double-blind outpatients trial comparing a protective and a probiotic agent. July 5 1999.

Bhalla A. Randomized placebo-controlled, double blind, multicentric trial on efficacy and safety of providac techsules (lactobacillus acidophilus LA-5 and bifidobacterium BB -12) for prevention of antibiotic-associated diarrhea in Indian patients. Journal of Clinical Pharmacology. 2011;51(9):1327.

**Borgia M, Sepe N, Brancato V. A controlled clinical study on Streptococcus faecium preparation for the prevention of side reactions during long termantibiotic treatments. . Curr Ther Res. 1982;31:265–271.

Bravo MV, Bunout D, Leiva L, et al. [Effect of probiotic Saccharomyces boulardii on prevention of antibiotic-associated diarrhea in adult outpatients with amoxicillin treatment]. Rev Med Chil. Aug 2008;136(8):981-988.

Can M, Besirbellioglu BA, Avci IY, Beker CM, Pahsa A. Prophylactic Saccharomyces boulardii in the prevention of antibiotic-associated diarrhea: a prospective study. Med Sci Monit. Apr 2006;12(4):PI19-22.

Cao YJ, Qu CM, Yuan Q. Control of intestinal flora alteration induced by eradication therapy of Helicobacter pylori infection in the elders. Chin J Gastroenterol Hepatol. 2005;14:195-199.

*Cimperman L, Bayless G, Best K, et al. A randomized, double-blind, placebo-controlled pilot study of Lactobacillus reuteri ATCC 55730 for the prevention of antibiotic-associated diarrhea in hospitalized adults. Journal of Clinical Gastroenterology. 2011;45(9):785-789.

Cindoruk M, Erkan G, Karakan T, Dursun A, Unal S. Efficacy and safety of Saccharomyces boulardii in the 14-day triple anti-Helicobacter pylori therapy: a prospective randomized placebo-controlled double-blind study. Helicobacter. Aug 2007;12(4):309-316.

Contardi I. [Oral bacterial therapy in prevention of antibiotic-induced diarrhea in childhood]. Clin Ter. Mar 31 1991;136(6):409-413.

**Conway S, Hart A, Clark A, Harvey I. Does eating yogurt prevent antibiotic-associated diarrhoea? A placebo-controlled randomised controlled trial in general practice. Br J Gen Pract. Dec 2007;57(545):953-959.

**Correa NB, Peret Filho LA, Penna FJ, Lima FM, Nicoli JR. A randomized formula controlled trial of Bifidobacterium lactis and Streptococcus thermophilus for prevention of antibiotic-associated diarrhea in infants. J Clin Gastroenterol. May-Jun 2005;39(5):385-389.

Cremonini F, Di Caro S, Covino M, et al. Effect of different probiotic preparations on anti-helicobacter pylori therapy-related side effects: a parallel group, triple blind, placebo-controlled study. Am J Gastroenterol. Nov 2002;97(11):2744-2749.

**de Bortoli N, Leonardi G, Ciancia E, et al. Helicobacter pylori eradication: a randomized prospective study of triple therapy versus triple therapy plus lactoferrin and probiotics. Am J Gastroenterol. May 2007;102(5):951-956.

de Vrese M, Kristen H, Rautenberg P, Laue C, Schrezenmeir J. Probiotic lactobacilli and bifidobacteria in a fermented milk product with added fruit preparation reduce antibiotic associated diarrhea and Helicob acter pylori activity. J Dairy Res. Aug 26 2011;4:396-403.

**Duman DG, Bor S, Ozutemiz O, et al. Efficacy and safety of Saccharomyces boulardii in prevention of antibiotic-associated diarrhoea due to Helicobacterpylori eradication. Eur J Gastroenterol Hepatol. Dec 2005;17(12):1357-1361.

Engelbrektson A, Korzenik JR, Pittler A, et al. Probiotics to minimize the disruption of faecal microbiota in healthy subjects undergoing antibiotic therapy. J Med Microbiol. May 2009;58(Pt 5):663-670.

*Erdeve O, Tiras U, Dallar Y. The probiotic effect of Saccharomyce boulardii in a pediatric age group. Journal of Tropical Pediatrics. 2004;50(4):234-236.

**Felley CP, Corthesy-Theulaz I, Rivero JL, et al. Favourable effect of an acidified milk (LC-1) on Helicobacter pylori gastritis in man. Eur J Gastroenterol Hepatol. Jan 2001;13(1):25-29.

*Frigerio G. A lactic acid produce enterococcus in the prevention of antibiotic-associated diarrhea and in the treatment of acute diarrheal disorders. A double blind multicenter placebo-controlled clinical study. Dig Dis Sci 1986;31:496S.

Gao XW, Mubasher M, Fang CY, Reifer C, Miller LE. Dose-response efficacy of a proprietary probiotic formula of Lactobacillus acidophilus CL1285 and Lactobacillus casei LBC80R for antibiotic-associated diarrhea and Clostridium difficile-associated diarrhea prophylaxis in adult patients. Am J Gastroenterol. Jul 2010;105(7):1636-1641.

Giovannone M, Barberani F, Boschetto S, al. E. Lactobacillus casei Dg effectiveness on Helicobacter pylori eradication treatment side effects; A placebo-controlled, double-blind randomized pilot study. Gastroenterology. 2007;132(4 Suppl 2):A614.

Gotz V, Romankiewicz JA, Moss J, Murray HW. Prophylaxis against ampicillin-associated diarrhea with a lactobacillus preparation. Am J Hosp Pharm. Jun 1979;36(6):754-757.

**Hickson M, D'Souza AL, Muthu N, et al. Use of probiotic Lactobacillus preparation to prevent diarrhoea associated with antibiotics: randomised double blind placebo controlled trial. BMJ. Jul 14 2007;335(7610):80.

Hurduc V, Plesca D, Dragomir D, Sajin M, Vandenplas Y. A randomized, open trial evaluating the effect of Saccharomyces boulardii on the eradication rate of Helicobacter pylori infection in children. Acta Paediatr. Jan 2009;98(1):127-131.

**Jirapinyo P, Densupsoontorn N, Thamonsiri N, Wongarn R. Prevention of antibiotic-associated diarrhea in infants by probiotics. J Med Assoc Thai. Aug 2002;85 Suppl 2:S739-742.

Kim MN, Kim N, Lee SH, et al. The effects of probiotics on PPI-triple therapy for Helicobacter pylori eradication. Helicobacter. Aug 2008;13(4):261-268.

Klarin B, Wullt M, Palmquist I, Molin G, Larsson A, Jeppsson B. Lactobacillus plantarum 299v reduces colonisation of Clostridium difficile in critically ill patients treated with antibiotics. Acta Anaesthesiol Scand. Sep 2008;52(8):1096-1102.

Klarin B, Johansson ML, Molin G, Larsson A, Jeppsson B. Adhesion of the probiotic bacterium Lactobacillus plantarum 299v onto the gut mucosa in critically ill patients: a randomised open trial. Crit Care. Jun 2005;9(3):R285-293.

**Koning CJ, Jonkers DM, Stobberingh EE, Mulder L, Rombouts FM, Stockbrugger RW. The effect of a multispecies probiotic on the intestinal microbiota and bowel movements in healthy volunteers taking the antibiotic amoxycillin. Am J Gastroenterol. Jan 2008;103(1):178-189.

Koning CJ, Jonkers D, Smidt H, et al. The effect of a multispecies probiotic on the composition of the faecal microbiota and bowel habits in chronic obstructive pulmonary disease patients treated with antibiotics. Br J Nutr. May 2010;103(10):1452-1460.

**Kotowska M, Albrecht P, Szajewska H. Saccharomyces boulardii in the prevention of antibiotic-associated diarrhoea in children: a randomized double-blind placebo-controlled trial. Aliment Pharmacol Ther. Mar 1 2005;21(5):583-590.

La Rosa M, Bottaro G, Gulino N, et al. [Prevention of antibiotic-associated diarrhea with Lactobacillus sporogens and fructo-oligosaccharides in children. A multicentric double-blind vs placebo study]. Minerva Pediatr. Oct 2003;55(5):447-452.

**Lewis SJ, Potts LF, Barry RE. The lack of therapeutic effect of Saccharomyces boulardii in the prevention of antibiotic-related diarrhoea in elderly patients. J Infect. Mar 1998;36(2):171-174.

Li D, Wang H, Tan M, Shao Y. Use of probiotics for prevention of antibiotic-associated diarrhea in elderly patients. Chinese Journal of Gastroenterology. 2010;15(3):154-156.

*Lighthouse J, Naito Y, Helmy A, et al. Endotoxinemia and benzodiazepine-like substances in compensated cirrhotic patients: A randomized study comparing the effect of rifaximine

alone and in association with a symbiotic preparation. Hepatology Research. March 1 2004;28(3):155-160.

Ligny G. Le traitment par le pereterol des troubles intestinaux secondaires a l'antibiotherapie - etude en double aveugle et etude clinique simple. Ars medici. 1976;31:989-995.

Lionetti E, Miniello VL, Castellaneta SP, et al. Lactobacillus reuteri therapy to reduce side-effects during anti-Helicobacter pylori treatment in children: a randomized placebo controlled trial. Aliment Pharmacol Ther. Nov 15 2006;24(10):1461-1468.

*Lonnermark E, Friman V, Lappas G, Sandberg T, Berggren A, Adlerberth I. Intake of Lactobacillus plantarum reduces certain gastrointestinal symptoms during treatment with antibiotics. J Clin Gastroenterol. Feb 2010;44(2):106-112.

Martinez RC, Franceschini SA, Patta MC, et al. Improved treatment of vulvovaginal candidiasis with fluconazole plus probiotic Lactobacillus rhamnosus GR-1 and Lactobacillus reuteri RC-14. Lett Appl Microbiol. Mar 2009;48(3):269-274.

**McFarland LV, Surawicz CM, Greenberg RN, et al. Prevention of beta-lactam-associated diarrhea by Saccharomyces boulardii compared with placebo. Am J Gastroenterol. Mar 1995;90(3):439-448.

**Merenstein DJ, Foster J, D'Amico F. A randomized clinical trial measuring the influence of kefir on antibiotic-associated diarrhea: the measuring the influence of Kefir (MILK) Study. Arch Pediatr Adolesc Med. Aug 2009;163(8):750-754.

Monteiro E, Fernandes JP, Vieira MR, et al. [Double blind clinical trial on the use of ultra-levure in the prophylaxis of antibiotic induced gastro-intestinal and mucocutaneous disorders] Acta Med Port 1981;3:143-145.

*Myllyluoma E, Veijola L, Ahlroos T, et al. Probiotic supplementation improves tolerance to Helicobacter pylori eradication therapy--a placebo-controlled, double-blind randomized pilot study. Aliment Pharmacol Ther. May 15 2005;21(10):1263-1272.

Nista EC, Candelli M, Cremonini F, et al. Bacillus clausii therapy to reduce side-effects of anti-Helicobacter pylori treatment: randomized, double-blind, placebo controlled trial. Aliment Pharmacol Ther. Nov 15 2004;20(10):1181-1188.

Ozkinay E, Terek MC, Yayci M, Kaiser R, Grob P, Tuncay G. The effectiveness of live lactobacilli in combination with low dose oestriol (Gynoflor) to restore the vaginal flora after treatment of vaginal infections. BJOG. Feb 2005;112(2):234-240.

**Park SK, Park DI, Choi JS, et al. The effect of probiotics on Helicobacter pylori eradication. Hepatogastroenterology. Oct-Nov 2007;54(79):2032-2036.

Plummer S, Weaver MA, Harris JC, Dee P, Hunter J. Clostridium difficile pilot study: effects of probiotic supplementation on the incidence of C. difficile diarrhoea. Int Microbiol. Mar 2004;7(1):59-62.

*Reid G, Bruce AW, Taylor M. Influence of three-day antimicrobial therapy and lactobacillus vaginal suppositories on recurrence of urinary tract infections. Clin Ther. Jan-Feb 1992;14(1):11-16.

**Ruszczynski M, Radzikowski A, Szajewska H. Clinical trial: effectiveness of Lactobacillus rhamnosus (strains E/N, Oxy and Pen) in the prevention of antibiotic-associated diarrhoea in children. Aliment Pharmacol Ther. Jul 2008;28(1):154-161.

**Safdar N, Barigala R, Said A, McKinley L. Feasibility and tolerability of probiotics for prevention of antibiotic-associated diarrhoea in hospitalized US military veterans. J Clin Pharm Ther. Dec 2008;33(6):663-668.

Sahagun-Flores J, Lopez-Pena L, de la Cruz-Ramirez Jaimes J, al. E. Eradication fo Helicobacter pylori: triple treatment scheme plus Lactobacillus vs. triple treatment alone. Cir Cir 2007;75(5):333-336.

**Sampalis J. et al. Efficacy of BIO K+ CL1285(registered trademark) in the reduction of antibiotic-associated diarrhea - A placebo controlled double-blind randomized, multi-center study. Archives of Medical Science. 2010;6(1):56-64.

Saneeyan H, Layegh S, Rahimi H. Effectivness of probiotic on treatment of Helicobacter pylori infection in children. Journal of Isfahan Medical School. 2011;29(146):882-889.

**Schrezenmeir J, Heller K, McCue M, et al. Benefits of oral supplementation with and without synbiotics in young children with acute bacterial infections. Clin Pediatr (Phila). Apr 2004;43(3):239-249.

*Selinger C, Lockett M, Bell A, Sebastian S, Haslam N. VSL#3 for the prevention of antibiotic associated diarrhoea (AAD) and clostridium difficile associated diarrhoea (CDAD): An interim analysis. Gut. 2011;60:A4.

*Sheu BS, Wu JJ, Lo CY, et al. Impact of supplement with Lactobacillus- and Bifidobacterium-containing yogurt on triple therapy for Helicobacter pylori eradication. Aliment Pharmacol Ther. Sep 2002;16(9):1669-1675.

Siitonen S, Vapaatalo H, Salminen S, et al. Effect of Lactobacillus GG yoghurt in prevention of antibiotic associated diarrhoea. Ann Med. Feb 1990;22(1):57-59.

Simakachorn N, Bibiloni R, Yimyaem P, et al. Tolerance, safety, and effect on the faecal microbiota of an enteral formula supplemented with pre- and probiotics in critically ill children. J Pediatr Gastroenterol Nutr. Aug 2011;53(2):174-181.

**Song MJ, Park DI, Park JH, et al. The effect of probiotics and mucoprotective agents on PPI-based triple therapy for eradication of Helicobacter pylori. Helicobacter. Jun 2010;15(3):206-213.

**Song HJ, Kim JY, Jung SA, et al. Effect of probiotic Lactobacillus (Lacidofil® cap) for the prevention of antibiotic-associated diarrhea: a prospective, randomized, double-blind, multicenter study. Journal of Korean medical science. 2010(12):1784-1791.

*Stein GY, Nanim R, Karniel E, Moskowitz I, Zeidman A. [Probiotics as prophylactic agents against antibiotic-associated diarrhea in hospitalized patients]. Harefuah. Jul 2007;146(7):520-522, 575.

*Sullivan A, Barkholt L, Nord CE. Lactobacillus acidophilus, Bifidobacterium lactis and Lactobacillus F19 prevent antibiotic-associated ecological disturbances of Bacteroides fragilis in the intestine. J Antimicrob Chemother. Aug 2003;52(2):308-311.

Surawicz CM, Elmer GW, Speelman P, McFarland LV, Chinn J, van Belle G. Prevention of antibiotic-associated diarrhea by Saccharomyces boulardii: a prospective study. Gastroenterology. Apr 1989;96(4):981-988.

Sykora J, Valeckova K, Amlerova J, et al. Effects of a specially designed fermented milk product containing probiotic Lactobacillus casei DN-114 001 and the eradication of H. pylori in children: a prospective randomized double-blind study. J Clin Gastroenterol. Sep 2005;39(8):692-698.

**Szajewska H, Albrecht P, Topczewska-Cabanek A. Randomized, double-blind, placebo-controlled trial: effect of lactobacillus GG supplementation on Helicobacter pylori eradication rates and side effects during treatment in children. J Pediatr Gastroenterol Nutr. Apr 2009;48(4):431-436.

**Szymanski H, Armanska M, Kowalska-Duplaga K, Szajewska H. Bifidobacterium longum PL03, Lactobacillus rhamnosus KL53A, and Lactobacillus plantarum PL02 in the prevention of antibiotic-associated diarrhea in children: a randomized controlled pilot trial. Digestion. 2008;78(1):13-17.

Tankanow RM, Ross MB, Ertel IJ, Dickinson DG, McCormick LS, Garfinkel JF. A double-blind, placebo-controlled study of the efficacy of Lactinex in the prophylaxis of amoxicillin-induced diarrhea. DICP. Apr 1990;24(4):382-384.

**Thomas MR, Litin SC, Osmon DR, Corr AP, Weaver AL, Lohse CM. Lack of effect of Lactobacillus GG on antibiotic-associated diarrhea: a randomized, placebo-controlled trial. Mayo Clinic Proceedings. Sep 2001;76(9):883-889.

Tursi A, Brandimarte G, Giorgetti GM, Modeo ME. Effect of Lactobacillus casei supplementation on the effectiveness and tolerability of a new second-line 10-day quadruple therapy after failure of a first attempt to cure Helicobacter pylori infection. Med Sci Monit. Dec 2004;10(12):CR662-666.

*Vanderhoof JA, Whitney DB, Antonson DL, Hanner TL, Lupo JV, Young RJ. Lactobacillus GG in the prevention of antibiotic-associated diarrhea in children. J Pediatr. Nov 1999;135(5):564-568.

**Wenus C, Goll R, Loken EB, Biong AS, Halvorsen DS, Florholmen J. Prevention of antibiotic-associated diarrhoea by a fermented probiotic milk drink. Eur J Clin Nutr. Feb 2008;62(2):299-301.

Witsell DL, Garrett CG, Yarbrough WG, Dorrestein SP, Drake AF, Weissler MC. Effect of Lactobacillus acidophilus on antibiotic-associated gastrointestinal morbidity: a prospective randomized trial. J Otolaryngol. Aug 1995;24(4):230-233.

*Wunderlich PF, Braun L, Fumagalli I, et al. Double-blind report on the efficacy of lactic acid-producing Enterococcus SF68 in the prevention of antibiotic-associated diarrhoea and in the treatment of acute diarrhoea. J Int Med Res. Jul-Aug 1989;17(4):333-338.

Wisniewski JA, Li XM. Alternative and Complementary Treatment for Food Allergy. Immunology and Allergy Clinics of North America. 2012;32(1):135-150.

**Yoon H, Kim N, Kim JY, et al. Effects of multistrain probiotic-containing yogurt on second-line triple therapy for Helicobacter pylori infection. Journal of Gastroenterology and Hepatology. 2011;26(1):44-48.

Zoppi G, Cinquetti M, Benini A, Bonamini E, Minelli EB. Modulation of the intestinal ecosystem by probiotics and lactulose in children during treatment with ceftriaxone. Current Therapeutic Research - Clinical and Experimental June 30 2001;62(5):418-435.

Palmer, S. C., M. Lucia Di Micco, et al. (2012). "Effects of Antiplatelet Therapy on Mortality and Cardiovascular and Bleeding Outcomes in Persons With Chronic Kidney Disease." Ann Intern Med 156: 445-459.

*Jardine MJ, Ninomiya T, Perkovic V, Cass A, Turnbull F, Gallagher MP. et al. Aspirin is beneficial in hypertensive patients with chronic kidney disease: a post-hoc subgroup analysis of a randomized controlled trial. J Am Coll Cardiol. 2010;56.956-65

Use of a monoclonal antibody directed against the platelet glycoprotein IIb/IIIa receptor in high-risk coronary angioplasty. The EPIC Investigation. N Engl J Med. 1994;330:956-61. [PMID: 8121459] Use of a monoclonal antibody directed against the platelet glycoprotein IIb/IIIa receptor in high-risk coronary angioplasty. The EPIC Investigation. N Engl J Med. 1994;330.956-61

EPISTENT Investigators. Randomised placebo-controlled and balloon-angioplasty-controlled trial to assess safety of coronary stenting with use of platelet glycoprotein-IIb/IIIa blockade. Lancet. 1998;352:87-92. [PMID: 9672272]EPISTENT Investigators. Randomised placebo-controlled and balloon-angioplasty-controlled trial to assess safety of coronary stenting with use of platelet glycoprotein-IIb/IIIa blockade. Lancet. 1998;352.87-92

**Brener SJ, Barr LA, Burchenal JE, Katz S, George BS, Jones AA. et al. Randomized, placebo-controlled trial of platelet glycoprotein IIb/IIIa blockade with primary angioplasty for acute myocardial infarction. ReoPro and Primary PTCA Organization and Randomized Trial (RAPPORT) Investigators. Circulation. 1998;98.734-41

Randomised placebo-controlled trial of effect of eptifibatide on complications of percutaneous coronary intervention: IMPACT-II. Integrilin to Minimise Platelet Aggregation and Coronary Thrombosis-II. Lancet. 1997;349:1422-8. [PMID: 9164315] Randomised placebo-controlled trial of effect of eptifibatide on complications of percutaneous coronary intervention: IMPACT-II. Integrilin to Minimise Platelet Aggregation and Coronary Thrombosis-II. Lancet. 1997;349.1422-8

Platelet glycoprotein IIb/IIIa receptor blockade and low-dose heparin during percutaneous coronary revascularization. The EPILOG Investigators. N Engl J Med. 1997;336:1689-96. [PMID: 9182212] Platelet glycoprotein IIb/IIIa receptor blockade and low-dose heparin during percutaneous coronary revascularization. The EPILOG Investigators. N Engl J Med.

1997;336.1689-96

Inhibition of platelet glycoprotein IIb/IIIa with eptifibatide in patients with acute coronary syndromes. The PURSUIT Trial Investigators. Platelet Glycoprotein IIb/IIIa in Unstable Angina: Receptor Suppression Using Integrilin Therapy. N Engl J Med. 1998;339:436-43. [PMID: 9705684] Inhibition of platelet glycoprotein IIb/IIIa with eptifibatide in patients with acute coronary syndromes. The PURSUIT Trial Investigators. Platelet Glycoprotein IIb/IIIa in Unstable Angina: Receptor Suppression Using Integrilin Therapy. N Engl J Med. 1998;339.436-43

**Januzzi JL Jr, Snapinn SM, DiBattiste PM, Jang IK, Theroux P. Benefits and safety of tirofiban among acute coronary syndrome patients with mild to moderate renal insufficiency: results from the Platelet Receptor Inhibition in Ischemic Syndrome Management in Patients Limited by Unstable Signs and Symptoms (PRISM-PLUS) trial. Circulation. 2002;105.2361-6

*Keltai M, Tonelli M, Mann JF, Sitkei E, Lewis BS, Hawken S, et al. CURE Trial Investigators. Renal function and outcomes in acute coronary syndrome: impact of clopidogrel. Eur J Cardiovasc Prev Rehabil. 2007;14.312-8

*Best PJ, Steinhubl SR, Berger PB, Dasgupta A, Brennan DM, Szczech LA, et al. CREDO Investigators. The efficacy and safety of short- and long-term dual antiplatelet therapy in patients with mild or moderate chronic kidney disease: results from the Clopidogrel for the Reduction of Events During Observation (CREDO) trial. Am Heart J. 2008;155.687-93

Donadio JV Jr, Anderson CF, Mitchell JC 3rd, Holley KE, Ilstrup DM, Fuster V. et al. Membranoproliferative glomerulonephritis. A prospective clinical trial of platelet-inhibitor therapy. N Engl J Med. 1984;310.1421-6

Aspirin effects on mortality and morbidity in patients with diabetes mellitus. Early Treatment Diabetic Retinopathy Study report 14. ETDRS Investigators. JAMA. 1992;268:1292-300. [PMID: 1507375] Aspirin effects on mortality and morbidity in patients with diabetes mellitus. Early Treatment Diabetic Retinopathy Study report 14. ETDRS Investigators. JAMA. 1992;268.1292-300

**Zäuner I, Böhler J, Braun N, Grupp C, Heering P, Schollmeyer P. Effect of aspirin and dipyridamole on proteinuria in idiopathic membranoproliferative glomerulonephritis: a

multicentre prospective clinical trial. Collaborative Glomerulonephritis Therapy Study Group (CGTS). Nephrol Dial Transplant. 1994;9.619-22

**Cheng IK, Fang GX, Wong MC, Ji YL, Chan KW, Yeung HW. A randomized prospective comparison of nadolol, captopril with or without ticlopidine on disease progression in IgA nephropathy. Nephrology. 1998;4.19-26

**Giustina A, Perini P, Desenzani P, Bossoni S, Ianniello P, Milani M. et al. Long-term treatment with the dual antithromboxane agent picotamide decreases microalbuminuria in normotensive type 2 diabetic patients. Diabetes. 1998;47.423-30

Khajehdehi P, Roozbeh J, Mostafavi H. A comparative randomized and placebo-controlled short-term trial of aspirin and dipyridamole for overt type-2 diabetic nephropathy. Scand J Urol Nephrol. 2002;36.145-8

*Dasgupta A, Steinhubl SR, Bhatt DL, Berger PB, Shao M, Mak KH, et al. CHARISMA Investigators. Clinical outcomes of patients with diabetic nephropathy randomized to clopidogrel plus aspirin versus aspirin alone (a post hoc analysis of the Clopidogrel for High Atherothrombotic Risk and Ischemic Stabilization, Management, and Avoidance [CHARISMA] trial). Am J Cardiol. 2009;103.1359-63

Kaegi A, Pineo GF, Shimizu A, Trivedi H, Hirsh J, Gent M. Arteriovenous-shunt thrombosis. Prevention by sulfinpyrazone. N Engl J Med. 1974;290.304-6

Michie DD, Wombolt DG. Use of sulfinpyrazone to prevent thrombus formation in arteriovenous fistulas and bovine grafts of patients on chronic hemodialysis. Cur Ther Res. 1977;22.196-204

Harter HR, Burch JW, Majerus PW, Stanford N, Delmez JA, Anderson CB. et al. Prevention of thrombosis in patients on hemodialysis by low-dose aspirin. N Engl J Med. 1979;301.577-9

Kobayashi K, Maeda K, Koshikawa S, Kawaguchi Y, Shimizu N, Naito C. Antithrombotic therapy with ticlopidine in chronic renal failure patients on maintenance hemodialysis—a multicenter collaborative double blind study. Thromb Res. 1980;20.255-61

Ell S, Mihindukulasuriya JC, O'Brien JR, Polak A, Vernham G. Ticlopidine in the prevention of blockage of fistulae and shunts. Haemostasis. 1982;12.180

Creek R. Ticlopidine—patency of haemodialysis access sites [Internal Report]. Guildford Sanofi Winthrop1990.

Middleton DA, Deichsel G. The prophylaxis of thrombosis in new arteriovenous dialysis shunts in the arm by low-dose acetylsalicyclic acid and dipyridamole [Internal Report]. Berlin, Germany Boehringer Ingelheim1992.

*Kooistra MP, van Es A, Marx JJ, Hertsig ML, Struyvenberg A. Low-dose aspirin does not prevent thrombovascular accidents in low-risk haemodialysis patients during treatment with recombinant human erythropoietin. Nephrol Dial Transplant. 1994;9.1115-20

**Sreedhara R, Himmelfarb J, Lazarus JM, Hakim RM. Anti-platelet therapy in graft thrombosis: results of a prospective, randomized, double-blind study. Kidney Int. 1994;45.1477-83

Mileti M, De PG, Bacchi M, Ogliari V, Pecchini F, Bufano G. et al. A trial to evaluate the efficacy of picotamide in preventing thrombotic occlusion of the vascular access in hemodialysis patients. J Nephrol. 1995;8.167-72

**Kaufman JS, O'Connor TZ, Zhang JH, Cronin RE, Fiore LD, Ganz MB, et al. Veterans Affairs Cooperative Study Group on Hemodialysis Access Graft Thrombosis. Randomized controlled trial of clopidogrel plus aspirin to prevent hemodialysis access graft thrombosis. J Am Soc Nephrol. 2003;14.2313-21

**Abdul-Rahman IS, AlHowaish AK. Warfarin versus aspirin in preventing tunneled hemodialysis catheter thrombosis: a prospective randomized study. Hong Kong Journal of Nephrology. 2007;9.23-30

**Dember LM, Beck GJ, Allon M, Delmez JA, Dixon BS, Greenberg A, et al. Dialysis Access Consortium Study Group. Effect of clopidogrel on early failure of arteriovenous fistulas for hemodialysis: a randomized controlled trial. JAMA. 2008;299.2164-71

**Dixon BS, Beck GJ, Vazquez MA, Greenberg A, Delmez JA, Allon M, et al. DAC Study Group. Effect of dipyridamole plus aspirin on hemodialysis graft patency. N Engl J Med.

2009;360.2191-201

**Ghorbani A, Aalamshah M, Shahbazian H, Ehsanpour A, Aref A. Randomized controlled trial of clopidogrel to prevent primary arteriovenous fistula failure in hemodialysis patients. Indian J Nephrol. 2009;19.57-61

Anderson M, Dewar P, Fleming LB, Hacking PM, Morley AR, Murray S. et al. A controlled trial of dipyridamole in human renal transplantation and an assessment of platelet function studies in rejection. Clin Nephrol. 1974;2.93-9

Schulze R, Langkopf B, Sziegoleit W. [The effect of dipyridamole on the results of allogenic kidney transplantation]. Z Urol Nephrol. 1990;83.255-9

Quarto Di Palo F, Elli A, Rivolta R, Parenti M, Palazzi P, Zanussi C. Prevention of chronic cyclosporine nephrotoxicity in renal transplantation by picotamide. Transplant Proc. 1991;23.969-71

Baigent C, Landray M, Leaper C, Altmann P, Armitage J, Baxter A. et al. First United Kingdom Heart and Renal Protection (UK-HARP-I) study: biochemical efficacy and safety of simvastatin and safety of low-dose aspirin in chronic kidney disease. Am J Kidney Dis. 2005;45.473-84

Orrow, G., A. L. Kinmonth, et al. (2012). "Effectiveness of physical activity promotion based in primary care: systematic review and meta-analysis of randomised controlled trials." BMJ: British Medical Journal 344.

*Harland J, White M, Drinkwater C, Chinn D, Farr L, Howel D. The Newcastle exercise project: a randomised controlled trial of methods to promote physical activity in primary care. BMJ1999;319:828-32.

**Chambers R, Chambers C, Campbell I. Exercise promotion for patients with significant medical problems. Health Educ J2000;59:90-8.

**The Writing Group for the Activity Counseling Trial Research Group. Effects of physical activity counseling in primary care: the activity counseling trial: a randomized controlled trial. JAMA2001;286:677-87.

**Elley CR, Kerse N, Arroll B, Robinson E. Effectiveness of counselling patients on physical activity in general practice: cluster randomised controlled trial. BMJ2003;326:793.

*Hillsdon M, Thorogood M, White I, Foster C. Advising people to take more exercise is ineffective: a randomized controlled trial of physical activity promotion in primary care. Int J Epidemiol2002;31:808-15.

**Petrella RJ, Koval JJ, Cunningham DA, Paterson DH. Can primary care doctors prescribe exercise to improve fitness?: The step test exercise prescription (STEP) project. Am J Prev Med2003;24:316-22.

*Kinmonth AL, Wareham NJ, Hardeman W, Sutton S, Prevost AT, Fanshawe T, et al. Efficacy of a theory-based behavioural intervention to increase physical activity in an at-risk group in primary care (ProActive UK): a randomised trial. Lancet2008;371:41-8.

**Halbert JA, Silagy CA, Finucane PM, Withers RT, Hamdorf PA. Physical activity and cardiovascular risk factors: effect of advice from an exercise specialist in Australian general practice. Med J Aust2000;173:84-7.

*Lamb SE, Bartlett HP, Ashley A, Bird W. Can lay-led walking programmes increase physical activity in middle aged adults? A randomised controlled trial. J Epidemiol Community Health2002;56:246-52.

**Van Sluijs EM, van Poppel MN, Twisk JW, Chin APM, Calfas KJ, van MW. Effect of a tailored physical activity intervention delivered in general practice settings: results of a randomized controlled trial. Am J Public Health2005;9510:1825-31.

**Jimmy G, Martin BW. Implementation and effectiveness of a primary care based physical activity counselling scheme. Patient Educ Couns2005;56:323-31.

*Harrison RA, Roberts C, Elton PJ. Does primary care referral to an exercise programme increase physical activity one year later? A randomized controlled trial. J Public Health2004;271:25-32.

**Lawton BA, Rose SB, Elley CR, Dowell AC, Fenton A, Moyes SA. Exercise on prescription for women aged 40-74 recruited through primary care: two year randomised controlled trial. BMJ2008;337:a2509.

**Kolt GS, Schofield GM, Kerse N, Garrett N, Oliver M. Effect of telephone counseling on physical activity for low-active older people in primary care: a randomized, controlled trial. J Am Geriatr Soc 2007;55:986-92.

**Morey MC, Peterson MJ, Pieper CF, Sloane R, Crowley GM, Cowper PA, et al. The Veterans Learning to Improve Fitness and Function in Elders Study: a randomized trial of primary care-based physical activity counseling for older men. J Am Geriatr Soc2009;57:1166-74.

Umpierre, D., P. A. B. Ribeiro, et al. (2011). "Physical Activity Advice Only or Structured Exercise Training and Association With HbA1c Levels in Type 2 Diabetes A Systematic Review and Meta-analysis." JAMA: the journal of the American Medical Association 305(17): 1790-1799.

**Aas AM, Bergstad I, Thorsby PM, Johannesen O, Solberg M, Birkeland KI. An intensified lifestyle intervention programme may be superior to insulin treatment in poorly controlled type 2 diabetic patients on oral hypoglycaemic agents: results of a feasibility study. *Diabet Med*. 2005;22(3):316-322

**Agurs-Collins TD, Kumanyika SK, Ten Have TR, Adams-Campbell LL. A randomized controlled trial of weight reduction and exercise for diabetes management in older African-American subjects. *Diabetes Care*. 1997;20(10):1503-1511

**Christian JG, Bessesen DH, Byers TE, Christian KK, Goldstein MG, Bock BC. Clinic-based support to help overweight patients with type 2 diabetes increase physical activity and lose weight. *Arch Intern Med*. 2008;168(2):141-146

Dasgupta K, Grover SA, Da Costa D, Lowensteyn I, Yale JF, Rahme E. Impact of modified glucose target and exercise interventions on vascular risk factors. *Diabetes Res Clin Pract*. 2006;72(1):53-60

Di Loreto C, Fanelli C, Lucidi P, et al. Validation of a counseling strategy to promote the adoption and the maintenance of physical activity by type 2 diabetic subjects. *Diabetes Care*. 2003;26(2):404-408

Hordern MD, Coombes JS, Cooney LM, Jeffriess L, Prins JB, Marwick TH. Effects of exercise intervention on myocardial function in type 2 diabetes. *Heart*. 2009;95(16):1343-1349

Kim SH, Lee SJ, Kang ES, et al. Effects of lifestyle modification on metabolic parameters and carotid intima-media thickness in patients with type 2 diabetes mellitus. *Metabolism*. 2006;55(8):1053-1059

**Jakicic JM, Jaramillo SA, Balasubramanyam A, et al; Look AHEAD Study Group. Effect of a lifestyle intervention on change in cardiorespiratory fitness in adults with type 2 diabetes: results from the Look AHEAD Study. *Int J Obes (Lond)*. 2009;33(3):305-316

**Mayer-Davis EJ, D'Antonio AM, Smith SM, et al. Pounds off with empowerment (POWER): a clinical trial of weight management strategies for black and white adults with diabetes who live in medically underserved rural communities. *Am J Public Health*. 2004;94(10):1736-1742

**Ménard J, Payette H, Baillargeon JP, et al. Efficacy of intensive multitherapy for patients with type 2 diabetes mellitus: a randomized controlled trial. *CMAJ*. 2005;173(12):1457-1466

**Vanninen E, Uusitupa M, Siitonen O, Laitinen J, Länsimies E. Habitual physical activity, aerobic capacity and metabolic control in patients with newly-diagnosed type 2 (non-insulin-dependent) diabetes mellitus: effect of 1-year diet and exercise intervention. *Diabetologia*. 1992;35(4):340-346

**Wing RR, Epstein LH, Paternostro-Bayles M, Kriska A, Nowalk MP, Gooding W. Exercise in a behavioural weight control programme for obese patients with type 2 (non-insulin-dependent) diabetes. *Diabetologia*. 1988;31(12):902-909

**Brun JF, Bordenave S, Mercier J, Jaussent A, Picot MC, Préfaut C. Cost-sparing effect of twice-weekly targeted endurance training in type 2 diabetics: a one-year controlled randomized trial. *Diabetes Metab*. 2008;34(3):258-265

**Cheung NW, Cinnadaio N, Russo M, Marek S. A pilot randomised controlled trial of resistance exercise bands in the management of sedentary subjects with type 2 diabetes. *Diabetes Res Clin Pract*. 2009;83(3):e68-e71

**Diedrich A, Munroe DJ, Romano M. Promoting physical activity for persons with diabetes. *Diabetes Educ*. 2010;36(1):132-140

*Kim CJ, Kang DH. Utility of a Web-based intervention for individuals with type 2 diabetes: the impact on physical activity levels and glycemic control. *Comput Inform Nurs*. 2006;24(6):337-345

*Kirk A, Mutrie N, MacIntyre P, Fisher M. Increasing physical activity in people with type 2 diabetes. *Diabetes Care*. 2003;26(4):1186-1192

*Kirk A, Barnett J, Leese G, Mutrie N. A randomized trial investigating the 12-month changes in physical activity and health outcomes following a physical activity consultation delivered by a person or in written form in type 2 diabetes: Time2Act. *Diabet Med*. 2009;26(3):293-301

*Krousel-Wood MA, Berger L, Jiang X, Blonde L, Myers L, Webber L. Does home-based exercise improve body mass index in patients with type 2 diabetes? results of a feasibility

trial. *Diabetes Res Clin Pract*. 2008;79(2):230-236

*Leehey DJ, Moinuddin I, Bast JP, et al. Aerobic exercise in obese diabetic patients with chronic kidney disease: a randomized and controlled pilot study. *Cardiovasc Diabetol*. 2009;862

Rönnemaa T, Mattila K, Lehtonen A, Kallio V. A controlled randomized study on the effect of long-term physical exercise on the metabolic control in type 2 diabetic patients. *Acta Med Scand*. 1986;220(3):219-224

**Samaras K, Ashwell S, Mackintosh AM, Fleury AC, Campbell LV, Chisholm DJ. Will older sedentary people with non-insulin-dependent diabetes mellitus start exercising? a health promotion model. *Diabetes Res Clin Pract*. 1997;37(2):121-128

**Tudor-Locke C, Bell RC, Myers AM, et al. Controlled outcome evaluation of the First Step Program: a daily physical activity intervention for individuals with type II diabetes. *Int J Obes Relat Metab Disord*. 2004;28(1):113-119

*van Rooijen AJ, Rheeder P, Eales CJ, Becker PJ. Effect of exercise versus relaxation on haemoglobin A_{1C} in Black females with type 2 diabetes mellitus. *QJM*. 2004;97(6):343-351

Bjørgaas M, Vik JT, Saeterhaug A, et al. Relationship between pedometer-registered activity, aerobic capacity and self-reported activity and fitness in patients with type 2 diabetes. *Diabetes Obes Metab*. 2005;7(6):737-744

**Cuff DJ, Meneilly GS, Martin A, Ignaszewski A, Tildesley HD, Frohlich JJ. Effective exercise modality to reduce insulin resistance in women with type 2 diabetes. *Diabetes Care*. 2003;26(11):2977-2982

Dela F, von Linstow ME, Mikines KJ, Galbo H. Physical training may enhance beta-cell function in type 2 diabetes. *Am J Physiol Endocrinol Metab*. 2004;287(5):E1024-E1031

Giannopoulou I, Fernhall B, Carhart R, et al. Effects of diet and/or exercise on the adipocytokine and inflammatory cytokine levels of postmenopausal women with type 2 diabetes. *Metabolism*. 2005;54(7):866-875

Goldhaber-Fiebert JD, Goldhaber-Fiebert SN, Tristán ML, Nathan DM. Randomized controlled community-based nutrition and exercise intervention improves glycemia and cardiovascular risk factors in type 2 diabetic patients in rural Costa Rica. *Diabetes Care*. 2003;26(1):24-29

**Kadoglou NP, Iliadis F, Angelopoulou N, et al. The anti-inflammatory effects of exercise training in patients with type 2 diabetes mellitus. *Eur J Cardiovasc Prev Rehabil*. 2007;14(6):837-843

**Kadoglou NP, Perrea D, Iliadis F, Angelopoulou N, Liapis C, Alevizos M. Exercise reduces resistin and inflammatory cytokines in patients with type 2 diabetes. *Diabetes Care*. 2007;30(3):719-721

Kadoglou NP, Iliadis F, Sailer N, et al. Exercise training ameliorates the effects of rosiglitazone on traditional and novel cardiovascular risk factors in patients with type 2 diabetes mellitus. *Metabolism*. 2010;59(4):599-607

**Lambers S, Van Laethem C, Van Acker K, Calders P. Influence of combined exercise training on indices of obesity, diabetes and cardiovascular risk in type 2 diabetes patients. *Clin Rehabil*. 2008;22(6):483-492

Ligtenberg PC, Hoekstra JB, Bol E, Zonderland ML, Erkelens DW. Effects of physical training on metabolic control in elderly type 2 diabetes mellitus patients. *Clin Sci (Lond)*. 1997;93(2):127-135

**Middlebrooke AR, Elston LM, Macleod KM, et al. Six months of aerobic exercise does not improve microvascular function in type 2 diabetes mellitus. *Diabetologia*. 2006;49(10):2263-2271

Raz I, Hauser E, Bursztyn M. Moderate exercise improves glucose metabolism in uncontrolled elderly patients with non-insulin-dependent diabetes mellitus. *Isr J Med Sci*. 1994;30(10):766-770

Ribeiro IC, Iborra RT, Neves MQ, et al. HDL atheroprotection by aerobic exercise training in type 2 diabetes mellitus. *Med Sci Sports Exerc*. 2008;40(5):779-786

**Sridhar B, Haleagrahara N, Bhat R, Kulur AB, Avabratha S, Adhikary P. Increase in the heart rate variability with deep breathing in diabetic patients after 12-month exercise training. *Tohoku J Exp Med*. 2010;220(2):107-113

Vancea DM, Vancea JN, Pires MI, Reis MA, Moura RB, Dib SA. Effect of frequency of physical exercise on glycemic control and body composition in type 2 diabetic patients. *Arq Bras Cardiol*. 2009;92(1):23-30

Verity LS, Ismail AH. Effects of exercise on cardiovascular disease risk in women with NIDDM. *Diabetes Res Clin Pract*. 1989;6(1):27-35

**Castaneda C, Layne JE, Munoz-Orians L, et al. A randomized controlled trial of resistance exercise training to improve glycemic control in older adults with type 2 diabetes. *Diabetes Care*. 2002;25(12):2335-2341

Dunstan DW, Daly RM, Owen N, et al. High-intensity resistance training improves glycemic control in older patients with type 2 diabetes. *Diabetes Care*. 2002;25(10):1729-1736

Balducci S, Leonetti F, Di Mario U, Fallucca F. Is a long-term aerobic plus resistance training program feasible for and effective on metabolic profiles in type 2 diabetic patients? *Diabetes Care*. 2004;27(3):841-842

**Loimaala A, Huikuri HV, Kööbi T, Rinne M, Nenonen A, Vuori I. Exercise training improves baroreflex sensitivity in type 2 diabetes. *Diabetes*. 2003;52(7):1837-1842

**Tessier D, Ménard J, Fülöp T, et al. Effects of aerobic physical exercise in the elderly with type 2 diabetes mellitus. *Arch Gerontol Geriatr*. 2000;31(2):121-132

**Sigal RJ, Kenny GP, Boulé NG, et al. Effects of aerobic training, resistance training, or both on glycemic control in type 2 diabetes: a randomized trial. *Ann Intern Med*. 2007;147(6):357-369

**Church TS, Blair SN, Cocreham S, et al. Effects of aerobic and resistance training on hemoglobin A_{1c} levels in patients with type 2 diabetes: a randomized controlled trial. *JAMA*. 2010;304(20):2253-2262

Coombes, B. K., L. Bisset, et al. (2010). "Efficacy and safety of corticosteroid injections and other injections for management of tendinopathy: a systematic review of randomised controlled trials." The Lancet 376(9754): 1751-1767.

**L Bisset, E Beller, G Jull, P Brooks, R Darnell, B Vicenzino
Mobilisation with movement and exercise, corticosteroid injection, or wait and see for tennis elbow: randomised trial BMJ, 333 (2006), p. 939

E Haker, T Lundeberg Elbow-band, splintage and steroids in lateral epicondylalgia (tennis elbow) Pain Clin, 6 (1993), pp. 103–112

*EM Hay, SM Paterson, M Lewis, G Hosie, P Croft Pragmatic randomised controlled trial of local corticosteroid injection and naproxen for treatment of lateral epicondylitis of elbow in primary care BMJ, 319 (1999), pp. 964–968

A Lindenhovius, M Henket, BP Gilligan, S Lozano-Calderon, JB Jupiter, D Ring
Injection of dexamethasone versus placebo for lateral elbow pain: a prospective, double-blind, randomized clinical trial J Hand Surg Am, 33 (2008), pp. 909–919

KL Newcomer, ER Laskowski, DM Idank, TJ McLean, KS Egan
Corticosteroid injection in early treatment of lateral epicondylitis
Clin J Sport Med, 11 (2001), pp. 214–222

G Okcu, H Yercan, U Ozic The comparison of single dose versus multi-dose local corticosteroid injections for tennis elbow Clin Res, 13 (2002), pp. 158–163

**JC Peerbooms, J Sluimer, DJ Bruijn, T Gosens Positive effect of an autologous platelet concentrate in lateral epicondylitis in a double-blind randomized controlled trial: platelet-rich plasma versus corticosteroid injection with a 1-year follow-up Am J Sports Med, 38 (2010), p. 255

R Price, H Sinclair, I Heinrich, T Gibson
Local injection treatment of tennis elbow—hydrocortisone, triamcinolone and lignocaine compared Br J Rheumatol, 30 (1991), pp. 39–44

T Saartok, E Eriksson Randomized trial of oral naproxen or local injection of betamethasone in lateral epicondylitis of the humerus Orthopedics, 9 (1986), pp. 191–194

**N Smidt, DA van der Windt, WJ Assendelft, WL Deville, IB Korthals-de Bos, LM Bouter Corticosteroid injections, physiotherapy, or a wait-and-see policy for lateral epicondylitis: a randomised controlled trial Lancet, 359 (2002), pp. 657–662

**JH Tonks, SK Pai, SR Murali Steroid injection therapy is the best conservative treatment for lateral epicondylitis: a prospective randomised controlled trial
Int J Clin Pract, 61 (2007), pp. 240–246

*Verhaar JA, Walenkamp GH, van Mameren H, Kester AD, van der Linden AJ. Local corticosteroid injection versus Cyriax-type physiotherapy for tennis elbow. J Bone Joint Surg Br199: 128–32.

AO Adebajo, P Nash, BL Hazleman A prospective double blind dummy placebo controlled study comparing triamcinolone hexacetonide injection with oral diclofenac 50 mg TDS in patients with rotator cuff tendinitis J Rheumatol, 17 (1990), pp. 1207–1210

K Akgün, M Birtane, U Akarimak Is local subacromial corticosteroid injection beneficial in subacromial impingement syndrome? Clin Rheumatol, 23 (2004), pp. 496–500

**CM Alvarez, R Litchfield, D Jackowski, S Griffin, A Kirkley
A prospective, double-blind, randomized clinical trial comparing subacromial injection of betamethasone and xylocaine to xylocaine alone in chronic rotator cuff tendinosis
Am J Sports Med, 33 (2005), pp. 255–262

**J Alvarez-Nemegyei, A Bassol-Perea, J Pasos
Efficacy of the local injection of methylprednisolone acetate in the subacromial impingement syndrome. A randomised, double blind trial Reumatol Clin, 4 (2008), pp. 49–54

**B Blair, AS Rokito, F Cuomo, K Jarolem, JD Zuckerman Efficacy of injections of corticosteroids for subacromial impingement syndrome J Bone Joint Surg Am, 78 (1996), pp. 1685–1689

*DJ Cloke, H Watson, S Purdy, IN Steen, JR Williams A pilot randomized, controlled trial of treatment for painful arc of the shoulderJ Shoulder Elbow Surg, 17 (Suppl) (2008), pp. S17–S21

**OM Ekeberg, E Bautz-Holter, EK Tveita, NG Juel, S Kvalheim, JI Brox
Subacromial ultrasound guided or systemic steroid injection for rotator cuff disease: randomised double blind study BMJ, 338 (2009), p. a3112

**EM Hay, E Thomas, SM Paterson, K Dziedzic, PR Croft A pragmatic randomised controlled trial of local corticosteroid injection and physiotherapy for the treatment of new episodes of unilateral shoulder pain in primary care Ann Rheum Dis, 62 (2003), pp. 394–399

S Karthikeyan, HT Kwong, PK Upadhyay, N Parsons, SJ Drew, D Griffin
A double-blind randomised controlled study comparing subacromial injection of tenoxicam or methylprednisolone in patients with subacromial impingement
J Bone Joint Surg Br, 92 (2010), pp. 77–82

**J McInerney, J Dias, S Durham, A Evan Randomised controlled trial of single, subacromial injection of methylprednisolone in patients with persistent, post-traumatic impingment of the shoulder Emerg Med J, 20 (2003), pp. 218–221

M Petri, R Dobrow, R Neiman, Q Whiting-O'Keefe, WE Seaman
Randomized, double-blind, placebo-controlled study of the treatment of the painful shoulder Arthritis Rheum, 30 (1987), pp. 1040–1045

**PC Vecchio, BL Hazleman, RH King A double-blind trial comparing subacromial methylprednisolone and lignocaine in acute rotator cuff tendinitis
Br J Rheumatol, 32 (1993), pp. 743–745

RH White, DM Paull, KW Fleming Rotator cuff tendinitis: comparison of subacromial injection of a long acting corticosteroid versus oral indomethacin therapy
J Rheumatol, 13 (1986), pp. 608–613

S Stahl, T Kaufman The efficacy of an injection of steroids for medial epicondylitis. A prospective study of sixty elbows J Bone Joint Surg Am, 79 (1997), pp. 1648–1652

G Capasso, V Testa, N Maffuli, G Bifulco Aprotinin, corticosteroids and normosaline in the management of patellar tendinopathy in athletes: a prospective randomised study Sports Exerc Injury, 3 (1997), pp. 111–115

U Fredberg, L Bolvig, M Pfeiffer-Jensen, D Clemmensen, BW Jakobsen, K Stengaard-Pedersen Ultrasonography as a tool for diagnosis, guidance of local steroid injection and, together with pressure algometry, monitoring of the treatment of athletes with chronic jumper's knee and Achilles tendinitis: a randomized, double-blind, placebo-controlled study Scand J Rheumatol, 33 (2004), pp. 94–101

C Akermark, H Crone, U Elsasser, B Forsskahl Glycosaminoglycan polysulfate injections in lateral humeral epicondylalgia: a placebo-controlled double-blind trial
Int J Sports Med, 16 (1995), pp. 196–200

**RJ Petrella, A Cogliano, J Decaria, N Mohamed, R Lee Management of tennis elbow with sodium hyaluronate periarticular injections Sports Med Arthrosc Rehabil Ther Technol, 2 (2010), p. 4

**M Scarpone, DP Rabago, A Zgierska, G Arbogast, E Snell The efficacy of prolotherapy for lateral epicondylosis: a pilot study Clin J Sport Med, 18 (2008), pp. 248–254

**SM Wong, AC Hui, PY Tong, DW Poon, E Yu, LK Wong Treatment of lateral epicondylitis with botulinum toxin: a randomized, double-blind, placebo-controlled trial Ann Intern Med, 143 (2005), pp. 793–797

**E Zeisig, M Fahlstrom, L Ohberg, H Alfredson Pain relief after intratendinous injections in patients with tennis elbow: results of a randomised study Br J Sports Med, 42 (2008), pp. 267–271

I Sengul, B Oz, O Yoleri, N Olmez, A Memis, E Uluc Sodium hyaluronate injections compared to local modalities for the treatment of shoulder impingement syndrome
Turk J Phys Med Rehab, 54 (2008), pp. 138–142

H Alfredson, L Ohberg Sclerosing injections to area s of neo-vascularisation reduce pain in chronic Achilles tendinopathy: a double-blind randomised controlled trial
Knee Surg Sports Traumatol Arthrosc, 13 (2005), pp. 338–344

*R Brown, J Orchard, M Kinchington, A Hooper, G Nalder Aprotinin in the management of Achilles tendinopathy: a randomised controlled trial
Br J Sports Med, 40 (2006), pp. 275–279

**RJ de Vos, A Weir, HT van Schie et al. Platelet-rich plasma injection for chronic Achilles tendinopathy: a randomized controlled trialJAMA, 303 (2010), pp. 144–149

**A Hoksrud, L Ohberg, H Alfredson, R Bahr Ultrasound-guided sclerosis of neovessels in painful chronic patellar tendinopathy: a randomized controlled trial
Am J Sports Med, 34 (2006), pp. 1738–1746

H Sundqvist, B Forsskahl, M Kvist A promising novel therapy for Achilles peritendinitis: double-blind comparison of glycosaminoglycan polysulfate and high-dose indomethacin Int J Sports Med, 8 (1987), pp. 298–303

*MJ Yelland, KR Sweeting, JA Lyftogt, SK Ng, PA Scuffham, KA Evans
Prolotherapy injections and eccentric loading exercises for painful Achilles tendinosis: a randomised trial Br J Sports Med (2009) http://dx.doi.org/10.1136/bjsm.2009.057968 published online June 22.

**Willberg, K Sunding, L Ohberg, M Forssblad, M Fahlstrom, H Alfredson
Sclerosing injections to treat midportion Achilles tendinosis: a randomised controlled study evaluating two different concentrations of Polidocanol
Knee Surg Sports Traumatol Arthrosc, 16 (2008), pp. 859–864

Neumann, I., G. Rada, et al. (2012). "Oral Direct Factor Xa Inhibitors Versus Low-Molecular-Weight Heparin to Prevent Venous Thromboembolism in Patients Undergoing Total Hip or Knee Replacement: A Systematic Review and Meta-analysis." Annals of internal medicine.

Weitz JI, Cao C, Eriksson BI, Fisher W, Kupfer S, Raskob G, et al.. A dose-finding study with TAK-442, an oral factor Xa inhibitor, in patients undergoing elective total knee replacement surgery. Thromb Haemost. 2010;104:1150-7.

**Lassen MR, Gallus A, Raskob GE, Pineo G, Chen D, Ramirez LM, ADVANCE-3 Investigators. Apixaban versus enoxaparin for thromboprophylaxis after hip replacement. N Engl J Med. 2010;363:2487-98.

Raskob G, Cohen AT, Eriksson BI, Puskas D, Shi M, Bocanegra T, et al.. Oral direct factor Xa inhibition with edoxaban for thromboprophylaxis after elective total hip replacement. A randomised double-blind dose-response study. Thromb Haemost. 2010;104:642-9.

**Eriksson BI, Turpie AG, Lassen MR, Prins MH, Agnelli G, Kälebo P, et al., ONYX-2 Study Group. Prevention of venous thromboembolism with an oral factor Xa inhibitor, YM150, after total hip arthroplasty. A dose finding study (ONYX-2). J Thromb Haemost. 2010;8:714-21.

**Lassen MR, Raskob GE, Gallus A, Pineo G, Chen D, Hornick P, ADVANCE-2 Investigators. Apixaban versus enoxaparin for thromboprophylaxis after knee replacement (ADVANCE-2): a randomised double-blind trial. Lancet. 2010;375:807-15.

**Lassen MR, Raskob GE, Gallus A, Pineo G, Chen D, Portman RJ. Apixaban or enoxaparin for thromboprophylaxis after knee replacement. N Engl J Med. 2009;361:594-604.

Turpie AG, Lassen MR, Davidson BL, Bauer KA, Gent M, Kwong LM, et al., RECORD4 Investigators. Rivaroxaban versus enoxaparin for thromboprophylaxis after total knee arthroplasty (RECORD4): a randomised trial. Lancet. 2009;373:1673-80.

Turpie AG, Bauer KA, Davidson BL, Fisher WD, Gent M, Huo MH, et al., EXPERT Study Group. A randomized evaluation of betrixaban, an oral factor Xa inhibitor, for prevention of thromboembolic events after total knee replacement (EXPERT). Thromb Haemost. 2009;101:68-76.

**Lassen MR, Ageno W, Borris LC, Lieberman JR, Rosencher N, Bandel TJ, et al., RECORD3 Investigators. Rivaroxaban versus enoxaparin for thromboprophylaxis after total knee arthroplasty. N Engl J Med. 2008;358:2776-86.

**Eriksson BI, Borris LC, Friedman RJ, Haas S, Huisman MV, Kakkar AK, et al., RECORD1 Study Group. Rivaroxaban versus enoxaparin for thromboprophylaxis after hip arthroplasty. N Engl J Med. 2008;358:2765-75.

Lassen MR, Davidson BL, Gallus A, Pineo G, Ansell J, Deitchman D. The efficacy and safety of apixaban, an oral, direct factor Xa inhibitor, as thromboprophylaxis in patients following total knee replacement. J Thromb Haemost. 2007;5:2368-75.

*Eriksson BI, Borris LC, Dahl OE, Haas S, Huisman MV, Kakkar AK, et al.. Dose-escalation study of rivaroxaban (BAY 59-7939)—an oral, direct factor Xa inhibitor—for the prevention of venous thromboembolism in patients undergoing total hip replacement. Thromb Res. 2007;120:685-93.

**Eriksson BI, Turpie AG, Lassen MR, Prins MH, Agnelli G, Kälebo P, et al., ONYX Study Group. A dose escalation study of YM150, an oral direct factor Xa inhibitor, in the prevention of venous thromboembolism in elective primary hip replacement surgery. J Thromb Haemost. 2007;5:1660-5.

Agnelli G, Haas S, Ginsberg JS, Krueger KA, Dmitrienko A, Brandt JT. A phase II study of the oral factor Xa inhibitor LY517717 for the prevention of venous thromboembolism after hip or knee replacement. J Thromb Haemost. 2007;5:746-53.

**Eriksson BI, Borris LC, Dahl OE, Haas S, Huisman MV, Kakkar AK, et al., ODIXa-HIP Study Investigators. A once-daily, oral, direct factor Xa inhibitor, rivaroxaban (BAY 59-7939), for thromboprophylaxis after total hip replacement. Circulation. 2006;114:2374-81.

**Turpie AG, Fisher WD, Bauer KA, Kwong LM, Irwin MW, Kälebo P, et al., OdiXa-Knee Study Group. BAY 59-7939: an oral, direct factor Xa inhibitor for the prevention of venous thromboembolism in patients after total knee replacement. A phase II dose-ranging study. J Thromb Haemost. 2005;3:2479-86.

Eriksson BI, Borris L, Dahl OE, Haas S, Huisman MV, Kakkar AK, et al., ODIXa-HIP Study Investigators. Oral, direct factor Xa inhibition with BAY 59-7939 for the prevention of venous thromboembolism after total hip replacement. J Thromb Haemost. 2006;4:121-8.

Silva Kanan P, Schwartsmann CR, Carbonera Boschin L, Conrad S, Faria Silva M. Comparative study between rivaroxaban and enoxaparin in deep venous thromboembolism prophylaxis in patients submitted to total hip arthroplasty. Revista Brasileira de Ortopedia. 2008;43:319-28.

Lassen MR, Davidson BL, Gallus A, Pineo A, Ansell J, Deitchman D. A phase II randomized, double-blind, five-arm, parallel-group, dose-response study of a new oral directly-acting factor Xa inhibitor, razaxaban, for the prevention of deep vein thrombosis in knee replacement surgery—on behalf of the razaxaban investigators [Abstract]. Blood. 2003;102: ((11 Pt 1)) 15a.

Fuji T, Wang CJ, Fujita S, Tachibana S, Kawai Y. Edoxaban in patients undergoing total hip arthroplasty: a phase IIb dose-finding study [Abstract]. Proceedings of the Annual Meeting of the American Society of Hematology, New Orleans, Louisiana, 5–8 December 2009. Abstract 2098.

Fuji T, Fujita S, Tachibana S, Kawai Y, Koretsune Y, Yamashita T, et al. Efficacy and safety of edoxaban versus enoxaparin for the prevention of venous thromboembolism following total hip arthroplasty: STARS J-V trial. Proceedings of the Annual Meeting of the American Society of Hematology, Orlando, Florida, 4–7 December 2010. Abstract 3320.

Fuji T, Wang C, Fujita S, Tachibana S, Kawai Y, Koretsune Y, et al. Edoxaban versus enoxaparin for thromboprophylaxis after total knee replacement: the STARS E-3 trial. Presented at 21st International Congress on Thrombosis, Milan, Italy, 6–9 July 2010.

Rutjes, A. W. S., P. Jüni, et al. (2012). "Viscosupplementation for Osteoarthritis of the KneeA Systematic Review and Meta-analysis." Annals of internal medicine.

Anika Therapeutics. Confidentially obtained report. Bedford, MA Anika Therapeutics2000.

Anika Therapeutics. Confidentially obtained report. Bedford, MA Anika Therapeutics2001.

Genzyme Biomaterials. Confidentially obtained report. Cambridge, MA Genzyme Biomaterials2005.

Personal communication (Rydell, 1972) described in: Peyron JG, Balazs EA. Preliminary clinical assessment of Na-hyaluronate injection into human arthritic joints. Pathol Biol (Paris). 1974;22:731-6.

Weiss C, Balazs EA, St Onge R, Denlinger JL. Clinical studies of the intra-articular injections of Healon (sodium hyaluronate) in the treatment of osteoarthritis of human knees. Semin Arthritis Rheum. 1981;11:Suppl 1143-4.

Shichikawa K, Igarashi M, Sugawara S, Iwasaki Y. Clinical evaluation of high molecular weight sodium hyaluronate (SPH) on osteoarthritis of the knee—a multi-center well controlled comparative study. Jpn J Clin Pharmacol Therapeut. 1983;14:545-58.

Shichikawa K, Maeda A, Ogawa N. [Clinical evaluation of sodium hyaluronate in the treatment of osteoarthritis of the knee]. Ryumachi. 1983;23:280-90.

*Bragantini A, Cassini M, de Bastiani G, Perbellini A. Controlled single-blind trial of intra-articularly injected hyaluronic acid (Hyalgan) in osteo-arthritis of the knee. Clin Trials J. 1987;24:333-40.

*Grecomoro G, Martorana U, Di Marco C. Intra-articular treatment with sodium hyaluronate in gonarthrosis: a controlled clinical trial versus placebo. Pharmatherapeutica. 1987;5:137-41.

*Dixon AS, Jacoby RK, Berry H, Hamilton EB. Clinical trial of intra-articular injection of sodium hyaluronate in patients with osteoarthritis of the knee. Curr Med Res Opin. 1988;11:205-13.

Ghirardini M, Betelemme L, Fatti F. Impiego intraarticolare di acido ialuronico estrattivo de orgoteina sia separatamente che in associazione in pazienti affetti da gonartrosi in fase sinovitica [Abstract]. Reumatismo. 1990;42:132.

*Russell IJ, Michalek JE, Lawrence VA, Lessard JA, Briggs BT, May GS. A randomized, placebo (PL) and no-intervention (NI) controlled, trial of intra-articular (IA) 1% sodium hyaluronate (HA) in the treatment of knee osteoarthritis (OA) [Abstract B94]. Arthritis Rheum. 1992;35:SupplS132.

**Dougados M, Nguyen M, Listrat V, Amor B. High molecular weight sodium hyaluronate (hyalectin) in osteoarthritis of the knee: a 1 year placebo-controlled trial. Osteoarthritis Cartilage. 1993;1:97-103.

Isdale AH, Hordon LD, Bird HA, Wright V. A controlled comparison of intra-articular Healon (hyaluronate), triamcinolone and saline in osteoarthritis of the knee [Abstract 118]. Br J Rheumatol. 1993;32:Suppl 161.

*Moreland LW, Arnold WJ, Saway A, Savory C, Sikes D. Efficacy and safety of intra-articular hylan G-F 20 (Synvisc), a viscoelastic derivative of hyaluronan, in patients with osteoarthritis of the knee [Abstract]. Arthritis Rheum. 1993;36:Suppl 9165.

*Pedersen PB. Intra-articular hyaluronic acid (HA) in the treatment of osteoarthritis (OA) of the knee [Abstract]. Osteoarthritis Cartilage. 1993;1:70.

**Puhl W, Bernau A, Greiling H, Köpcke W, Pförringer W, Steck KJ, et al.. Intra-articular sodium hyaluronate in osteoarthritis of the knee: a multicenter, double-blind study. Osteoarthritis Cartilage. 1993;1:233-41.

*Cohen MA, Shiroky JB, Ballachey ML, Neville C, Esdaile JM. Double-blind randomized trial of intra-articular (I/A) hyaluronate in the treatment of osteoarthritis of the knee [Abstract 62]. Arthritis Rheum. 1994;37:Suppl 6R31.

**Creamer P, Sharif M, George E, Meadows K, Cushnaghan J, Shinmei M, et al.. Intra-articular hyaluronic acid in osteoarthritis of the knee: an investigation into mechanisms of action. Osteoarthritis Cartilage. 1994;2:133-40.

*Henderson EB, Smith EC, Pegley F, Blake DR. Intra-articular injections of 750 kD hyaluronan in the treatment of osteoarthritis: a randomised single centre double-blind placebo-controlled trial of 91 patients demonstrating lack of efficacy. Ann Rheum Dis. 1994;53:529-34.

*Scale D, Wobig M, Wolpert W. Viscosupplementation of osteoarthritic knees with hylan: a treatment schedule study. Curr Ther Res Clin Exp. 1994;55:220-32.

**Adams ME, Atkinson MH, Lussier AJ, Schulz JI, Siminovitch KA, Wade JP, et al.. The role of viscosupplementation with hylan G-F 20 (Synvisc) in the treatment of osteoarthritis of the knee: a Canadian multicenter trial comparing hylan G-F 20 alone, hylan G-F 20 with non-steroidal anti-inflammatory drugs (NSAIDs) and NSAIDs alone. Osteoarthritis Cartilage. 1995;3:213-25.

**Carrabba M, Paresce E, Angelini M, Re KA, Torchiana EEM, Perbellini A. The safety and efficacy of different dose schedules of hyaluronic acid in the treatment of painful osteoarthritis of the knee with joint effusion. Eur J Rheumatol Inflamm. 1995;15:25-31.

**Corrado EM, Peluso GF, Gigliotti S, De DC, Palmieri D, Savoia N, et al.. The effects of intra-articular administration of hyaluronic acid on osteoarthritis of the knee: a clinical study with immunological and biochemical evaluations. Eur J Rheumatol Inflamm. 1995;15:47-56.

**Formiguera Sala S, Esteve de Miguel R. Intra-articular hyaluronic acid in the treatment osteoarthritis of the knee: a short term study. Eur J Rheumatol Inflamm. 1995;15:33-8.

Guler M, Kuran B, Parlar D, Guler M, Saglam H, Yapici S, et al. Clinical trial of intra-articular injection of hyaluronic acid in patients with osteoarthritis of the knee [Abstract]. Presented at X National Rheumatology Congress, Pamukkale-Denizli, Turkey, 29 October–3 November 1996.

**Lohmander LS, Dalén N, Englund G, Hämäläinen M, Jensen EM, Karlsson K, et al.. Intra-articular hyaluronan injections in the treatment of osteoarthritis of the knee: a randomised, double blind, placebo controlled multicentre trial. Hyaluronan Multicentre Trial Group. Ann Rheum Dis. 1996;55:424-31.

Graf von der Schulenburg JM, Allhoff PG. Cost-effectiveness and quality of life of treatment of gonarthrosis with hyaluronic acid [Abstract 187]. Rheumatol Eur. 1997;26:Suppl 2191.

Kalay S. The effectiveness of intra-articular hyaluronic acid treatment in primary gonarthrosis [Specialization thesis]. Ankara, Turkey Ministry of Health, Republic of Turkey1997.

**Listrat V, Ayral X, Patarnello F, Bonvarlet JP, Simonnet J, Amor B, et al.. Arthroscopic evaluation of potential structure modifying activity of hyaluronan (Hyalgan) in osteoarthritis of the knee. Osteoarthritis Cartilage. 1997;5:153-60.

Schneider U, Miltner O, Graf J, Thomsen M, Niethard FU. [Mechanism of action of hyaluronic acid in gonarthrosis of both knee joints in a right/left comparison. Study with dynamometry, oxygen partial pressure, temperature and Lequesne score]. Z Orthop Ihre Grenzgeb. 1997;135:341-7.

Wu JJ, Shih LY, Hsu HC, Chen TH. The double-blind test of sodium hyaluronate (ARTZ) on osteoarthritis knee. Zhonghua Yi Xue Za Zhi (Taipei). 1997;59:99-106.

**Altman RD, Moskowitz R. Intraarticular sodium hyaluronate (Hyalgan) in the treatment of patients with osteoarthritis of the knee: a randomized clinical trial. Hyalgan Study Group. J Rheumatol. 1998;25:2203-12.

**Wobig M, Dickhut A, Maier R, Vetter G. Viscosupplementation with hylan G-F 20: a 26-week controlled trial of efficacy and safety in the osteoarthritic knee. Clin Ther. 1998;20:410-

23.

**Huskisson EC, Donnelly S. Hyaluronic acid in the treatment of osteoarthritis of the knee. Rheumatology (Oxford). 1999;38:602-7.

**Payne MW, Petrella RJ. Viscosupplementation effect on proprioception in the osteoarthritic knee. Arch Phys Med Rehabil. 2000;81:598-603.

Renklitepe N, Atalay E. The effect of intra-articular sodium hyaluronate therapy in knee osteoarthritis [Abstract 328]. Ann Rheum Dis. 2000;59:Suppl142.

Ardiç F, Bolulu D, Topuz O, Cubukçu S. Efficacy of intra-articular hyaluronic acid injections in knee osteoarthritis [Abstract 75]. Ann Rheum Dis. 2001;60:Suppl 1232.

**Bunyaratavej N, Chan KM, Subramanian N. Treatment of painful osteoarthritis of the knee with hyaluronic acid. Results of a multicenter Asian study. J Med Assoc Thai. 2001;84:Suppl 2S576-81.

Caracuel MA, Muñoz-Villanueva MC, Escudero A, Veroz R, Frias G, Vacas J, et al.. Effects of joint lavage and hyaluronic acid infiltration in patients with osteoarthritis of the knee [Abstract 91]. Ann Rheum Dis. 2001;60:Suppl236.

Dickson DJ, Hosie G, English JR, Primary Care Rheumatology Society OA Knee Study Group. A double-blind, placebo-controlled comparison of hylan G-F 20 against diclofenac in knee osteoarthritis. J Clin Research. 2001;4:41-52.

Groppa LG, Moshneaga M. Studying of the efficiency of the synvisk in osteoartrosis [Abstract 144]. Ann Rheum Dis. 2001;60:Suppl 1230.

Seikagaku Corporation. Summary of safety and effectiveness data: Sodium hyaluronate. Bethesda, MD: U.S. Food and Drug Administration; 2001. Accessed at *www.accessdata.fda.gov/cdrh_docs/pdf/P980044b.pdf* on 27 February 2012.

*Tamir E, Robinson D, Koren R, Agar G, Halperin N. Intra-articular hyaluronan injections for the treatment of osteoarthritis of the knee: a randomized, double blind, placebo controlled study. Clin Exp Rheumatol. 2001;19:265-70.

Bütün B, Kaçar C, Evcik D. Intra-articular injection of sodium hyaluronate in the treatment of knee osteoarthritis. Romatizma. 2002;17:31-8.

Cogalgil S, Hatipoglu F. The effects of intra-articular sodium-hyaluronan in patients with gonarthrosis treated with physical therapy [Abstract 282]. Presented at the European League Against Rheumatism, 3rd Annual European Congress of Rheumatology, Stockholm, Sweden, 12–15 June 2002. Accessed at *www.abstracts2view.com/eular/view.php?nu=EULAR2L1_2002AB0282&terms=* on 21 May 2012.

Hizmetli S, Kocagil S, Kaptanoglu E, Elden H, Nacitarhan V. The efficacy and safety of intra-articular hyaluronic acid in osteoarthritis of the knee: a prospective, double-blind trial. Pamphlet provided at the European League Against Rheumatism, 3rd Annual European Congress of Rheumatology, Stockholm, Sweden, 12–15 June 2002.

Karlsson J, Sjögren LS, Lohmander LS. Comparison of two hyaluronan drugs and placebo in patients with knee osteoarthritis. A controlled, randomized, double-blind, parallel-design multicentre study. Rheumatology (Oxford). 2002;41:1240-8.

*Miltner O, Schneider U, Siebert CH, Niedhart C, Niethard FU. Efficacy of intraarticular hyaluronic acid in patients with osteoarthritis—a prospective clinical trial. Osteoarthritis Cartilage. 2002;10:680-6.

**Petrella RJ, DiSilvestro MD, Hildebrand C. Effects of hyaluronate sodium on pain and physical functioning in osteoarthritis of the knee: a randomized, double-blind, placebo-controlled clinical trial. Arch Intern Med. 2002;162:292-8.

**Raynauld JP, Torrance GW, Band PA, Goldsmith CH, Tugwell P, Walker V, et al., Canadian Knee OA Study Group. A prospective, randomized, pragmatic, health outcomes trial evaluating the incorporation of hylan G-F 20 into the treatment paradigm for patients with knee osteoarthritis (Part 1 of 2): clinical results. Osteoarthritis Cartilage. 2002;10:506-17.

Saravanan V, Morgan T, Stobbs D, Daymond TJ. Inflammatory effusion after viscosupplementation with Hylan G-F 20 [Abstract 336]. Rheumatology. 2002;41:Suppl 1121.

*Bayramoğlu M, Karataş M, Cetin N, Akman N, Sözay S, Dilek A. Comparison of two different viscosupplements in knee osteoarthritis—a pilot study. Clin Rheumatol. 2003;22:118-22.

**Jubb RW, Piva S, Beinat L, Dacre J, Gishen P. A one-year, randomised, placebo (saline) controlled clinical trial of 500-730 kDa sodium hyaluronate (Hyalgan) on the radiological change in osteoarthritis of the knee. Int J Clin Pract. 2003;57:467-74.

**Kahan A, Lleu PL, Salin L. Prospective randomized study comparing the medicoeconomic benefits of Hylan GF-20 vs. conventional treatment in knee osteoarthritis. Joint Bone Spine. 2003;70:276-81.

**Tetik S, Ones K, Tetik C. Efficacy of intra-articular Hylan G-F 20 on osteoarthritis of the knee. The Pain Clinic. 2003;15:459-66.

Tsai CL, Chang CC, Chen SC, Beinat L, Piva S. Treatment of knee osteoarthritis in Asian population with an intra-articular hyaluronan of MW500–730 KDa [Abstract P333]. Osteoarthritis Cartilage. 2003;11:Suppl A119.

**Altman RD, Akermark C, Beaulieu AD, Schnitzer T, Durolane International Study Group. Efficacy and safety of a single intra-articular injection of non-animal stabilized hyaluronic acid (NASHA) in patients with osteoarthritis of the knee. Osteoarthritis Cartilage. 2004;12:642-9.

**Day R, Brooks P, Conaghan PG, Petersen M, Multicenter Trial Group. A double blind, randomized, multicenter, parallel group study of the effectiveness and tolerance of intraarticular hyaluronan in osteoarthritis of the knee. J Rheumatol. 2004;31:775-82.

Groppa L, Golubciuc S, Vremis L, Dutca L, Sincari L. The efficacy of combined chondroprotective treatment in the osteoarthritis of the knee [Abstract 415]. Zurich, Switzerland: European League Against Rheumatism; 2004. Accessed at *www.abstracts2view.com/eular/view.php?nu=EULAR04L1_2004FRI0415&terms* on 27 February 2012.

**Pham T, Le Henanff A, Ravaud P, Dieppe P, Paolozzi L, Dougados M. Evaluation of the symptomatic and structural efficacy of a new hyaluronic acid compound, NRD101, in comparison with diacerein and placebo in a 1 year randomised controlled study in

symptomatic knee osteoarthritis. Ann Rheum Dis. 2004;63:1611-7.

Wu HB, Du JY, Yang SH, Shao ZW, Xiong XQ. [Evaluation on the effects of hyaluronan combined with different dosages of celecoxib for relieving pain and ankylosis induced by knee osteoarthritis]. Zhongguo Linchuang Kangfu. 2004;8:5491-3.

**Cubukçu D, Ardiç F, Karabulut N, Topuz O. Hylan G-F 20 efficacy on articular cartilage quality in patients with knee osteoarthritis: clinical and MRI assessment. Clin Rheumatol. 2005;24:336-41.

*Huang MH, Yang RC, Lee CL, Chen TW, Wang MC. Preliminary results of integrated therapy for patients with knee osteoarthritis. Arthritis Rheum. 2005;53:812-20.

**Neustadt D, Caldwell J, Bell M, Wade J, Gimbel J. Clinical effects of intraarticular injection of high molecular weight hyaluronan (Orthovisc) in osteoarthritis of the knee: a randomized, controlled, multicenter trial. J Rheumatol. 2005;32:1928-36.

**Sezgin M, Demirel AC, Karaca C, Ortancil O, Ulkar GB, Kanik A, et al.. Does hyaluronan affect inflammatory cytokines in knee osteoarthritis? Rheumatol Int. 2005;25:264-9.

**Kotevoglu N, Iyibozkurt PC, Hiz O, Toktas H, Kuran B. A prospective randomised controlled clinical trial comparing the efficacy of different molecular weight hyaluronan solutions in the treatment of knee osteoarthritis. Rheumatol Int. 2006;26:325-30.

**Petrella RJ, Petrella M. A prospective, randomized, double-blind, placebo controlled study to evaluate the efficacy of intraarticular hyaluronic acid for osteoarthritis of the knee. J Rheumatol. 2006;33:951-6.

**Atay T, Aslan A, Baydar ML, Ceylan B, Baykal B, Kirdemir V. [The efficacy of low- and high-molecular-weight hyaluronic acid applications after arthroscopic debridement in patients with osteoarthritis of the knee]. Acta Orthop Traumatol Turc. 2008;42:228-33.

**Blanco FJ, Fernández-Sueiro JL, Pinto-Tasende JA, Fernández-López JC, Ramallal M, Freire A, et al.. Intra-articular hyaluronan treatment of patients with knee osteoarthritis waiting for replacement surgery. Open Arthritis J. 2008;1:1-7.

*Heybeli N, Doral MN, Atay OA, Leblebicioğlu G, Uzümcügil A. [Intra-articular sodium hyaluronate injections after arthroscopic debridement for osteoarthritis of the knee: a prospective, randomized, controlled study]. Acta Orthop Traumatol Turc. 2008;42:221-7.

Lundsgaard C, Dufour N, Fallentin E, Winkel P, Gluud C. Intra-articular sodium hyaluronate 2 mL versus physiological saline 20 mL versus physiological saline 2 mL for painful knee osteoarthritis: a randomized clinical trial. Scand J Rheumatol. 2008;37:142-50.

**Petrella RJ, Cogliano A, Decaria J. Combining two hyaluronic acids in osteoarthritis of the knee: a randomized, double-blind, placebo-controlled trial. Clin Rheumatol. 2008;27:975-81.

**Altman RD, Rosen JE, Bloch DA, Hatoum HT, Korner P. A double-blind, randomized, saline-controlled study of the efficacy and safety of EUFLEXXA for treatment of painful osteoarthritis of the knee, with an open-label safety extension (the FLEXX trial). Semin Arthritis Rheum. 2009;39:1-9.

**Baltzer AW, Moser C, Jansen SA, Krauspe R. Autologous conditioned serum (Orthokine) is an effective treatment for knee osteoarthritis. Osteoarthritis Cartilage. 2009;17:152-60.

*Baraf HS, Strand V, Hosokawa H, Akahane O, Lim S, Yaguchi M. Effectiveness and safety of a single intraarticular injection of gel-200, a new cross-linked formulation of hyaluronic acid [HA] in the treatment of symptomatic osteoarthritis [OA] of the knee [Abstract 326]. Osteoarthritis Cartilage. 2009;17:Suppl 1S174.

**Diracoglu D, Vural M, Baskent A, Dikici F, Aksoy C. The effect of viscosupplementation on neuromuscular control of the knee in patients with osteoarthritis. J Back Musculoskelet Rehabil. 2009;22:1-9.

Petrella RJ, Decaria JE, Wolfe D, Chesworth B, Shapiro S, Montero-Odasso M. The effect of hyaluronic acid on gait in knee osteoarthritis patients: preliminary results for a randomized, double-blind, placebo controlled study [Abstract 375]. Ann Rheum Dis. 2009;68:Suppl 3479.

**Chevalier X, Jerosch J, Goupille P, van Dijk N, Luyten FP, Scott DL, et al.. Single, intra-articular treatment with 6 ml hylan G-F 20 in patients with symptomatic primary osteoarthritis of the knee: a randomised, multicentre, double-blind, placebo controlled trial. Ann Rheum Dis. 2010;69:113-9.

**Jørgensen A, Stengaard-Pedersen K, Simonsen O, Pfeiffer-Jensen M, Eriksen C, Bliddal H, et al.. Intra-articular hyaluronan is without clinical effect in knee osteoarthritis: a multicentre, randomised, placebo-controlled, double-blind study of 337 patients followed for 1 year. Ann Rheum Dis. 2010;69:1097-102.

Kosuwon W, Sirichatiwapee W, Visanuyotin T, Jeeravipoolvarn P, Laupattarakasem W. An efficacy study on cartilage volume by MRI findings in patient with knee osteoarthritis between 25 mg of sodium hyaluronate (2.5 ml) to placebo [Abstract 370]. Ann Rheum Dis. 2010;69:Suppl 3267.

**Kul-Panza E, Berker N. Is hyaluronate sodium effective in the management of knee osteoarthritis? A placebo-controlled double-blind study. Minerva Med. 2010;101:63-72.

*Pavelka K, Niethard FU, Giordan N. A multicentre, international, double blind, randomized, placebo-controlled study to assess the efficacy and safety of 2 different regimens of HYADD4-G in knee osteoarthritis [Abstract 326]. Osteoarthritis Cartilage. 2010;18:Suppl 2S144.

Sanofi-Aventis. TREAD-20: Trial of Hyalgan three injection-regimen for the treatment of knee pain due to osteoarthritis [clinical trial]. Accessed at *http://clinicaltrials.gov/ct2/show/NCT00130468* on 21 February 2012.

**Huang TL, Chang CC, Lee CH, Chen SC, Lai CH, Tsai CL. Intra-articular injections of sodium hyaluronate (Hyalgan®) in osteoarthritis of the knee. a randomized, controlled, double-blind, multicenter trial in the Asian population. BMC Musculoskelet Disord. 2011;12:221.

**Navarro-Sarabia F, Coronel P, Collantes E, Navarro FJ, de la Serna AR, Naranjo A, et al., AMELIA study group. A 40-month multicentre, randomised placebo-controlled study to assess the efficacy and carry-over effect of repeated intra-articular injections of hyaluronic acid in knee osteoarthritis: the AMELIA project. Ann Rheum Dis. 2011;70:1957-62.

Brandt KD, Block JA, Michalski JP, Moreland LW, Caldwell JR, Lavin PT. Efficacy and safety of intraarticular sodium hyaluronate in knee osteoarthritis. ORTHOVISC Study Group. Clin Orthop Relat Res. 2001;130-43.

Anothaisintawee, T., J. Attia, et al. (2011). "Management of Chronic Prostatitis/Chronic Pelvic Pain Syndrome." JAMA: the journal of the American Medical Association 305(1): 78-86.

**Nickel JC, Krieger JN, McNaughton-Collins M, et al; Chronic Prostatitis Collaborative Research Network. Alfuzosin and symptoms of chronic prostatitis-chronic pelvic pain syndrome. N Engl J Med. 2008;359(25):2663-2673

*Tuğcu V, Taşçi AI, Fazlioğlu A, et al. A placebo-controlled comparison of the efficiency of triple- and monotherapy in category III B chronic pelvic pain syndrome (CPPS). Eur Urol. 2007;51(4):1113-1117
PubMed | Link to Article

Goldmeier D, Madden P, McKenna M, Tamm N. Treatment of category III A prostatitis with zafirlukast: a randomized controlled feasibility study. Int J STD

Nickel JC, Pontari M, Moon T, et al; Rofecoxib Prostatitis Investigator Team. A randomized, placebo controlled, multicenter study to evaluate the safety and efficacy of rofecoxib in the treatment of chronic nonbacterial prostatitis. J Urol. 2003;169(4):1401-1405

*Elist J. Effects of pollen extract preparation Prostat/Poltit on lower urinary tract symptoms in patients with chronic nonbacterial prostatitis/chronic pelvic pain syndrome: a randomized, double-blind, placebo-controlled study. Urology. 2006;67(1):60-63

**Shoskes DA, Zeitlin SI, Shahed A, Rajfer J. Quercetin in men with category III chronic prostatitis. Urology. 1999;54(6):960-963
.

**Wagenlehner FM, Schneider H, Ludwig M, et al. A pollen extract (Cernilton) in patients with inflammatory chronic prostatitis-chronic pelvic pain syndrome. Eur Urol. 2009;56(3):544-551

**Alexander RB, Propert KJ, Schaeffer AJ, et al; Chronic Prostatitis Collaborative Research Network. Ciprofloxacin or tamsulosin in men with chronic prostatitis/chronic pelvic pain syndrome. Ann Intern Med. 2004;141(8):581-589

**Bates SM, Hill VA, Anderson JB, et al. A prospective, randomized, double-blind trial to evaluate the role of a short reducing course of oral corticosteroid therapy in the treatment of chronic prostatitis/chronic pelvic pain syndrome. BJU Int. 2007;99(2):355-359

*Nickel JC, Narayan P, McKay J, Doyle C. Treatment of chronic prostatitis/chronic pelvic pain syndrome with tamsulosin. J Urol. 2004;171(4):1594-1597

**Cheah PY, Liong ML, Yuen KH, et al. Terazosin therapy for chronic prostatitis/chronic pelvic pain syndrome. J Urol. 2003;169(2):592-596
PubMed | Link to Article

*Evliyaoğlu Y, Burgut R. Lower urinary tract symptoms, pain and quality of life assessment in chronic non-bacterial prostatitis patients treated with α-blocking agent doxazosin; vs placebo. Int Urol Nephrol. 2002;34(3):351-356
PubMed | Link to Article

*Gül O, Eroğlu M, Ozok U. Use of terazosine in patients with chronic pelvic pain syndrome and evaluation by Prostatitis Symptom Score Index. Int Urol Nephrol. 2001;32(3):433-436

*Jeong CW, Lim DJ, Son H, et al. Treatment for chronic prostatitis/chronic pelvic pain syndrome: levofloxacin, doxazosin and their combination. Urol Int. 2008;80(2):157-161

**Leskinen M, Lukkarinen O, Marttila T. Effects of finasteride in patients with inflammatory chronic pelvic pain syndrome. Urology. 1999;53(3):502-505

*Mehik A, Alas P, Nickel JC, et al. Alfuzosin treatment for chronic prostatitis/chronic pelvic pain syndrome. Urology. 2003;62(3):425-429

**Nickel JC, Downey J, Clark J, et al. Levofloxacin for chronic prostatitis/chronic pelvic pain syndrome in men. Urology. 2003;62(4):614-617

**Nickel JC, Downey J, Pontari MA, et al. A randomized placebo-controlled multicentre study to evaluate the safety and efficacy of finasteride for male chronic pelvic pain syndrome (category IIIA chronic nonbacterial prostatitis). BJU Int. 2004;93(7):991-995

Nickel JC, Forrest JB, Tomera K, et al. Pentosan polysulfate sodium therapy for men with chronic pelvic pain syndrome. J Urol. 2005;173(4):1252-1255

*Ye ZQ, Lan RZ, Yang WM, et al. Tamsulosin treatment of chronic non-bacterial prostatitis. J Int Med Res. 2008;36(2):244-252

*Zhao WP, Zhang ZG, Li XD, et al. Celecoxib reduces symptoms in men with difficult chronic pelvic pain syndrome (category IIIA). Braz J Med Biol Res. 2009;42(10):963-967

*Zhou Z, Hong L, Shen X, et al. Detection of nanobacteria infection in type III prostatitis. Urology. 2008;71(6):1091-1095

**Pontari MA, Krieger JN, Litwin MS, et al; Chronic Prostatitis Collaborative Research Network-2. Pregabalin for the treatment of men with chronic prostatitis/chronic pelvic pain syndrome. Arch Intern Med. 2010;170(17):1586-1593

**Thangaratinam, S., E. Rogozińska, et al. (2012). "Effects of interventions in pregnancy on maternal weight and obstetric outcomes: meta-analysis of randomised evidence." BMJ: British Medical Journal 344.

*Badrawi H, Hassanein MK, Badraoui MHH, Wafa YA, Shawky HA, Badrawi N. Pregnancy outcome in obese pregnant mothers. J Perinat Med1992;20:203.

*Bechtel-Blackwell DA. Computer-assisted self-interview and nutrition education in pregnant teens. Clin Nurs Res2002;11:450-62.

*Briley C, Flanagan NL, Lewis N. In-home prenatal nutrition intervention increased dietary iron intakes and reduced low birthweight in low-income African-American women. J Am Diet Assoc2002;102:984-7.

Clapp IJF. Diet, exercise, and fete-placental growth. Arch Gynecol Obstet1997;260:101-8.

**Crowther CA, Hiller JE, Moss JR, McPhee AJ, Jeffries WS, Robinson JS. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. N Engl J Med2005;352:2477-86.

Gomez TG, Delgado JG, Agudelo AA, Hurtado H. Diet effects on the perinatal result of obese pregnant patient. [Spanish]. Rev Colomb Obstet Ginecol1994;45:313-6.

**Khoury J, Henriksen T, Christophersen B, Tonstad S. Effect of a cholesterol-lowering diet on maternal, cord, and neonatal lipids, and pregnancy outcome: a randomized clinical trial. Am J Obstet Gynecol2005;193:1292-301.

**Landon MB, Spong CY, Thom E, Carpenter MW, Ramin SM, Casey B, et al. A multicenter, randomized trial of treatment for mild gestational diabetes. N Engl J Med2009;361:1339-48.

*Ney D, Hollingsworth DR, Cousins L. Decreased insulin requirement and improved control of diabetes in pregnant women given a high-carbohydrate, high-fiber, low-fat diet. Diabetes Care1982;5:529-33.

**Quinlivan JA, Lam LT, Fisher J. A randomised trial of a four-step multidisciplinary approach to the antenatal care of obese pregnant women. Aust N Z J Obstet Gynaecol2011;51:141-6.

*Rae A, Bond D, Evans S, North F, Roberman B, Walters B. A randomised controlled trial of dietary energy restriction in the management of obese women with gestational diabetes. Aust N Z J Obstet Gynaecol2000;40:416-22.

*Thornton YS, Smarkola C, Kopacz SM, Ishoof SB. Perinatal outcomes in nutritionally monitored obese pregnant women: a randomized clinical trial. J Natl Med Assoc2009;101:569-77.

**Wolff S, Legarth J, Vangsgaard K, Toubro S, Astrup A. A randomized trial of the effects of dietary counseling on gestational weight gain and glucose metabolism in obese pregnant women. Int J Obes2008;32:495-501.

**Baciuk EP, Pereira RI, Cecatti JG, Braga AF, Cavalcante SR. Water aerobics in pregnancy: cardiovascular response, labor and neonatal outcomes. Reprod Health2008;5:10.

**Barakat R, Lucia A, Ruiz JR. Resistance exercise training during pregnancy and newborn's birth size: a randomised controlled trial. Int J Obes2009;33:1048-57.

Barakat R, Cordero Y, Coteron J, Luaces M, Montejo R. Exercise during pregnancy improves maternal glucose screen at 24-28 weeks: a randomised controlled trial. Br J Sports Med2011 Sep 26, epub ahead of print.

*Bell RJ, Palma SM. Antenatal exercise and birth-weight. Aust N Z J Obstet Gynaecol2000;40:70-3.

**Cavalcante SR, Cecatti JG, Pereira RI, Baciuk EP, Bernardo AL, Silveira C. Water aerobics II: maternal body composition and perinatal outcomes after a program for low risk pregnant women. Reprod Health2009;6:1.

*Clapp JF, III, Kim H, Burciu B, Lopez B. Beginning regular exercise in early pregnancy: effect on fetoplacental growth. Am J Obstet Gynecol2000;183:1484-8.

*Erkkola R. The influence of physical exercise during pregnancy upon physical work capacity and circulatory parameters. Scand J Clin Lab Invest1976;6:747-9.

**Erkkola R, Makela M. Heart volume and physical fitness of parturients. Ann Clin Res1976;8:15-21.

**Garshasbi A, Faghih ZS. The effect of exercise on the intensity of low back pain in pregnant women. Int J Gynaecol Obstet2005;88:271-5.

**Haakstad L, Bo K. Exercise in pregnant women and birth weight: a randomized controlled trial. BMC Preg Childbirth2011;11:66.

**Hopkins SA, Baldi JC, Cutfield WS, McCowan L, Hofman PL. Exercise training in pregnancy reduces offspring size without changes in maternal insulin sensitivity. J Clin Endocrinol Metab2010;95:2080-8.

*Khaledan A, Sh, Motahari Tabari NS, Ahmad Shirvani M. Effect of an aerobic exercise program on fetal growth in pregnant women. HAYAT: J Faculty Nurs Midwifery 2010;16:78.

Lee G, Challenger S, McNabb M, Sheridan M. Exercise in pregnancy. Mod Midwife1996;6:28-33.
*Marquez-Sterling S, Perry AC, Kaplan TA, Halberstein RA, Signorile JF. Physical and psychological changes with vigorous exercise in sedentary primigravidae. Med Sci Sports Exerc2000;32:58-62.

*Ong MJ, Guelfi KJ, Hunter T, Wallman KE, Fournier PA, Newnham JP. Supervised home-based exercise may attenuate the decline of glucose tolerance in obese pregnant women. Diabetes Metab2009;35:418-21.

Prevedel T, Calderon I, DM, Adami H-O, RM. Maternal and perinatal effects of hydrotherapy in pregnancy. Rev Bras Ginecol Obstet2003;25:53-9.

**Santos IA, Stein R, Fuchs SC, Duncan BB, Ribeiro JP, Kroeff LR, et al. Aerobic exercise and submaximal functional capacity in overweight pregnant women: a randomized trial. Obstet Gynecol2005;106:243-9.

Sedaghati P, Ziaee V, Ardjmand A. The effect of an ergometric training program on pregnants weight gain and low back pain. Gazzetta Medica Italiana Archivio per le Scienze Mediche2007;166:209-13.

*Yeo S, Steele NM, Chang MC, Leclaire SM, Ronis DL, Hayashi R. Effect of exercise on blood pressure in pregnant women with a high risk of gestational hypertensive disorders. J Reprod Med2000;45:293-8.

**Asbee SM, Jenkins TR, Butler JR, White J, Elliot M, Rutledge A. Preventing excessive weight gain during pregnancy through dietary and lifestyle counseling: a randomized controlled trial. Obstet Gynecol2009;113:305-12.

Bung P, Artal R, Khodiguian N, Kjos S. Exercise in gestational diabetes. An optional therapeutic approach? Diabetes1991;40(suppl 2):182-5.

*Ferrara A, Hedderson MM, Albright CL, Ehrlich SF, Quesenbery CP, Peng TP, et al. A pregnancy and postpartum lifestyle intervention in women with gestational diabetes mellitus reduces diabetes risk factors. A feasibility randomized control trial. Diabetes Care2011;34:1519-25.

**Guelinckx I, Devlieger R, Mullie P, Vansant G. Effect of lifestyle intervention on dietary habits, physical activity, and gestational weight gain in obese pregnant women: a randomized controlled trial. Am J Clin Nutr2010;91:373-80.

*Huang TT, Yeh CY, Tsai YC. A diet and physical activity intervention for preventing weight retention among Taiwanese childbearing women: a randomised controlled trial. Midwifery2011;27:257-64.

**Hui A, Back L, Ludwig S, Gardiner P, Sevenhuysen G, Dean H, et al. Lifestyle intervention on diet and exercise reduced excessive gestational weight gain in pregnant women under a randomised controlled trial. BJOG2011;119:70-7.

**Hui AL, Ludwig SM, Gardiner P, Sevenhuysen G, Murray R, Morris M, et al. Community-based exercise and dietary intervention during pregnancy: a pilot study. Can J Diabetes2006;30:169-75.

Jeffries K, Shub A, Walker SP, Hiscock R, Permezel M. Reducing excessive weight gain in pregnancy: a randomised controlled trial. Med J Aust2009;191:429-33.

*Kulpa PJ, White BM, Visscher R. Aerobic exercise in pregnancy. Am J Obstet Gynecol1987;156:1395-403.

**Phelan S, Phipps MG, Abrams B, Darroch F, Schaffner A, Wing RR. Randomized trial of a behavioral intervention to prevent excessive gestational weight gain: the Fit for Delivery Study. Am J Clin Nutr2011;93:772-9.

*Polley BA, Wing RR, Sims CJ. Randomized controlled trial to prevent excessive weight gain in pregnant women. Int J Obes2002;26:1494-502.

Vinter CA, Jensen DM, Ovesen P, Beck-Nielsen H, Jorgensen JS. The LiP (Lifestyle in Pregnancy) study: a randomized controlled trial of lifestyle intervention in 360 obese pregnant women. Diabetes Care2011;34:2502-7.

**Jackson RA, Stotland NE, Caughey AB, Gerbert B. Improving diet and exercise in pregnancy with Video Doctor counseling: a randomized trial. Patient Educ Couns*2011*;*83*:*203*-9

Heneghan, C., A. Ward, et al. (2011). "Self-monitoring of oral anticoagulation: systematic review and meta-analysis of individual patient data." The Lancet.

ME Cromheecke, M Levi, LP Colly et al. Oral anticoagulation self-management and management by a specialist anticoagulation clinic: a randomised cross-over comparison Lancet, 356 (2000), pp. 97–102

H Körtke, K Minami, T Breymann et al. INR self-management after mechanical heart valve replacement: ESCAT (Early Self-Controlled Anticoagulation Trial) Z Kardiol, 90 (suppl 6) (2001), pp. 118–124

**B Menendez-Jandula, JC Souto, A Oliver et al. Comparing self-management of oral anticoagulant therapy with clinic management: a randomized trial
Ann Intern Med, 142 (2005), pp. 1–10

*DA Fitzmaurice, ET Murray, D McCahon et al. Self management of oral anticoagulation: randomised trial BMJ, 331 (2005), p. 1057

**TD Christensen, M Maegaard, HT Sorensen, VE Hjortdal, JM Hasenkam Self-management versus conventional management of oral anticoagulant therapy: a randomized, controlled trial Eur J Intern Med, 17 (2006), pp. 260–266

**DB Matchar, A Jacobson, R Dolor et al. Effect of home testing of international normalized ratio on clinical events N Engl J Med, 363 (2010), pp. 1608–1620

**A Siebenhofer, I Rakovac, C Kleespies, B Piso, U Didjurgeit Self-management of oral anticoagulation reduces major outcomes in the elderly. A randomized controlled trial Thromb Haemost, 100 (2008), pp. 1089–1098

*S Kaatz, J Elston-Lafata, S Gooldy Anticoagulation therapy home and office monitoring evaluation study J Thromb Thrombolysis, 12 (2001), p. 111

**RJ Beyth, L Quinn, CS Landefeld A multicomponent intervention to prevent major bleeding complications in older patients receiving warfarin. A randomized, controlled trial Ann Intern Med, 133 (2000), pp. 687–695

**R Sunderji, K Gin, K Shalansky et al. A randomized trial of patient self-managed versus physician-managed oral anticoagulation Can J Cardiol, 20 (2004), pp. 1117–1123

**H Völler, J Glatz, U Taborski, A Bernardo, C Dovifat, K Heidinger Self-management of oral anticoagulation in nonvalvular atrial fibrillation (SMAAF study) Z Kardiol, 94 (2005), pp. 182–186

Hemmingsen, B., L. L. Christensen, et al. (2012). "Comparison of metformin and insulin versus insulin alone for type 2 diabetes: systematic review of randomised clinical trials with meta-analyses and trial sequential analyses." BMJ: British Medical Journal 344.

**Douek IF, Allen SE, Ewings P, Gale EAM, Bingley PJ. Continuing metformin when starting insulin in patients with type 2 diabetes: a double-blind randomized placebo-controlled trial. Diabet Med2005;22:634-40.

Wulffele MG, Kooy A, Lehert P, Bets D, Ogterop JC, Van Der Burg BB, et al. Combination of insulin and metformin in the treatment of type 2 diabetes. Diabetes Care2002;25:2133-40.

**Relimpio F, Pumar A, Losada F, Mangas MA, Acosta D, Astorga R. Adding metformin versus insulin dose increase in insulin-treated but poorly controlled type 2 diabetes mellitus: an open-label randomized trial. Diabet Med1998;15:997-1002

Heine RJ, Scheen A, Van Gaal L, Schmitt H, Van der Waal PS. Efficacy of bedtime NPH insulin alone, as compared to combination with metformin and/or glipizide in NIDDM patients with secondary failure on oral hypoglycaemic agents. Neth J Med1995;47:A59-60.

*Robinson AC, Burke J, Robinson S, Johnston DG, Elkeles RS. The effects of metformin on glycemic control and serum lipids in insulin-treated NIDDM patients with suboptimal metabolic control. Diabetes Care1998;21:701-5.

Schnack C, Biesenbach G, Kacerovsky G, Mihaljevic R, Pecnik I, Pieber T, et al. Evaluation of optimal therapy in type-2 diabetic patients insufficiently treated with sulfonylureas: the Austrian insulin intervention study. Diabetologia1996;39(suppl 1):A33.

Van der Waal PS, Scheen A, van Gaal L, Schmitt H, Heine RJ. Efficacy of bedtime NPH insulin alone, as compared to combination with metformin and/or gliplizide in NIDDM patients with secondary failure to oral hypoglycemic agents. Diabetes1996;45:286A.

Van der Waal PS, Scheen A, Van Gaal L, Schmitt H, Heine RJ. Predictors of glycaemic efficacy of four treatment strategies in NIDDM patients with secondary failure to oral hypoglycaemic agents. Neth J Med1996;48:A49.

*Hirsch IB. Metformin added to insulin therapy in poorly controlled type 2 diabetes. Diabetes Care1999;22:854.

*Ponssen HH, Elte JW, Lehert P, Schouten JP, Bets D. Combined metformin and insulin therapy for patients with type 2 diabetes mellitus. Clin Ther2000;22:709-18.

**Civera M, Merchante A, Salvador M, Sanz J, Martínez I. Safety and efficacy of repaglinide in combination with metformin and bedtime NPH insulin as an insulin treatment regimen in type 2 diabetes. Diabet Res Clin Pract2008;79:42-7.

*Galani V, Patel HM. Comparison of metformin and insulin monotherapy with combined metformin and insulin therapy in patients of type 2 diabetes with HbA1c > 7%. Int J Pharmaceutical Biol Arch2011;2:563-8.

**Kabadi UM, Kabadi M. Comparative efficacy of glimepiride and/or metformin with insulin in type 2 diabetes. Diabet Res Clin Pract2006;72:265-70.
**Kvapil M, Swatko A, Hilberg C, Shestakova M. Biphasic insulin aspart 30 plus metformin: an effective combination in type 2 diabetes. Diabet Obes Metab2006;8:39-48.

Onuchin SG, Elsukova OS, Solov'ev OV, Onuchina EL. [Capabilities of hypoglycemic therapy in women with decompensated type 2 diabetes mellitus.] [In Russian.] Terapevticheskii Arkhiv2010;82:34-41.

**Ushakova O, Sokolovskaya V, Morozova A, Valeeva F, Zanozina O, Sazonova O, et al. Comparison of biphasic insulin aspart 30 given three times daily or twice daily in combination with metformin versus oral antidiabetic drugs alone in patients with poorly controlled type 2 diabetes: a 16-week, randomized, open-label, parallel-group trial conducted in Russia. Clin Ther2007;29:2374-84.

**Kokic S, Bukovic D, Radman M, Capkun V, Gabric N, Lesko V, et al. Lispro insulin and metformin versus other combination in the diabetes mellitus type 2 management after secondary oral antidiabetic drug failure. Collegium Antropologicum2003;27:181-7.

**Altuntas Y, Ozen B, Ozturk B, Sengul A, Ucak S, Ersoy O, et al. Comparison of additional metformin or NPH insulin to mealtime insulin lispro therapy with mealtime human insulin therapy in secondary OAD failure. Diabet Obes Metab2003;5:371-8.

*Gram J, Henriksen JE, Grodum E, Juhl H, Hansen TB, Christiansen C, et al. Pharmacological treatment of the pathogenetic defects in type 2 diabetes: the randomized multicenter South Danish Diabetes Study. Diabetes Care2011;34:27-33.

Kokic S, Kokic V, Krnic M, Miric L, Jovanovic Z, Orlic-Crncevic Z. Advantage of prandial insulin as a therapeutic approach in initial secondary pancreatic beta-cell exhaustion in type 2 diabetic patients. Diabetologia Croatica2010;39:37-42.

*Vähätalo M, Rönnemaa T, Viikari J. Recognition of fasting or overall hyperglycaemia when starting insulin treatment in patients with type 2 diabetes in general practice. Scand J Prim Health Care2007;25:147-53.

**Strowig SM, Avilés-Santa ML, Raskin P. Comparison of insulin monotherapy and combination therapy with insulin and metformin or insulin and troglitazone in type 2 diabetes. Diabetes Care2002;25:1691-8.

**Avilés-Santa L, Sinding J, Raskin P. Effects of metformin in patients with poorly controlled, insulin-treated type 2 diabetes mellitus. A randomized, double-blind, placebo-controlled trial. Ann Intern Med1999;131:182-8.

**Giugliano D, Quatraro A, Consoli G, Minei A, Ceriello A, De Rosa N, et al. Metformin for obese, insulin-treated diabetic patients: improvement in glycaemic control and reduction of metabolic risk factors. Eur J Clin Pharmacol1993;44:107-12.

**Hermann LS, Kalén J, Katzman P, Lager I, Nilsson A, Norrhamn O, et al. Long-term glycaemic improvement after addition of metformin to insulin in insulin-treated obese type 2 diabetes patients. Diabet Obes Metab2001;3:428-34.

**Yilmaz H, Gursoy A, Sahin M, Guvener Demirag N. Comparison of insulin monotherapy and combination therapy with insulin and metformin or insulin and rosiglitazone or insulin and acarbose in type 2 diabetes. Acta Diabetologica2007;44:187-92.

Leucht, S., M. Tardy, et al. (2012). "Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis." The Lancet.

*Andrews P, Hall JN, Snaith RP. A controlled trial of phenothiazine withdrawal in chronic schizophrenic
patients. *Br J Psychiatry* 1976;**128**:451-5.

Arato M, O´Connor R, Meltzer HY. A 1-year, double-blind, placebo-controlled trial of ziprasidone 40, 80 and 160 mg/day in chronic schizophrenia: the Ziprasidone Extended Use in Schizophrenia (ZEUS) study. *Int Clin Psychopharmacol* 2002;17**:207-15.

*Baro F, Brugmans J, Dom R, van Lommel R. Maintenance therapy of chronic psychotic patients with a weekly oral dose of R 16341. A controlled double-blind study. *J Clin Pharmacol* 1970;**10**:330-41.

Beasley CM, Sutton VK, Hamilton SH et al. A double-blind, randomized, placebo-controlled trial of olanzapine in the prevention of psychotic relapse. *J Clin Psychopharmacol* 2003;23**:582-94.

*Blackburn HL, Allen JL. Behavioral effects of interrupting and resuming tranquilizing medication among schizophrenics. *J Nerv Ment Dis* 1981;**133**:303-8.

Boonstra G, Burger H, Grobbee DE, Kahn RS. Antipsychotic prophylaxis is needed after remission from a first psychotic episode in schizophrenia patients: Results from an aborted randomised trial. *Int J Psychiatry Clin Pract* 2011;15** (2):128-34.

*Caffey EM, Diamond LS, Frank TV et al. Discontinuation or reduction of chemotherapy in chronic schizophrenics. *J Chronic Dis* 1964;**17**:347-58.

*Channabasavanna SM, Michael A. Penfluridol maintenance therapy in schizophrenia: a controlled study. *Indian J Psychiatry* 1987;**29**:333-6.

Chen EYH, Hui CLM, Lam M et al. A double-blind randomized placebo-controlled study of relapseprevention in remitted first-episode psychosis patients following one year of maintenance therapy 119. *BMJ* 2010;**341**:c4024.

Cheung HK. Schizophrenics fully remitted on neuroleptics for 3-5 years - to stop or continue drugs? *Br J Psychiatry* 1981;**138**:490-4.

Clark ML, Huber WK, Hill D, Wood F, Costiloe JP. Pimozide in chronic schizophrenic outpatients. *Dis Nerv Syst* 1975;36**:137-41.

Cooper SJ, Butler A, Tweed J, Welch C, Raniwalla J. Zotepine in the prevention of recurrence: a randomised, double-blind, placebo-controlled study for chronic schizophrenia. *Psychopharmacology* 2000;150**:237-43.

Crow TJ, MacMillan JF, Johnson AL, Johnstone EC. A randomised controlled trial of prophylactic neuroleptic treatment. *Br J Psychiatry* 1986;148**:120-7.

*Denijs EL, Vereeken JL. Pimozide (OrapR, R 6238) in residual schizophrenia. A clinical evaluation with long-term double-blind follow-up. *Psychiatr Neurol Neurochir* 1973;**76**:47-59.

Dotti A, Bersani G, Rubino IA, Eliseo C. Double-blind trial of fluphenazine decanoate against placebo in ambulant maintenance treatment of chronic schizophrenics. *Riv Psichiatr* 1979;**14**:374-83.

Eklund K, Forsman A. Minimal effective dose and relapse - double-blind trial: haloperidol decanoate vs. placebo. *Clin Neuropharmacol* 1991;**14**:S7-S15.

Elie R, Gagnon MA, Gauthier R, Jequier JC. Effects of neuroleptic withdrawal on the drug-induced extrapyramidal syndrome of chronic schizophrenia. *Union Med Can* 1975;**104**:909-14.

Freeman LS, Alson E. Prolonged withdrawal of chlorpromazine in chronic patients. *Dis Nerv Syst* 1962;**23**:522-5.

Gallant DM, Mielke DH, Spirtes MA, Swanson WC, Bost R. Penfluridol: an efficacious long-acting oral antipsychotic compound. *Am J Psychiatry* 1974;**131**:699-702.

*Gardos G, Cole JO, Rapkin RM et al. Anticholinergic challenge and neuroleptic withdrawal. Changes in dyskinesia and symptom measures. *Arch Gen Psychiatry* 1984;**41**:1030-5.

*Garfield SL, Gershon S, Sletten I, Neubauer H, Ferrel E. Withdrawal of ataractic medication in schizophrenic patients. *Dis Nerv Syst* 1966;**27**:321-5.

*Gitlin M, Nuechterlein K, Subotnik KL et al. Clinical outcome following neuroleptic discontinuation in
patients with remitted recent-onset schizophrenia. *Am J Psychiatry* 2001;**158**:1835-42.

*Goldberg SC, Shenoy RS, Sadler A, Hamer R, Ross B. The effects of a drug holiday on relapse and tardive dyskinesia in chronic schizophrenics. *Psychopharmacol Bull* 1981;**17**:116-7.

Gross M, Hitchman IL, Reeves WP, Lawrence J, Newell PC. Discontinuation of treatment with ataractic drugs. A preliminary report. *Am J Psychiatry* 1960;**116**:931-2.

Gross HS. A double-blind comparison of once-a-day pimozide, trifluoperazine, and placebo in the maintenance care of chronic schizophrenic outpatients. *Curr Ther Res* 1974;16**:696-705.

*Hershon HI, Kennedy PF, McGuire RJ. Persistence of extra-pyramidal disorders and psychiatric relapse
after withdrawal of long-term phenothiazine therapy. *Br J Psychiatry* 1972;**120**:41-50.

Hirsch SR, Gaind R, Rohde PD, Stevens BC, Wing JK. Outpatient maintenance of chronic schizophrenic patients with long-acting fluphenazine: double-blind placebo trial. Report to the Medical Research Council Committee on Clinical Trials in Psychiatry. *BMJ* 1973;1**:633-7.

Hirsch SR, Bowen JT, Emami J et al. A one year prospective study of the effect of life events and medication in the aetiology of schizophrenic relapse. *Br J Psychiatry* 1996;**168**:49-56.

Hogarty GE, Goldberg SC. Drug and sociotherapy in the aftercare of schizophrenic patients. One-year relapse rates. *Arch Gen Psychiatry* 1973;**28**:54-64.

Hough D, Gopal S, Vijapurkar U et al. Paliperidone palmitate maintenance treatment in delaying the time-to-relapse in patients with schizophrenia: a randomized, double-blind, placebo-controlled study. *Schizophr Res* 2010;**116**:107-17.

Kane JM, Rifkin A, Quitkin F et al. Low dose fluphenazine decanoate in maintenance treatment of schizophrenia. *Psychiatry Res* 1979;1**:341-8.

Kane JM, Rifkin A, Quitkin F, Nayak D, Ramos-Lorenzi J. Fluphenazine vs. placebo in patients with remitted, acute first-episode schizophrenia. *Arch Gen Psychiatry* 1982;39**:70-3.

Keskiner A, Holden JMC, Itil TM. Maintenance treatment of schizophrenic outpatients with a depot phenothiazine. *Psychosomatics* 1968;**9**:166-71.

Kramer M, Simpson G, Maciulis V et al. Paliperidone extended-release tablets for prevention of symptom recurrence in patients with schizophrenia: a randomized, double-blind, placebo-controlled study. *J Clin Psychopharmacol* 2007;27**:6-14.

Kurland AA, Ota KY, Slotnick VB. Penfluridol: a long-acting oral neuroleptic. A controlled study. *J Clin Pharmacol* 1975;15**:611-21.

*Leff JP, Wing JK. Trial of maintenance therapy in schizophrenia. *BMJ* 1971;**3**:599-604.

Levine J, Schooler NR, Severe J et al. Discontinuation of oral and depot fluphenazine in schizophrenic patients after one year of continuous medication: a controlled study. *Adv Biochem Psychopharmacol* 1980;24**:483-93.

*Marjerrison G, Irvine D, Stewart CN et al. Withdrawal of long-term phenothiazines from chronically hospitalized psychiatric patients. *Can Psychiatr Assoc J* 1964;**9**:290-8.

*McCreadie RG, Wiles D, Grant S et al. The Scottish first episode schizophrenia study. *Acta Psychiatr Scand* 1989;**80**:597-602.

*Melnyk WT, Worthington AG, Laverty SG. Abrupt withdrawal of chlorpromazine and thioridazine from schizophrenic in-patients. *Can Psychiatr Assoc J* 1966;**11**:410-3.

Morton MR. A study of the withdrawal of chlorpromazine or trifluoperazine in chronic schizophrenia. *Am J Psychiatry* 1968;**124**:1585-8.

Nishikawa T, Tsuda A, Tanaka M, Koga I, Uchida Y. Prophylactic effect of neuroleptics in symptomfree schizophrenia. *Psychopharmacology* 1982;**77**:301-4.

*Nishikawa T, Tsuda A, Tanaka M et al. Prophylactic effect of neuroleptics in symptom-free schizophrenics: a comparative dose response study of haloperidol and propericiazine. *Psychopharmacology* 1984;**82**:153-6.

*Odejide OA, Aderounmu AF. Double-blind placebo substitution: withdrawal of fluphenazine decanoate in schizophrenic patients. *J Clin Psychiatry* 1982;**43**:195-6.

Olson GW, Peterson DB. Sudden removal of tranquilizing drugs from chronic psychiatric patients. *J Nerv Ment Dis* 1960;**131**:252-5.

*Ota KY, Kurland AA. A double-blind comparison of haloperidol oral concentrate, haloperidol solutabs and placebo in the treatment of chronic schizophrenia. *J Clin Pharmacol New Drugs* 1973;**13**:99-110.

Peuskens J, Trivedi J, Malyarov S et al. Prevention of schizophrenia relapse with extended release quetiapine fumarate dosed once daily: a randomized, placebo-controlled trial in clinically stable patients. *Psychiatry* 2007;4**:34-50.

Pietzcker A, Gaebel W, Koepcke W et al. Intermittent versus maintenance neuroleptic long-termtreatment in schizophrenia - 2-year results of a German multicenter study. *J Psychiatr Res* 1993;**27**:321-39.
*Pigott TA, Carson WH, Saha AR et al. Aripiprazole for the prevention of relapse in stabilized patients with chronic schizophrenia: a placebo-controlled 26-week study. *J Clin Psychiatry* 2003;**64**:1048-56.

*Prien RF, Cole JO, Belkin NF. Relapse in chronic schizophrenics following abrupt withdrawal of tranquillizing medication. *Br J Psychiatry* 1968;**115**:679-86.

Prien RF, Levine J, Cole JO. High dose trifluoperazine therapy in chronic schizophrenia. *Am J Psychiatry* 1969;**126**:305-13.

Rifkin A, Quitkin F, Kane J, Klein DF, Ross D. The effect of fluphenazine upon social and vocational functioning in remitted schizophrenics. *Biol Psychiatry* 1979;**14**:499-508.

*Roelofs GA. Penfluridol (R 16341) as a maintenance therapy in chronic psychotic patients: a doubleblind clinical evaluation. *Acta Psychiatr Scand* 1974;**50**:219-24.
54. Ruskin PE, Nyman G. Discontinuation of neuroleptic medication in older, outpatient schizophrenics. A placebo-controlled, double-blind trial. *J Nerv Ment Dis* 1991;**179**:212-4.

Sampath G, Shah A, Krska J, Soni SD. Neuroleptic discontinuation in the very stable schizophrenic patient - relapse rates and serum neuroleptic levels. *Hum Psychopharmacol* 1992;7**:255-64.

**Kane JM, Mackle M, Snow-Adami L et al. A randomized placebo-controlled trial of asenapine for the
prevention of relapse of schizophrenia after long-term treatment. *J Clin Psychiatry* 2011;**72**:349-55.

*Schiele BC, Vestre ND, Stein KE. A comparison of thioridazine, trifluoperazine, chlorpromazine, and placebo: a double-blind controlled study on the treatment of chronic,

hospitalized, schizophrenic patients. *J Clin Exp Psychopathol Q Rev Psychiatry Neurol* 1961;**22**:151-62.

*Shawver JR, Gorham DR, Leskin LW, Good WW, Kabnick DE. Comparison of chlorpromazine and
reserpine in maintenance drug therapy. *Dis Nerv Syst* 1959;**20**:452-7.
Spohn HE, Coyne L, Larson J et al. Episodic and residual thought pathology in chronic schizophrenics: effect of neuroleptics. *Schizophr Bull* 1986;**12**:394-407.

Troshinsky C, Aaronson HG, Stone RK. Maintenance phenothiazines in aftercare of schizophrenic patients. *Pennsylvania Psychiatric Bulletin* 1962;**2**:11-5.

Vandecasteele AJ, Vereecken JL. A double-blind clinical evaluation of penfluridol (R 16 341) as a maintenance therapy in schizophrenia. *Acta Psychiatr Scand* 1974;50**:346-53.

*Whittaker CB, Hoy RM. Withdrawal of perphenazine in chronic schizophrenia. *Br J Psychiatry* 1963;**109**:422-7.

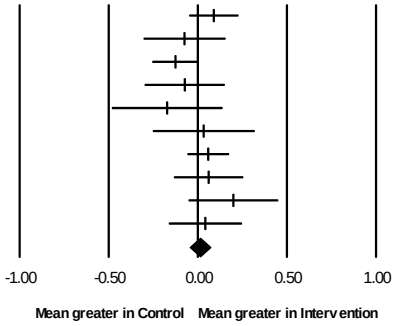
Wistedt B. A depot neuroleptic withdrawal study. A controlled study of the clinical effects of the withdrawal of depot fluphenazine decanoate and depot flupenthixol decanoate in chronic schizophrenic patients. *Acta Psychiatr Scand* 1981;64**:65-84.

Zissis NP, Psaras M, Lyketsos G. Haloperidol decanoate, a new long-acting antipsychotic, in chronic schizophrenics: double-blind comparison with placebo. *Curr Ther Res* 1982;31**:650-5.

Appendix B: Forest plots for the meta-analysis of baseline age for each of the 12 Systematic Reviews

Orrow et al 2012

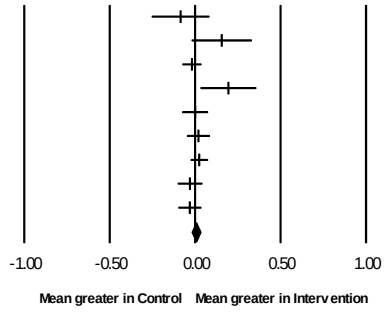
Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Acitivity Counselling trial 2001	0.090	0.072	0.005	-0.051	0.230	1.252	0.210
Chambers 2000	-0.075	0.119	0.014	-0.307	0.158	-0.628	0.530
Elley 2003	-0.126	0.068	0.005	-0.258	0.007	-1.858	0.063
Halbert 2000	-0.074	0.116	0.013	-0.300	0.153	-0.635	0.525
Jimmy 2004	-0.172	0.160	0.025	-0.484	0.141	-1.077	0.282
Kolt 2007	0.033	0.147	0.022	-0.254	0.320	0.225	0.822
Lawton 2009	0.058	0.061	0.004	-0.060	0.177	0.963	0.335
Morey 2009	0.061	0.100	0.010	-0.136	0.257	0.604	0.546
Petrella 2003	0.200	0.130	0.017	-0.055	0.454	1.539	0.124
Slujis 2005	0.042	0.106	0.011	-0.165	0.250	0.399	0.690
	0.010	0.029	0.001	-0.048	0.067	0.338	0.736



Q=10.7, p=0.29, I-squared=16.18

Neumann et al 2012

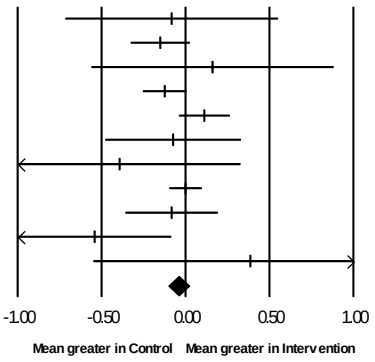
Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Eriksson 2006	-0.085	0.088	0.008	-0.256	0.087	-0.970	0.332
Eriksson 2007	0.156	0.091	0.008	-0.022	0.334	1.714	0.086
Eriksson 2008	-0.019	0.030	0.001	-0.077	0.039	-0.633	0.527
Eriksson 2010	0.195	0.085	0.007	0.029	0.360	2.303	0.021
Lassen 2008	0.000	0.040	0.002	-0.078	0.078	0.000	1.000
Lassen 2009	0.019	0.035	0.001	-0.050	0.089	0.549	0.583
Lassen 2010	0.024	0.027	0.001	-0.029	0.078	0.900	0.368
Lassen 2010 (ADVANCE 2)	-0.030	0.038	0.001	-0.105	0.045	-0.791	0.429
Turpie 2009	-0.031	0.036	0.001	-0.101	0.039	-0.868	0.386
	0.003	0.013	0.000	-0.023	0.029	0.226	0.821



Q=12.0, p=0.15, I-squared=33.46

Palmer et al 2012

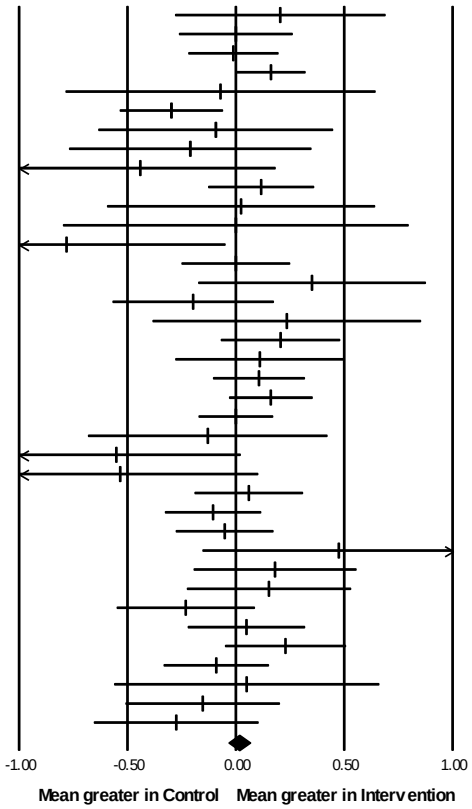
Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Abdul-Rahman 2007	-0.082	0.325	0.105	-0.718	0.554	-0.253	0.800
Brener 1998	-0.150	0.091	0.008	-0.329	0.029	-1.645	0.100
Cheng 1998	0.161	0.369	0.136	-0.563	0.884	0.435	0.664
Dember 2008	-0.124	0.068	0.005	-0.256	0.009	-1.830	0.067
Dixon 2009	0.112	0.079	0.006	-0.042	0.266	1.432	0.152
Ghorbani 2009	-0.074	0.207	0.043	-0.481	0.333	-0.357	0.721
Giustina 1998	-0.392	0.369	0.136	-1.115	0.330	-1.064	0.287
Januzzi 2002	0.000	0.050	0.003	-0.099	0.099	0.000	1.000
Kaufman 2003	-0.083	0.142	0.020	-0.360	0.195	-0.585	0.559
Sreedhara 1994	-0.542	0.234	0.055	-1.001	-0.082	-2.311	0.021
Zauner 1994	0.387	0.479	0.229	-0.552	1.325	0.808	0.419
	-0.043	0.031	0.001	-0.104	0.019	-1.364	0.173



Q=14.1, p=0.17, I-squared=29.03

Rutjes et al 2012

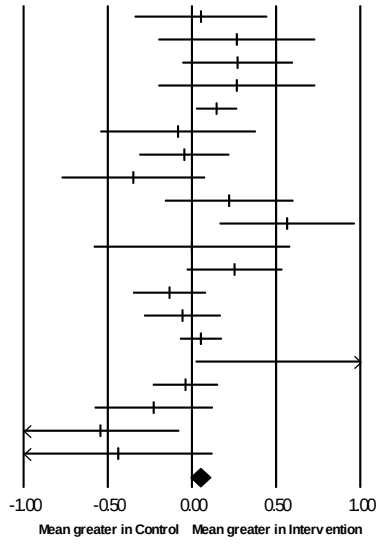
Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Adams 1995	0.205	0.249	0.062	-0.283	0.693	0.824	0.410
Altman 1998	0.000	0.135	0.018	-0.265	0.265	0.000	1.000
Altman 2004	-0.011	0.108	0.012	-0.222	0.199	-0.107	0.915
Altman 2009	0.162	0.083	0.007	-0.000	0.324	1.955	0.051
Atay 2008	-0.070	0.366	0.134	-0.788	0.647	-0.193	0.847
Baltzer 2008	-0.298	0.123	0.015	-0.538	-0.057	-2.426	0.015
Blanco 2008	-0.093	0.277	0.077	-0.637	0.451	-0.335	0.738
Bunyaratavej 2001	-0.210	0.287	0.082	-0.772	0.351	-0.734	0.463
Carrabba 1995	-0.441	0.320	0.102	-1.068	0.186	-1.378	0.168
Chevalier 2010	0.117	0.126	0.016	-0.130	0.364	0.929	0.353
Corrado 1995	0.024	0.317	0.100	-0.597	0.645	0.076	0.940
Creamer 1994	0.000	0.408	0.167	-0.800	0.800	0.000	1.000
Cubukcu 2005	-0.781	0.375	0.141	-1.517	-0.046	-2.081	0.037
Day 2003	0.000	0.129	0.017	-0.253	0.253	0.000	1.000
Diracoglu 2009	0.352	0.269	0.072	-0.176	0.879	1.306	0.191
Dougados 1993	-0.197	0.191	0.037	-0.572	0.178	-1.030	0.303
Formiguera Sala 1995	0.235	0.317	0.101	-0.387	0.857	0.740	0.459
Huang 2011	0.206	0.142	0.020	-0.072	0.484	1.453	0.146
Huskisson 1999	0.110	0.200	0.040	-0.282	0.503	0.552	0.581
Jorgensen 2010	0.107	0.109	0.012	-0.108	0.321	0.975	0.329
Judd 2003	0.161	0.099	0.010	-0.033	0.356	1.625	0.104
Kahan 2001	0.000	0.089	0.008	-0.174	0.174	0.000	1.000
Kotevoglu 2006	-0.130	0.283	0.080	-0.684	0.425	-0.458	0.647
Kul-Panza 2010	-0.552	0.294	0.087	-1.129	0.025	-1.874	0.061
Listrat 1997	-0.533	0.326	0.106	-1.172	0.106	-1.635	0.102
Lohmander 1996	0.060	0.129	0.017	-0.193	0.313	0.462	0.644
Navarro-Sarabia 2011	-0.105	0.114	0.013	-0.329	0.119	-0.919	0.358
Neustadt 2005	-0.052	0.116	0.013	-0.279	0.176	-0.446	0.656
Payne 2000	0.475	0.322	0.104	-0.156	1.107	1.475	0.140
Petrella 2002	0.181	0.193	0.037	-0.197	0.559	0.938	0.348
Petrella 2006	0.153	0.195	0.038	-0.228	0.534	0.786	0.432
Petrella b 2008	-0.231	0.164	0.027	-0.552	0.090	-1.412	0.158
Pham 2004	0.049	0.139	0.019	-0.224	0.322	0.351	0.725
Puhl 1993	0.229	0.144	0.021	-0.052	0.511	1.595	0.111
Raynauld 2002	-0.090	0.125	0.016	-0.336	0.155	-0.721	0.471
Sezgin 2004	0.050	0.313	0.098	-0.564	0.664	0.160	0.873
Tetik 2003	-0.153	0.183	0.033	-0.511	0.205	-0.837	0.403
Wobig 1998	-0.275	0.195	0.038	-0.657	0.108	-1.408	0.159
	0.013	0.025	0.001	-0.037	0.062	0.502	0.616



Q=46.5, p=0.14, I-squared=20.39

Thangaratinam et al 2012

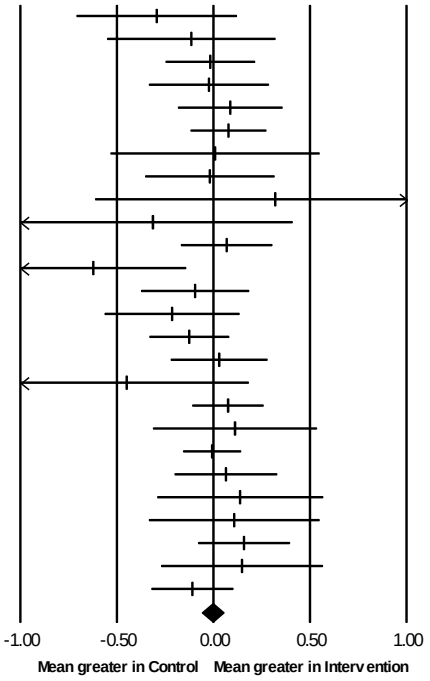
Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Asbee 2009	0.054	0.202	0.041	-0.342	0.450	0.265	0.791
Baciuk 2008	0.266	0.239	0.057	-0.202	0.734	1.115	0.265
Barakat 2009	0.271	0.169	0.028	-0.059	0.602	1.608	0.108
Cavalcante 2009	0.266	0.239	0.057	-0.202	0.734	1.115	0.265
Crowther 2005	0.147	0.063	0.004	0.023	0.271	2.317	0.021
Erkkola and Makela 1976	-0.083	0.237	0.056	-0.546	0.381	-0.349	0.727
Garshasbi 2005	-0.045	0.137	0.019	-0.314	0.224	-0.328	0.743
Guelinckx 2010	-0.348	0.219	0.048	-0.776	0.081	-1.591	0.112
Haakstad 2011	0.221	0.196	0.038	-0.163	0.605	1.130	0.259
Hopkins 2010	0.566	0.206	0.042	0.162	0.969	2.746	0.006
Hui 2006	0.000	0.299	0.089	-0.586	0.586	0.000	1.000
Hui 2011	0.253	0.146	0.021	-0.033	0.539	1.732	0.083
Jackson 2011	-0.133	0.112	0.012	-0.352	0.086	-1.192	0.233
Khoury 2005	-0.056	0.118	0.014	-0.287	0.174	-0.480	0.632
Landon 2009	0.053	0.065	0.004	-0.074	0.180	0.821	0.411
Marquez-Sterling 1998	1.129	0.566	0.320	0.020	2.238	1.995	0.046
Phelan 2011	-0.038	0.100	0.010	-0.234	0.157	-0.385	0.700
Quinlivan 2011	-0.227	0.180	0.032	-0.581	0.126	-1.262	0.207
Santos 2005	-0.544	0.240	0.058	-1.014	-0.073	-2.265	0.024
Wolff 2008	-0.438	0.287	0.082	-1.000	0.125	-1.525	0.127
	0.048	0.030	0.001	-0.011	0.107	1.583	0.113



Q=38.1, p=0.01, I-squared=50.11

Hempel et al 2012

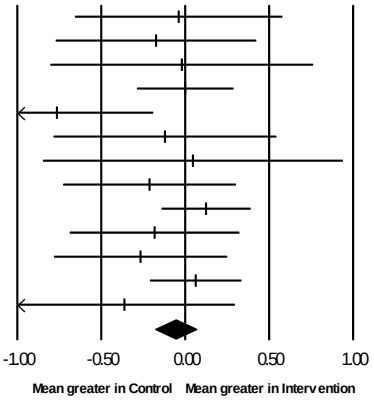
Studyname	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Beausoleil 2007	-0.294	0.213	0.045	-0.712	0.124	-1.378	0.168
Borgia 1982	-0.115	0.224	0.050	-0.553	0.324	-0.512	0.609
Conway 2007	-0.016	0.120	0.014	-0.251	0.218	-0.137	0.891
Correa 2005	-0.023	0.160	0.025	-0.336	0.290	-0.147	0.884
de Bortoli 2007	0.087	0.139	0.019	-0.186	0.361	0.627	0.531
Duman 2005	0.078	0.102	0.010	-0.122	0.277	0.763	0.445
Felley 2001	0.008	0.278	0.077	-0.536	0.552	0.030	0.976
Hickson 2007	-0.018	0.172	0.030	-0.356	0.319	-0.107	0.914
Jirapinyo 2002	0.320	0.477	0.228	-0.616	1.256	0.670	0.503
Koning 2008	-0.314	0.371	0.137	-1.040	0.413	-0.846	0.397
Kotowska 2005	0.069	0.122	0.015	-0.171	0.308	0.562	0.574
Lewis 1998	-0.622	0.247	0.061	-1.106	-0.139	-2.522	0.012
McFarland 1995	-0.095	0.144	0.021	-0.377	0.187	-0.659	0.510
Merenstein 2009	-0.214	0.179	0.032	-0.566	0.138	-1.193	0.233
Park 2007	-0.125	0.107	0.011	-0.334	0.084	-1.174	0.240
Ruszczynski 2008	0.030	0.129	0.017	-0.224	0.283	0.229	0.819
Safdar 2008	-0.449	0.324	0.105	-1.084	0.185	-1.387	0.165
Sampalis 2010	0.075	0.096	0.009	-0.112	0.263	0.786	0.432
Schrezenmeir 2004	0.112	0.218	0.047	-0.315	0.539	0.513	0.608
Song 2010	-0.007	0.078	0.006	-0.159	0.146	-0.089	0.929
Song 2010	0.064	0.137	0.019	-0.204	0.333	0.471	0.638
Szajewska 2009	0.138	0.220	0.048	-0.293	0.570	0.628	0.530
Szymanski 2008	0.107	0.227	0.051	-0.337	0.552	0.474	0.635
Thomas 2001	0.158	0.123	0.015	-0.082	0.398	1.290	0.197
Wenus 2008	0.148	0.215	0.046	-0.274	0.570	0.688	0.491
Yoon 2011	-0.109	0.110	0.012	-0.324	0.106	-0.997	0.319
	-0.007	0.028	0.001	-0.062	0.049	-0.231	0.818



Q=21.1, p=0.69, I-squared=0.00

Hemmingsen et al 2012

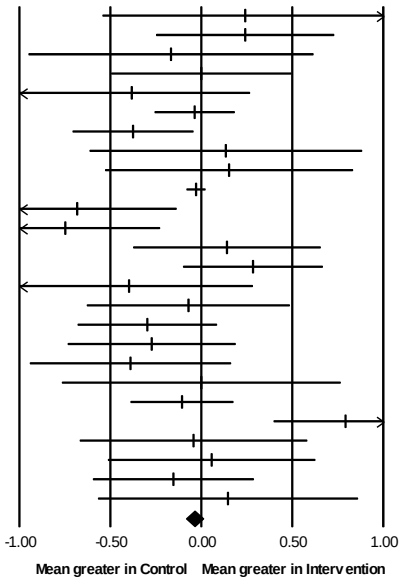
Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Altuntas et al (2003)	-0.039	0.316	0.100	-0.659	0.581	-0.123	0.902
Aviles-Santa et al (1999)	-0.174	0.306	0.093	-0.773	0.425	-0.569	0.569
Civera et al (2008)	-0.021	0.400	0.160	-0.805	0.764	-0.051	0.959
Doueket al (2005)	0.000	0.148	0.022	-0.290	0.290	0.000	1.000
Giugliano et al (1993)	-0.764	0.294	0.086	-1.340	-0.188	-2.600	0.009
Hermann et al (2001)	-0.121	0.340	0.115	-0.786	0.545	-0.356	0.722
Kabadi & Kabadi (2006)	0.046	0.456	0.208	-0.849	0.941	0.101	0.920
Kokic et al (2003)	-0.212	0.263	0.069	-0.729	0.304	-0.807	0.420
Kvapil et al (2006)	0.124	0.137	0.019	-0.143	0.392	0.909	0.363
Relimpio et al (1998)	-0.182	0.259	0.067	-0.690	0.325	-0.704	0.481
Strowig et al (2002)	-0.266	0.264	0.070	-0.784	0.252	-1.006	0.314
Ushakova et al (2007)	0.062	0.140	0.020	-0.212	0.337	0.444	0.657
Yilmaz et al (2007)	-0.362	0.337	0.113	-1.022	0.298	-1.075	0.282
	-0.059	0.063	0.004	-0.182	0.064	-0.940	0.347



Q=10.7, p=0.56, I-squared=0.00

Umpierre et al 2011

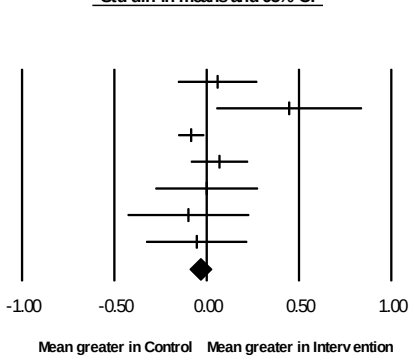
Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Aas, 2005	0.242	0.402	0.161	-0.546	1.029	0.601	0.548
Agurs-Collins, 1997	0.241	0.251	0.063	-0.250	0.733	0.962	0.336
Brun, 2008	-0.167	0.401	0.161	-0.953	0.619	-0.416	0.677
Casteneda, 2002	0.000	0.254	0.065	-0.498	0.498	0.000	1.000
Cheung, 2008	-0.382	0.333	0.111	-1.035	0.270	-1.148	0.251
Christian, 2008	-0.036	0.114	0.013	-0.259	0.186	-0.321	0.748
Church, 2010	-0.375	0.171	0.029	-0.710	-0.041	-2.197	0.028
Cuff, 2003	0.135	0.383	0.147	-0.617	0.887	0.352	0.725
Diedrich, 2009	0.153	0.349	0.122	-0.531	0.836	0.437	0.662
Jakicic, 209	-0.029	0.028	0.001	-0.084	0.025	-1.047	0.295
Kadoglou, 2007 (Exercise reduces resistin…)	-0.682	0.280	0.078	-1.232	-0.133	-2.436	0.015
Kadoglou, 2007 (The anti-inflammatory…)	-0.748	0.267	0.071	-1.271	-0.224	-2.800	0.005
Kim, 2006	0.141	0.264	0.070	-0.377	0.659	0.533	0.594
Kirk, 2009	0.284	0.197	0.039	-0.103	0.671	1.438	0.150
Lambers, 2008	-0.397	0.348	0.121	-1.079	0.285	-1.140	0.254
Louimaala, 2003	-0.071	0.286	0.082	-0.631	0.489	-0.249	0.803
Mayer Davis, 2004	-0.297	0.197	0.039	-0.682	0.089	-1.510	0.131
Menard, 2005	-0.273	0.237	0.056	-0.737	0.191	-1.151	0.250
Middlebrooke, 2006	-0.389	0.283	0.080	-0.945	0.166	-1.374	0.169
Samaras, 1997	0.000	0.392	0.154	-0.769	0.769	0.000	1.000
Sigal, 2007	-0.106	0.146	0.021	-0.392	0.179	-0.729	0.466
Sridhar, 2010	0.793	0.203	0.041	0.395	1.191	3.907	0.000
Tessier, 2000	-0.043	0.320	0.103	-0.671	0.585	-0.133	0.894
Tudor-Locke, 2004	0.057	0.292	0.085	-0.515	0.629	0.195	0.846
Vanninen, 1992	-0.154	0.227	0.051	-0.598	0.291	-0.678	0.498
Wing, 1988	0.147	0.366	0.134	-0.570	0.863	0.402	0.688
	-0.040	0.024	0.001	-0.087	0.007	-1.657	0.098



Q=45.8, p=0.01, I-squared=45.46

Heneghan et al 2011

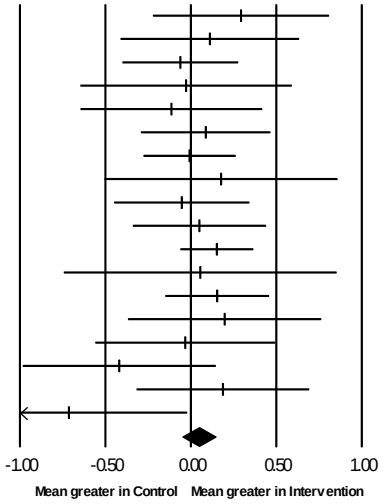
Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Beyth 2000	0.059	0.111	0.012	-0.158	0.277	0.534	0.593
Christensen 2006	0.447	0.202	0.041	0.050	0.844	2.206	0.027
Matcher 2010	-0.084	0.037	0.001	-0.156	-0.011	-2.263	0.024
Mendendez-Jandula 2005	0.069	0.080	0.006	-0.088	0.226	0.863	0.388
Siebenhofer 2008	0.000	0.143	0.021	-0.281	0.281	0.000	1.000
Sunderji 2004	-0.099	0.169	0.029	-0.430	0.233	-0.585	0.559
Voller 2005	-0.054	0.141	0.020	-0.330	0.222	-0.384	0.701
	-0.036	0.030	0.001	-0.095	0.022	-1.218	0.223



Q=10.0, p=0.12, I-squared=40.13

Coombes et al 2010

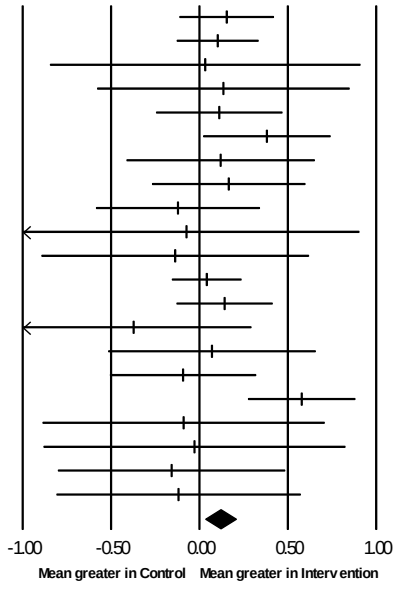
Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Alvarez 2005	0.293	0.264	0.070	-0.224	0.811	1.110	0.267
Alvarez-Nemegyei 2008	0.111	0.268	0.072	-0.413	0.636	0.415	0.678
Bisset 2006	-0.061	0.174	0.030	-0.403	0.280	-0.352	0.725
Blair 1996	-0.028	0.317	0.100	-0.649	0.592	-0.089	0.929
de Vos 2010	-0.114	0.272	0.074	-0.648	0.420	-0.418	0.676
Ekeberg 2009	0.087	0.194	0.038	-0.294	0.468	0.447	0.655
Hay 2003	-0.007	0.139	0.019	-0.280	0.265	-0.053	0.958
Holsrud 2006	0.177	0.349	0.122	-0.507	0.861	0.506	0.613
McInerney 2003	-0.053	0.203	0.041	-0.451	0.345	-0.260	0.795
Peerbooms 2010	0.050	0.200	0.040	-0.342	0.442	0.250	0.803
Petrella 2010	0.152	0.110	0.012	-0.064	0.368	1.382	0.167
Scarpone 2008	0.055	0.408	0.167	-0.745	0.855	0.135	0.893
Smidt 2002	0.154	0.156	0.024	-0.152	0.459	0.984	0.325
Tonks 2006	0.197	0.289	0.084	-0.370	0.765	0.682	0.495
Vecchio 1993	-0.033	0.270	0.073	-0.562	0.496	-0.122	0.903
Willberg 2008	-0.419	0.289	0.084	-0.986	0.148	-1.448	0.148
Wong 2005	0.187	0.259	0.067	-0.320	0.695	0.724	0.469
Zeisig 2008	-0.714	0.354	0.126	-1.408	-0.019	-2.014	0.044
	0.045	0.049	0.002	-0.052	0.142	0.911	0.362



Q=11.5, p=0.83, I-squared=0.00

Leucht et al 2012

Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Arato 2002	0.154	0.138	0.019	-0.116	0.424	1.120	0.263
Beasley 2003	0.103	0.120	0.014	-0.131	0.337	0.864	0.388
Boonstra 2011	0.033	0.449	0.202	-0.848	0.914	0.073	0.942
Clark 1975	0.136	0.365	0.134	-0.581	0.852	0.371	0.711
Cooper 2000	0.112	0.184	0.034	-0.247	0.472	0.613	0.540
Crow 1986	0.381	0.185	0.034	0.018	0.744	2.060	0.039
Gross 1974	0.120	0.273	0.075	-0.415	0.655	0.440	0.660
Hirsch 1973	0.166	0.223	0.050	-0.270	0.602	0.745	0.456
Hirsch 1996	-0.122	0.238	0.056	-0.587	0.344	-0.512	0.609
Kane 1979	-0.073	0.500	0.250	-1.054	0.907	-0.146	0.884
Kane 1982	-0.137	0.387	0.150	-0.896	0.622	-0.354	0.724
Kane 2011	0.041	0.102	0.010	-0.158	0.241	0.407	0.684
Kramer 2007	0.142	0.140	0.020	-0.132	0.416	1.016	0.310
Kurland 1975	-0.372	0.341	0.116	-1.041	0.297	-1.090	0.276
Levine 1980	0.071	0.301	0.090	-0.519	0.660	0.235	0.814
Nishikawa 1984	-0.093	0.212	0.045	-0.509	0.323	-0.437	0.662
Peuskens 2007	0.578	0.156	0.024	0.272	0.884	3.705	0.000
Sampath 1992	-0.090	0.408	0.167	-0.890	0.711	-0.219	0.826
Vandecasteele 1974	-0.028	0.437	0.191	-0.884	0.828	-0.064	0.949
Wistedt 1981	-0.158	0.329	0.108	-0.803	0.487	-0.479	0.632
Zissis 1982	-0.118	0.354	0.125	-0.812	0.576	-0.334	0.739
	0.117	0.044	0.002	0.030	0.203	2.645	0.008



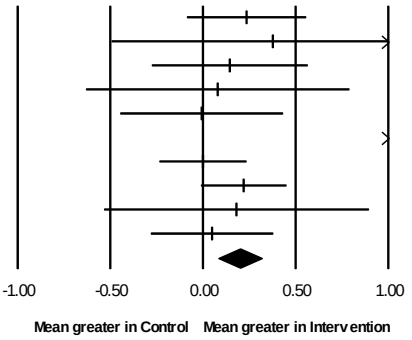
Q=17.7, p=0.61, I-squared=0.00

Anothaisintawee et al 2012

Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Alexander 2004	0.235	0.165	0.027	-0.089	0.559	1.421	0.155
Bates 2007	0.377	0.445	0.198	-0.495	1.248	0.847	0.397
Cheah 2003	0.145	0.216	0.047	-0.278	0.568	0.671	0.502
Leskinen 1999	0.080	0.364	0.132	-0.633	0.793	0.219	0.827
Nicke, Downey, Clark et al 2003	-0.007	0.225	0.051	-0.449	0.435	-0.030	0.976
Nickel, downey et al 2004	2.731	0.348	0.121	2.050	3.412	7.857	0.000
Nickel, Krieger et al 2008	0.000	0.121	0.015	-0.238	0.238	0.000	1.000
Pontari 2010	0.220	0.119	0.014	-0.013	0.452	1.850	0.064
Shoskes 1999	0.181	0.366	0.134	-0.536	0.898	0.495	0.621
Wagenlehner 2009	0.049	0.170	0.029	-0.284	0.381	0.288	0.774
	0.198	0.060	0.004	0.081	0.315	3.307	0.001

Statistics for each study

Std diff in means and 95% CI

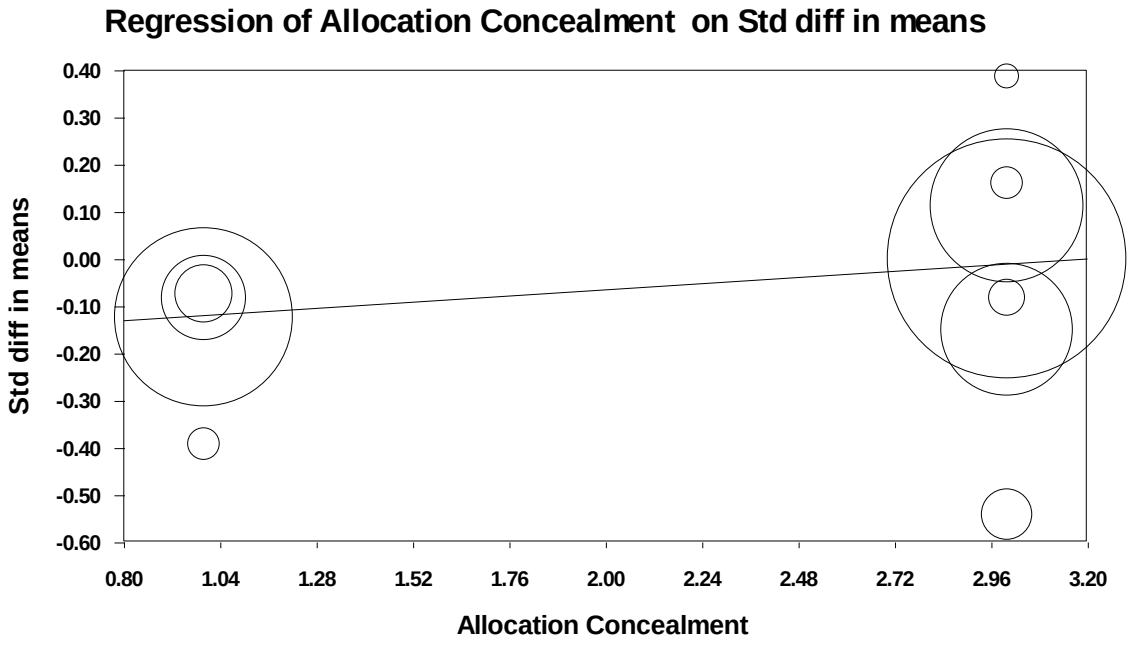


Q=57.8, p<0.001, I-squared=84.42

Appendix C. Bubble plots and meta-regression statistics for each of the 12 Systematic review.

For each bubble plot, the allocation concealments relate to: 1 Adequate, 2 Inadequate and 3 Unclear.

Palmer et al 2012

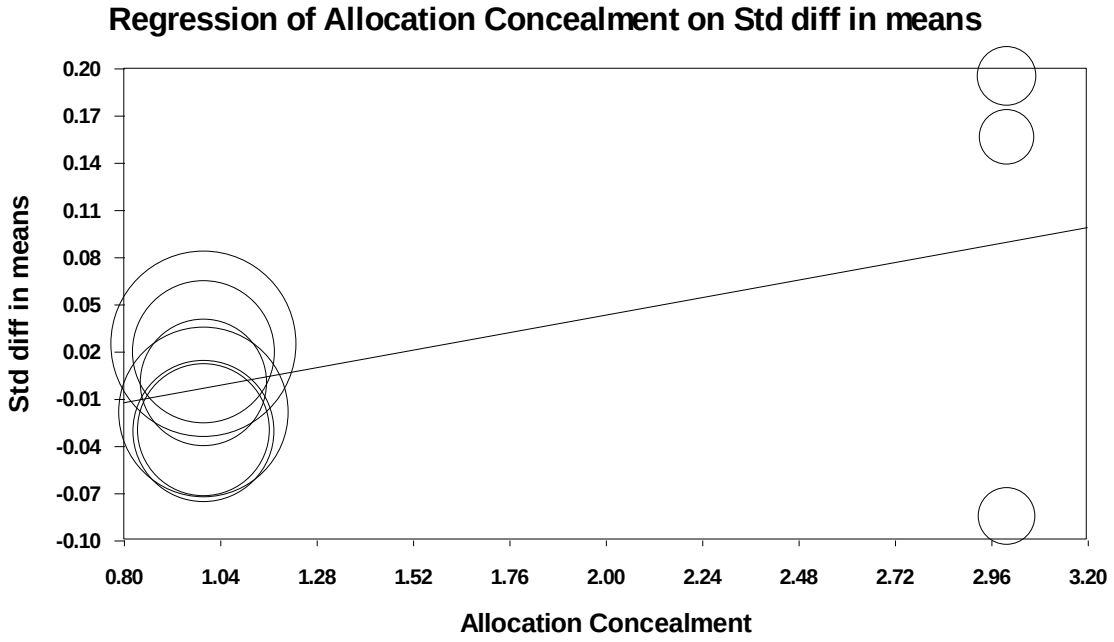


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	0.05449	0.03443	-0.01299	0.12197	1.58263	0.11350
Intercept	-0.17412	0.08870	-0.34797	-0.00027	-1.96300	0.04965
Tau-squared	0.00454					

	Q	df	p-value
Model	2.50473	1.00000	0.11350
Residual	11.58557	9.00000	0.23769
Total	14.09030	10.00000	0.16891

Neumann et al 2012

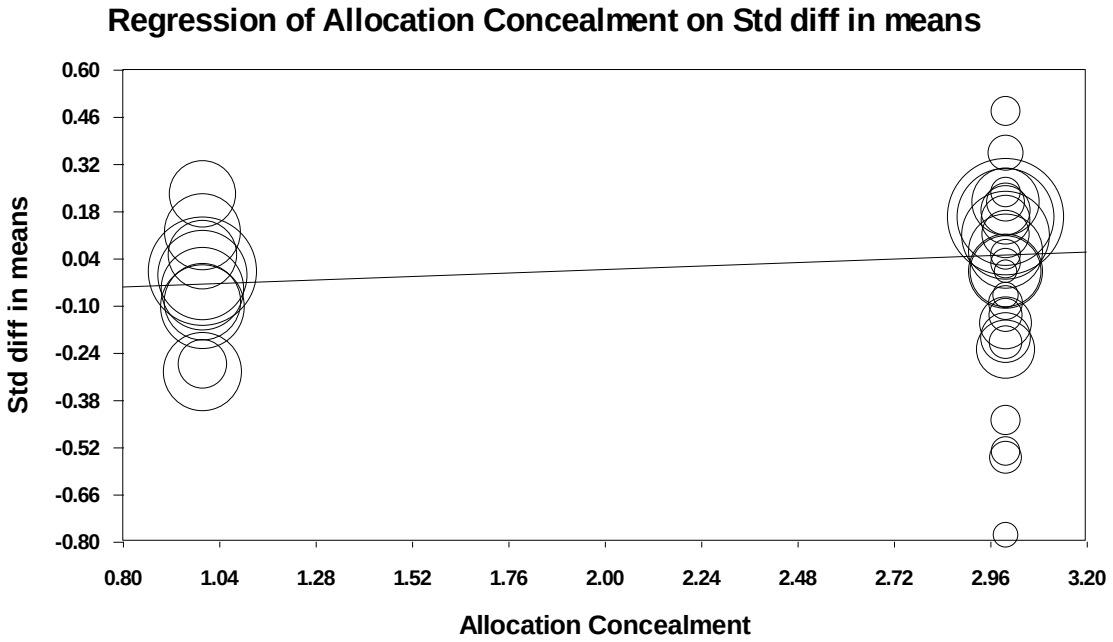


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	0.04638	0.02618	-0.00493	0.09770	1.77165	0.07645
Intercept	-0.04968	0.03250	-0.11338	0.01401	-1.52878	0.12632
Tau-squared	0.00040					

	Q	df	p-value
Model	3.13874	1.00000	0.07645
Residual	8.88328	7.00000	0.26115
Total	12.02202	8.00000	0.15022

Rutjes et al 2012

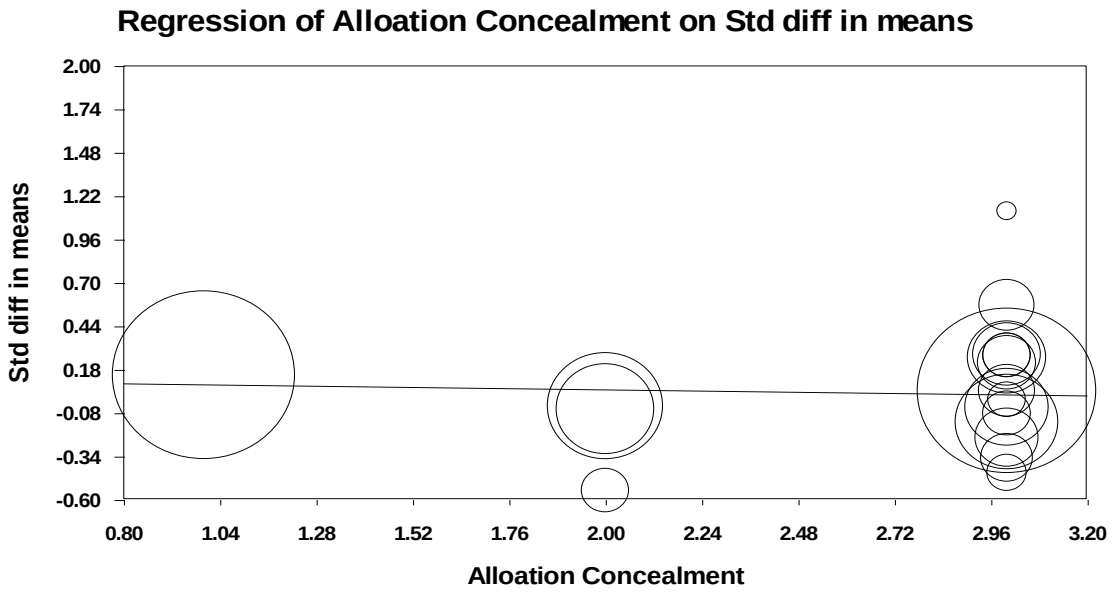


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	0.04309	0.02545	-0.00680	0.09297	1.69294	0.09047
Intercept	-0.07912	0.05980	-0.19633	0.03808	-1.32313	0.18579
Tau-squared	0.00533					

	Q	df	p-value
Model	2.86604	1.00000	0.09047
Residual	43.60935	36.00000	0.17937
Total	46.47539	37.00000	0.13665

Thangaratinam et al 2012

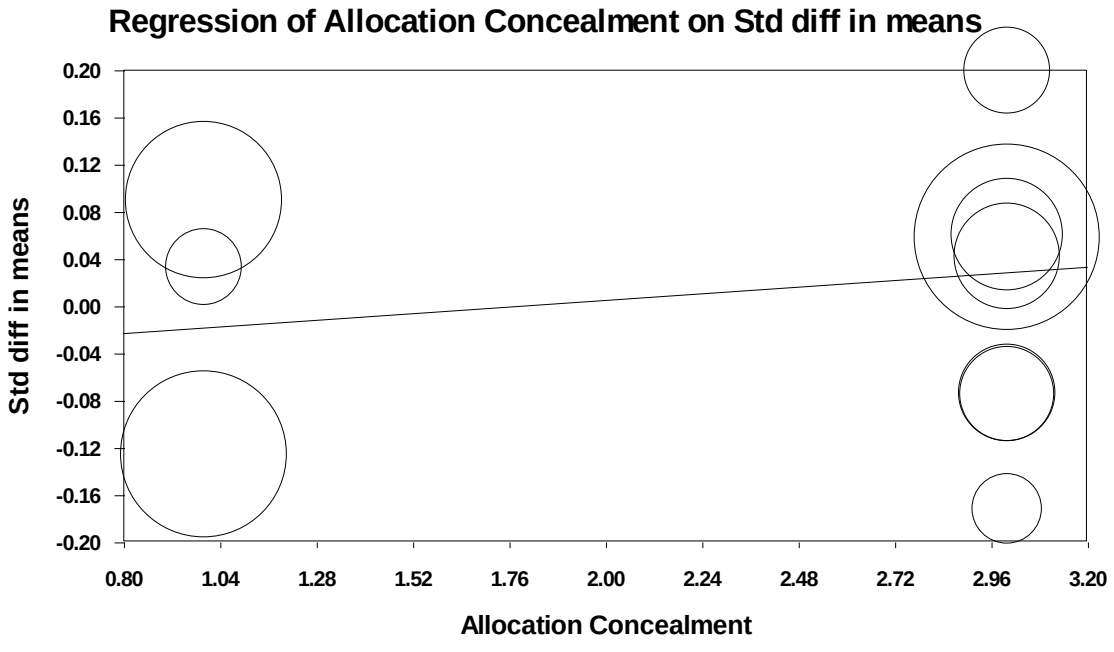


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	-0.03032	0.03638	-0.10162	0.04098	-0.83337	0.40463
Intercept	0.11971	0.09146	-0.05956	0.29897	1.30881	0.19060
Tau-squared	0.02538					

	Q	df	p-value
Model	0.69451	1.00000	0.40463
Residual	37.38887	18.00000	0.00466
Total	38.08338	19.00000	0.00579

Orrow et al 2012

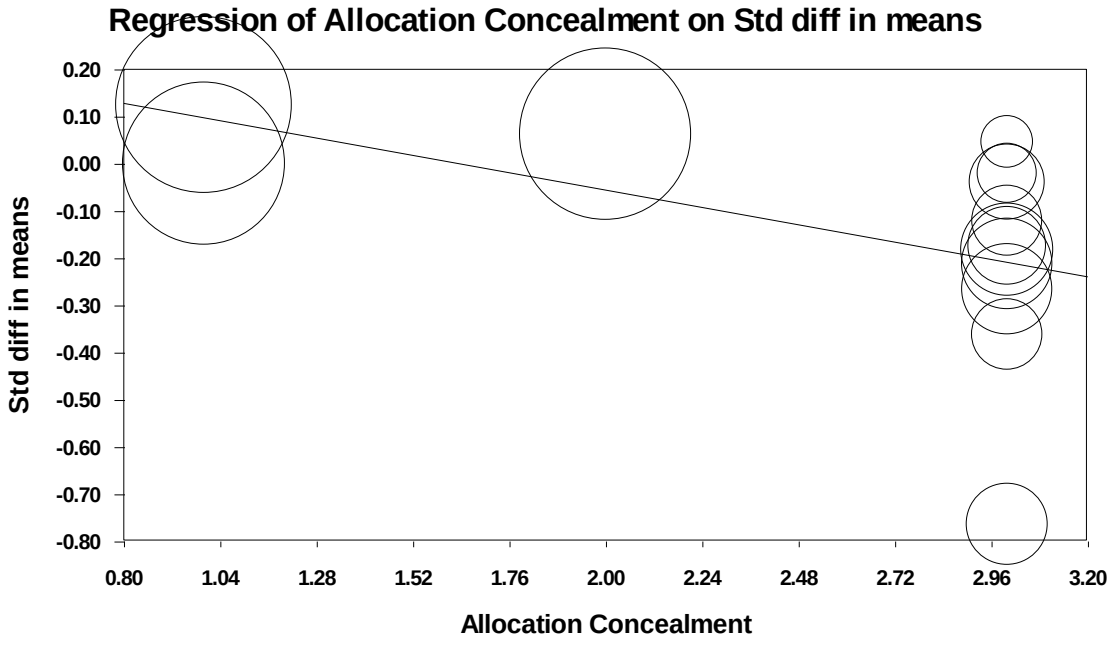


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	0.02344	0.02997	-0.03530	0.08219	0.78226	0.43406
Intercept	-0.04195	0.07245	-0.18395	0.10005	-0.57897	0.56261
Tau-squared	0.00259					

	Q	df	p-value
Model	0.61193	1.00000	0.43406
Residual	10.12493	8.00000	0.25637
Total	10.73686	9.00000	0.29418

Hemmingsen et al 2012

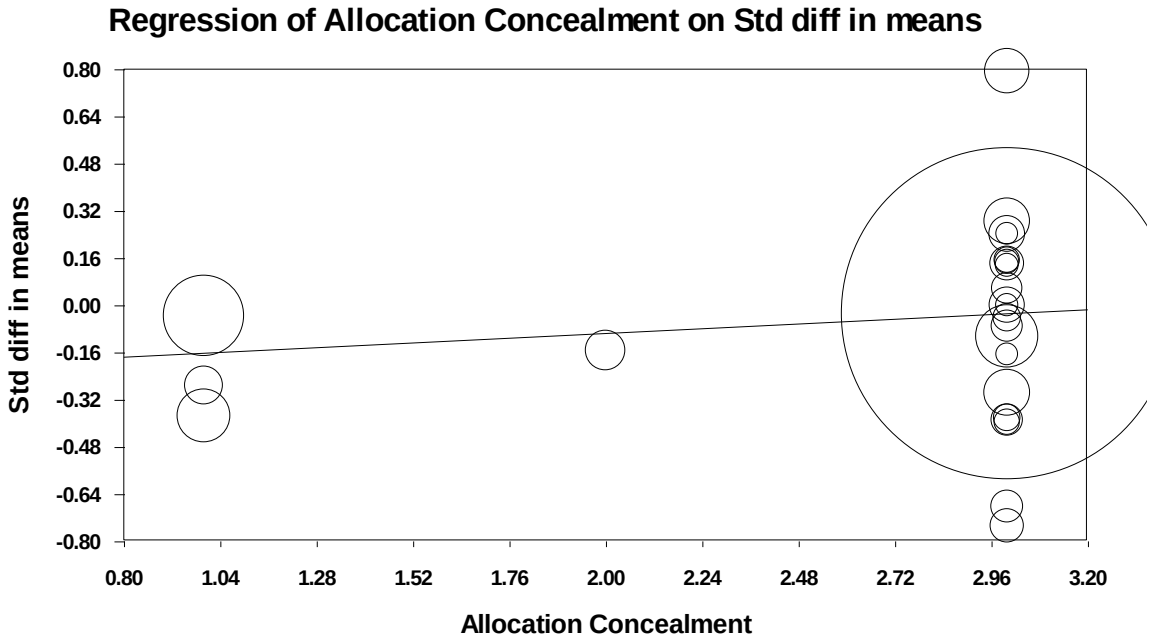


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	-0.15312	0.07006	-0.29043	-0.01581	-2.18571	0.02884
Intercept	0.25038	0.15475	-0.05292	0.55369	1.61797	0.10567
Tau-squared	0.00000					

	Q	df	p-value
Model	4.77731	1.00000	0.02884
Residual	5.91734	11.00000	0.87882
Total	10.69465	12.00000	0.55525

Umpierre et al 2011

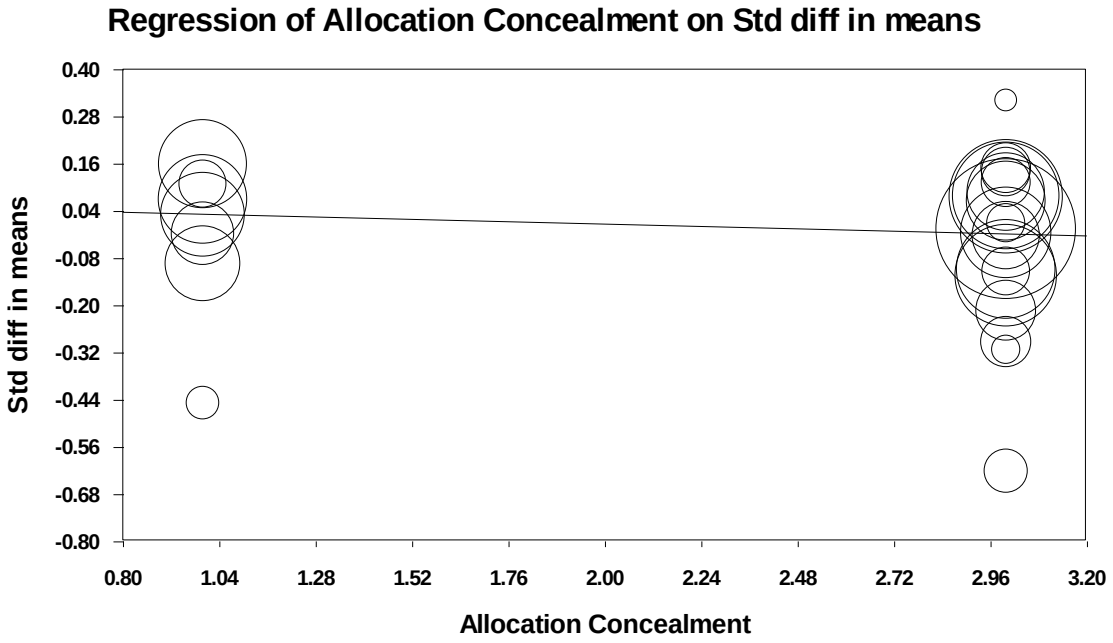


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	0.06693	0.04504	-0.02133	0.15520	1.48623	0.13722
Intercept	-0.22983	0.13011	-0.48484	0.02518	-1.76641	0.07733
Tau-squared	0.03089					

	Q	df	p-value
Model	2.20888	1.00000	0.13722
Residual	43.62677	24.00000	0.00843
Total	45.83566	25.00000	0.00671

Hempel et al 2012

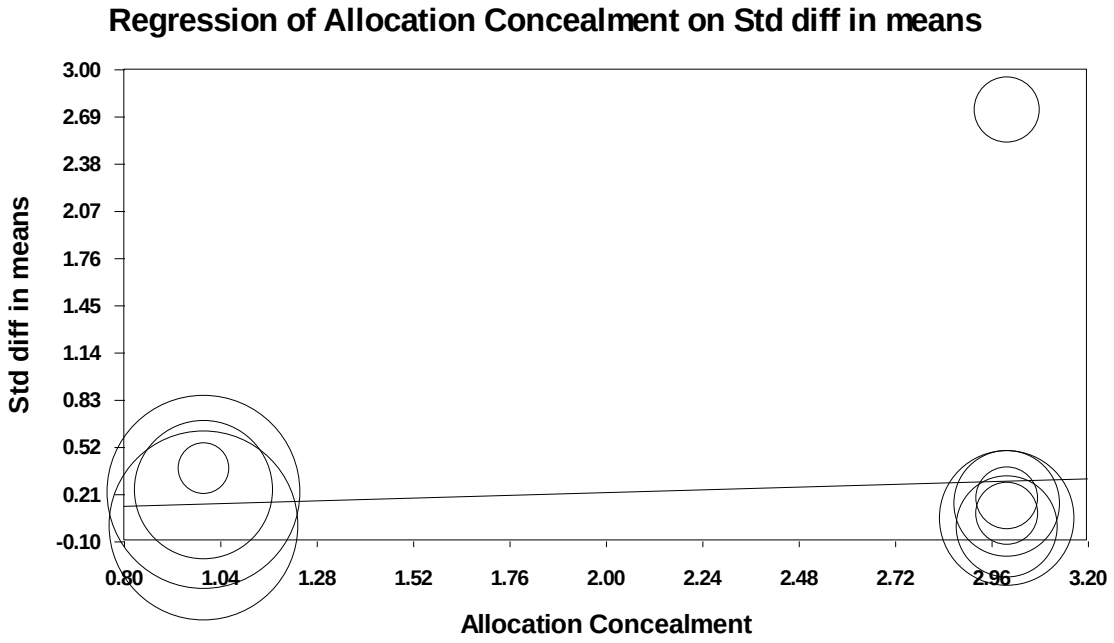


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	-0.02506	0.03299	-0.08972	0.03960	-0.75963	0.44748
Intercept	0.05630	0.08748	-0.11517	0.22776	0.64351	0.51989
Tau-squared	0.00000					

	Q	df	p-value
Model	0.57704	1.00000	0.44748
Residual	20.47321	24.00000	0.66957
Total	21.05025	25.00000	0.68980

Anothaisintawee et al 2012

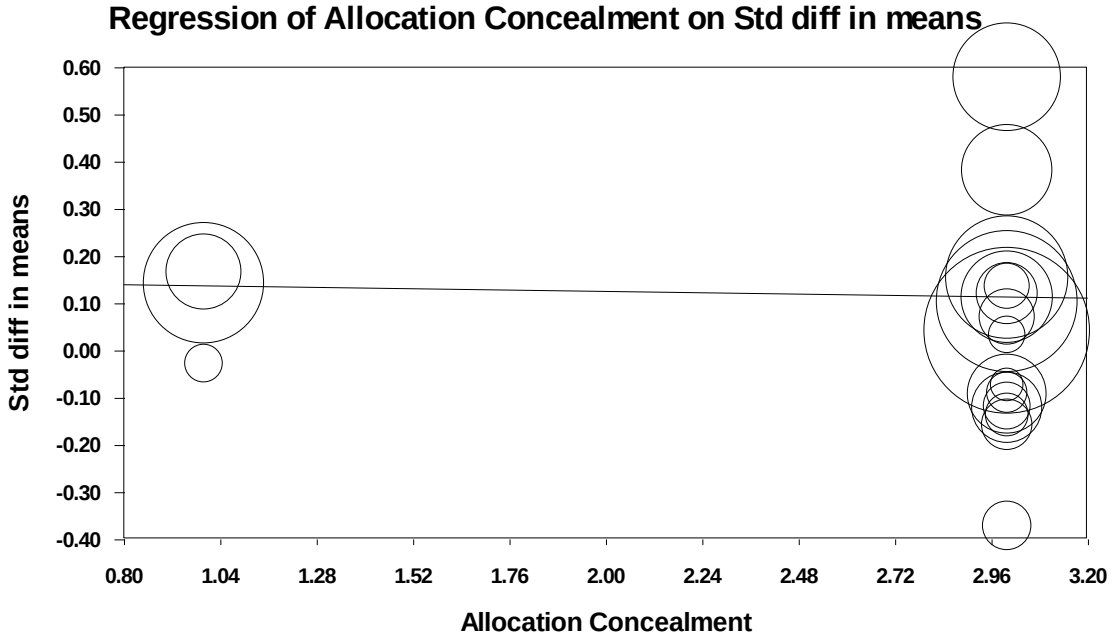


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	0.07510	0.06249	-0.04738	0.19759	1.20178	0.22945
Intercept	0.06936	0.12241	-0.17055	0.30928	0.56668	0.57093
Tau-squared	0.24641					

	Q	df	p-value
Model	1.44427	1.00000	0.22945
Residual	56.33335	8.00000	0.00000
Total	57.77762	9.00000	0.00000

Leucht et al 2012

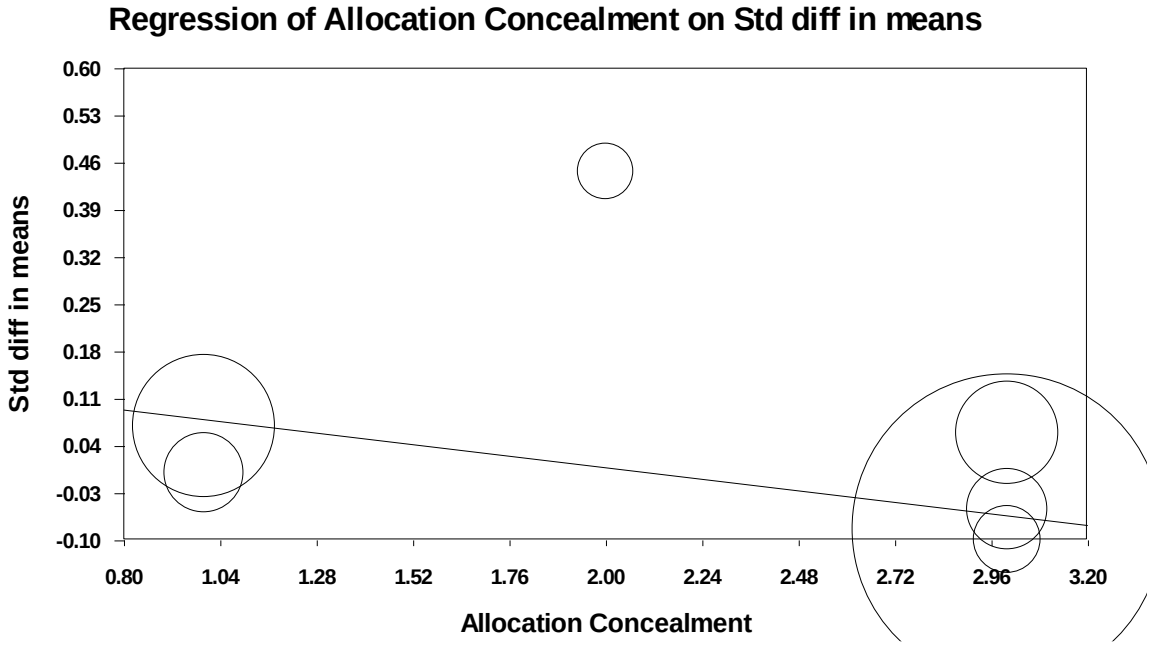


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	-0.01184	0.06194	-0.13325	0.10957	-0.19112	0.84843
Intercept	0.14858	0.17313	-0.19074	0.48790	0.85822	0.39077
Tau-squared	0.00000					

	Q	df	p-value
Model	0.03653	1.00000	0.84843
Residual	17.63371	19.00000	0.54702
Total	17.67023	20.00000	0.60912

Heneghan et al 2011

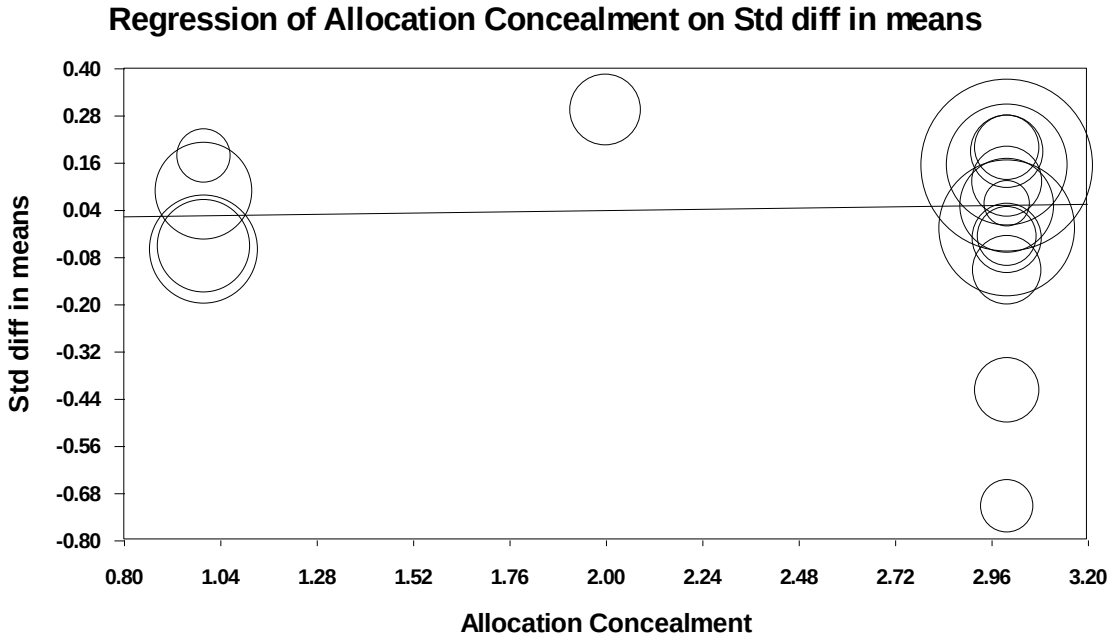


Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	-0.07133	0.03851	-0.14681	0.00414	-1.85240	0.06397
Intercept	0.15026	0.10504	-0.05562	0.35614	1.43048	0.15258
Tau-squared	0.00402					

	Q	df	p-value
Model	3.43140	1.00000	0.06397
Residual	6.59079	5.00000	0.25290
Total	10.02218	6.00000	0.12372

Coombes et al 2010



Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	0.01332	0.05910	-0.10252	0.12915	0.22532	0.82173
Intercept	0.01145	0.15675	-0.29577	0.31868	0.07306	0.94176
Tau-squared	0.00000					

	Q	df	p-value
Model	0.05077	1.00000	0.82173
Residual	11.47565	16.00000	0.77922
Total	11.52642	17.00000	0.82793

Appendix 4: paper 2

Important outcome predictors showed greater baseline heterogeneity than age in two systematic reviews

Appendices for paper 2:

Appendix 1a: Meta-regression bubble plot and corresponding statistics for hip protectors review for age (low versus high or unclear risk of bias)

Appendix 1b Meta-regression bubble plot and corresponding statistics for hip protectors review for body mass (low versus high or unclear risk of bias).

Appendix 1c: Meta-regression bubble plot and corresponding statistics for low back pain review for age.

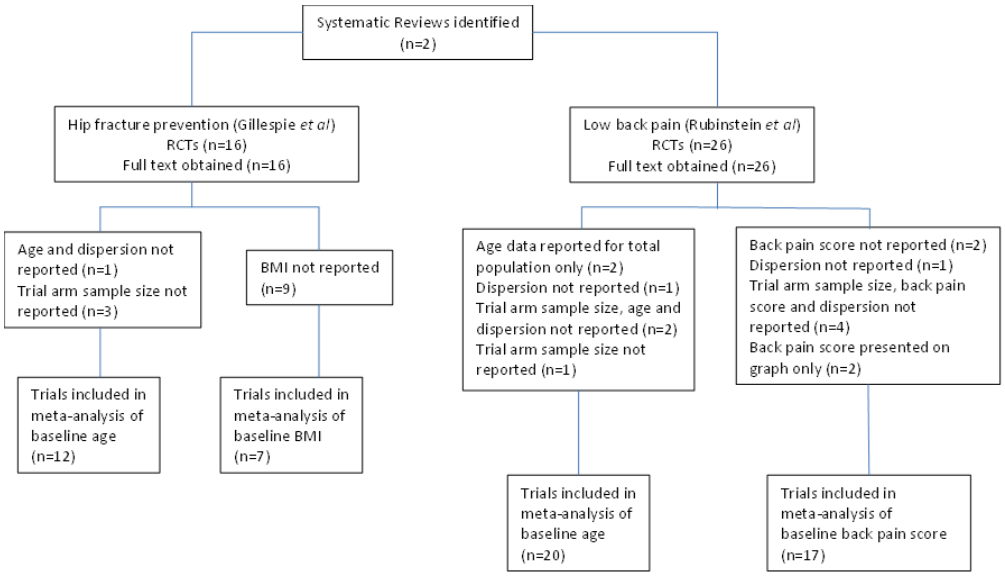
Appendix 1d: Meta-regression bubble plot and corresponding statistics for low back pain review for back pain score.

Reference

Clark, L., Fairhurst, C., Cook, E. and Torgerson, D.J., 2015. Important outcome predictors showed greater baseline heterogeneity than age in two systematic reviews. *Journal of clinical epidemiology*, *68*(2), pp.175-181

Additional note

In figure 1 'flow diagram of studies through the review' there is a minor labelling error. The Hip fracture prevention study and Low back pain labels were assigned to the incorrect boxes, they should be the other way around. The text in the manuscript is correct, the error relates to the flow diagram only. It has been requested that the journal updates the diagram in the publication. Below is the updated flow diagram.



Journal of Clinical Epidemiology 68 (2015) 175–181

Journal of Clinical Epidemiology

Important outcome predictors showed greater baseline heterogeneity than age in two systematic reviews

Laura Clark*, Caroline Fairhurst, Elizabeth Cook, David J. Torgerson

York Trials Unit, Department of Health Sciences, University of York, Heslington, York, YO10 5DD, North Yorkshire, UK

Accepted 1 September 2014; Published online 12 November 2014

Abstract

Objectives: An unknown number of randomized controlled trials (RCTs) have their treatment allocation subverted. If such trials are included in systematic reviews, biased results may be used to change policy. To assess whether a systematic review contains subverted trials, a meta-analysis of group differences regarding a baseline variable can be undertaken. In this article, the performance of age with another prognostic variable in detecting selection bias within systematic reviews is compared.

Study Design and Setting: Two Cochrane systematic reviews, one of low back pain and one of hip protectors for fracture prevention, were identified. The component RCT texts were obtained, and data were extracted on age, baseline back pain score (low back pain review), and baseline body mass (hip protector review). In this exemplar, we tested for baseline heterogeneity with a fixed-effects meta-analysis.

Results: Heterogeneity in age between the intervention and control groups was found. The observed heterogeneity increased with baseline back pain and body mass relative to age in each review.

Conclusion: We found that covariates predictive of outcome demonstrate greater heterogeneity than age. However, there were fewer missing data relating to age. Reviewers should consider using age and another prognostic covariate in baseline meta-analyses to check the validity of their results.

Keywords: Systematic review; Meta-analysis; Meta-regression; RCTs; Bias; Heterogeneity

1. Introduction

Systematic reviews of randomized controlled trials (RCTs) are powerful tools in shaping healthcare interventions and decisions [1]. It is imperative that when systematic reviews are undertaken, a rigorous assessment of bias is performed. There is evidence to suggest that some component RCTs in systematic reviews are biased because of inadequate methods of allocation concealment [2–6]. It is important that systematic reviews containing biased RCTs are identified as they jeopardize the validity of the review conclusions.

A suggested strategy is to assess intervention and control group differences in baseline covariates [7,8] and investigate the heterogeneity of the baseline imbalance between control and intervention groups. This approach of using a baseline variable, which randomization should ensure would differ only by chance, allows us to assess whether a meta-analysis of RCTs is reliable. We have shown that in a sample of 12 recently published systematic reviews of randomized trials, there was either heterogeneity or baseline differences in age between the intervention and control groups in six reviews [9]. If randomization had been adequately and faithfully conducted, then there should have been zero heterogeneity and minimal baseline imbalance. In a recent systematic review of 11 published and unpublished studies on the effectiveness of oseltamivir for treating influenza [10], the diagnostic measure of influenza (polymerase chain reaction) was found to be significantly different between the intervention and control groups [11]. The proportion of patients who were diagnosed with "true" influenza was statistically significantly lower in the intervention group than in the placebo group [relative risk 0.95; 95% confidence interval (CI): 0.91, 0.99], and it was hypothesized that this was due a possible failure to protect the randomization. Similarly in a systematic review of nine randomized trials of calcium supplementation for weight loss, Trowman et al. [8] found a statistically significantly lower body weight at baseline among those

Conflict of interest: None.

* Corresponding author. Tel.: 00 44 1904321115; fax: 00 44 1904321387.

E-mail address: laura.clark@york.ac.uk (L. Clark).

http://dx.doi.org/10.1016/j.jclinepi.2014.09.023

What is new?

- Heterogeneity is observed in age between the intervention and control groups in systematic reviews. The observed heterogeneity increased in a specific prognostic factor relevant to the outcome relative to age.
- An unknown number of randomized controlled trials (RCTs) have their allocation subverted, leading to potentially biased RCTs contributing to systematic reviews. RCTs could be subverted on age or a prognostic variable, which is measured at baseline. A meta-analysis of these covariates could identify systematic reviews that contain such biased trials.
- Systematic reviews are used to change policy and practice, and it is important that the conclusions reached are not based on biased results. We suggest that baseline heterogeneity is assessed relating to age and a covariate predictive of outcome when undertaking systematic review to assess the validity of the findings.

''randomized'' to the intervention group compared with those allocated to the control group. There is, therefore, accumulating evidence that allocation problems affect component trials in systematic reviews. In this study, we investigate the presence of group imbalances and heterogeneity across trials in two published systematic reviews based on age and another important prognostic factor. This article builds on our previous work that looked at single covariate imbalance (ie, age) and tests whether other covariates associated with outcome might be more important. The aims of this study were first to assess differences in heterogeneity between age and another covariate and second to determine whether, in a meta-regression, allocation concealment predicted any heterogeneity observed in the baseline meta-analyses of age and the prognostic factor.

2. Methods

We identified two systematic reviews in the Cochrane Library: one which measured the primary outcome (low back pain) at both baseline and the primary endpoint (spinal manipulative therapy for chronic low back pain [12]) and one which measured a prognostic factor [body mass (kg)] but could not measure the primary outcome (in this case, hip fractures) prerandomization (hip protectors for preventing hip fractures in older people [13]).

2.1. Data extraction and allocation concealment

The full texts for the component RCTs from the two reviews were retrieved. The following information was extracted for each trial arm where possible: summary of age and the prognostic factor (mean or median) and their measures of dispersion [standard deviation (SD), standard error, range, interquartile range], and number of participants. Double, independent data extraction was performed by two researchers (L.C. and E.C.), and disagreements were resolved by discussion. No other aspect of trial design was extracted from the articles, and we did not contact the trial authors of randomized trials where data were not available in the published articles.

In addition, we extracted the judgements of low, high, or unclear risk of bias awarded the RCTs of each review by the Cochrane reviewers.

If age and the prognostic factor were not summarized as a mean and SD, for example, if median and range were presented, then measures were converted using standard approximation formulas [14].

If the RCTs were constituted of more than two trial arms, then appropriate arms were combined to make one ''intervention group'' and one ''control group'' in accordance with the original systematic reviews. In the hip protector review, there were four cluster RCTs. We used the intracluster correlation coefficient from one of the cluster trials to estimate a design effect (ie, 1.24) and divided this into the sample size to reduce the sample size in line with the likely clustering [15].

2.2. Meta-analysis

Fixed-effect meta-analyses of age and the prognostic factor were performed for each review on the assumption that there was a common treatment estimate (ie, zero) across the randomized trials. Because of the principle of random allocation, we would assume that all trials that had conducted ''true'' randomization would result in comparability of covariates by group allocation except for chance differences. Following this, then the assumption is that all rigorous randomized trials, when meta-analyzed, will result in an estimated difference of zero in mean baseline variable between groups. Therefore, the assumption that the true ''treatment'' estimate of zero on baseline covariates must be true across all the trials. The *P*-value for the difference in the variables between the control and intervention groups for each systematic review was calculated. The I^2 value of heterogeneity from the meta-analysis was interpreted in line with the Cochrane hand book guidelines: 0% to 40% might not be important; 30% to 60% may represent moderate heterogeneity; 50% to 90% may represent substantial heterogeneity; and 75% to 100% may represent considerable heterogeneity [16]. The I^2 values are presented with 95% CI computed by hand using the ''test-based'' method [17]. Separate meta-regressions were performed for each review to assess whether allocation concealment adequacy was a predictor of heterogeneity in the variables [16]. For the hip protector review, two ''collapsed'' meta-regressions were conducted for each

Table 1. Results of the analyses for the two systematic reviews

Variable	Intervention group, mean (SD)	Control group, mean (SD)	Standardized mean difference	Difference between groups (*P*-value)	I^2 value	Meta-regression (*P*-value)
Hip protector review (n = 16[a])						
Age (n = 12)	83.6 (6.9)	83.4 (7.1)	−0.007	0.71	53.7	Both <0.01
BMI (n = 7)	55.0 (11.5)	55.9 (11.4)	−0.093	<0.001	76.4	Both <0.01
Low back pain review (n = 26)						
Age (n = 20)	44.2 (11.4)	44.0 (11.5)	0.003	0.91	22.0	0.19
Back pain (n = 17)	17.1 (8.5)	17.2 (8.7)	−0.021	0.51	55.8	0.40

Abbreviations: SD, standard deviation; BMI, body mass index.
[a] n indicates the number of RCTs.

factor: first comparing the low-risk trials with those with a high or unclear risk and second comparing those with a low or unclear risk with high-risk trials. Analyses were performed in Comprehensive Meta-Analysis 2 (Biostat, Englewood, NJ, USA).

3. Results

The results of the meta-analyses for both reviews are presented in Table 1.

3.1. Hip protector review (n = 16 RCTs)

3.1.1. Age

Twelve trials were included in the meta-analysis of baseline age, of which 42% (n = 5) were classified as having a low risk of bias, 25% (n = 3) as having a high risk of bias, and 33% (n = 4) as unclear (Fig. 1). The number of participants randomized into each trial ranged from 72 to 6,868 (median 581). The standardized mean difference in age was −0.007 (intervention group on average younger than control group). This difference was not significant (P = 0.71; Fig. 2); however, there was evidence of moderate heterogeneity (I^2 = 53.7; 95% CI: 10.9%, 76.0%), with the CI suggesting that the heterogeneity is statistically significantly different from zero. The meta-regression found evidence to suggest that allocation concealment was a statistically significant predictor of heterogeneity (P < 0.01, Appendix at www.jclinepi.com), when trials with an unclear risk of bias were combined with both low- or high-risk trials and compared with the remaining subset.

3.1.2. Body mass

Seven trials were included in the meta-analysis of baseline body mass, of which 43% (n = 3) were classified as having a low risk of bias, 29% (n = 2) as having a high risk of bias, and 29% (n = 2) as unclear (Fig. 1). The number of participants randomized into each trial ranged from 164 to 6,868 (median 672). There was evidence of a statistically significant difference in body mass between the intervention and control groups (standardized mean difference −0.093; P < 0.001; Fig. 2) and of considerable heterogeneity (I^2 = 76.4; 95% CI: 50.3%, 88.8%), with the CI suggesting that the heterogeneity is statistically significantly different from zero. Both meta-regression analyses yielded a P-value of less than 0.01; thus, there was evidence to suggest that allocation concealment explained

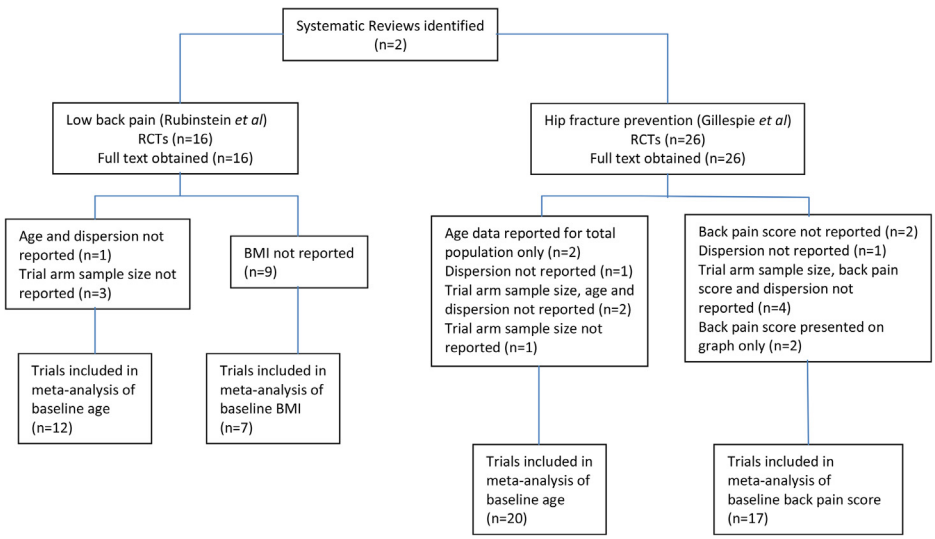


Fig. 1. Flow diagram of studies through the review.

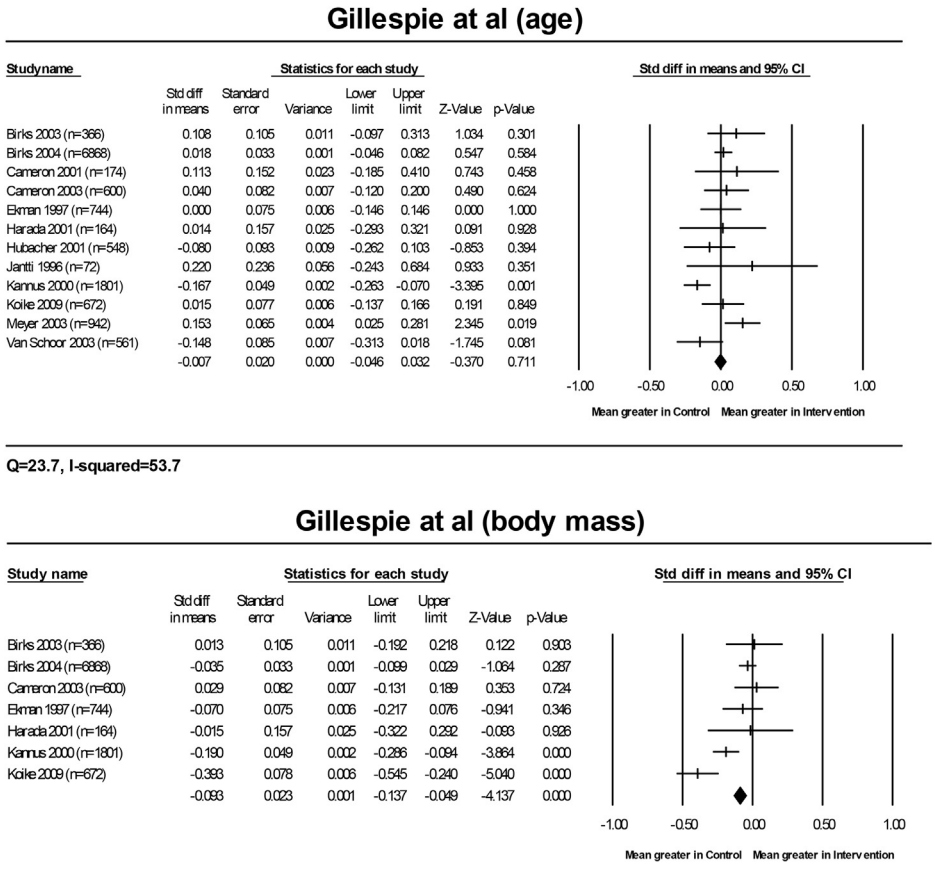


Fig. 2. Forest plots for the hip protector review.

some of the heterogeneity observed (Appendix at www.jclinepi.com).

3.2. Low back pain review ($n = 26$ RCTs)

3.2.1. Age

Twenty trials were included in the meta-analysis of baseline age, of which 50% ($n = 10$) were classified as having a low risk of bias and 50% ($n = 10$) as unclear (Fig. 1). The number of participants randomized into each trial ranged from 30 to 1,334 (median 154). No difference in age was detected between the intervention and control groups (standardized mean difference 0.003; $P = 0.91$; Fig. 3) and only weak heterogeneity ($I^2 = 22.0$; 95% CI: 0.0%, 54.7%), with a CI that includes zero. The meta-regression did not suggest that allocation concealment predicted this heterogeneity ($P = 0.19$; Appendix at www.jclinepi.com).

3.2.2. Back pain score

Seventeen trials were included in the meta-analysis of baseline back pain score, of which 47% ($n = 8$) were classified as having a low risk of bias and 53% ($n = 9$) as unclear (Fig. 1). The number of participants randomized into each trial ranged from 47 to 1,334 (median 174). No difference between the age of the intervention and control groups was detected (standardized mean difference −0.021; $P = 0.50$; Fig. 3), but there was evidence of substantial statistically significant heterogeneity ($I^2 = 55.8$; 95% CI: 23.7%, 74.4%). As with age, the meta-regression did not show allocation concealment to be a predictor of this heterogeneity ($P = 0.40$; Appendix at www.jclinepi.com).

4. Discussion

The purpose of this work was to ascertain whether it is viable to use a prognostic factor to assess baseline heterogeneity within systematic reviews.

Our meta-analyses suggested that there is evidence of heterogeneity between the control and intervention groups across both variables in the two reviews, when there should be none if patients were randomized and allocated to their treatment group securely. Indeed, we have shown that although I^2 values calculated from the same meta-analysis are likely to be correlated, the prognostic factors exhibited more heterogeneity than age, which suggests that they may be more sensitive to selection bias. However, although in

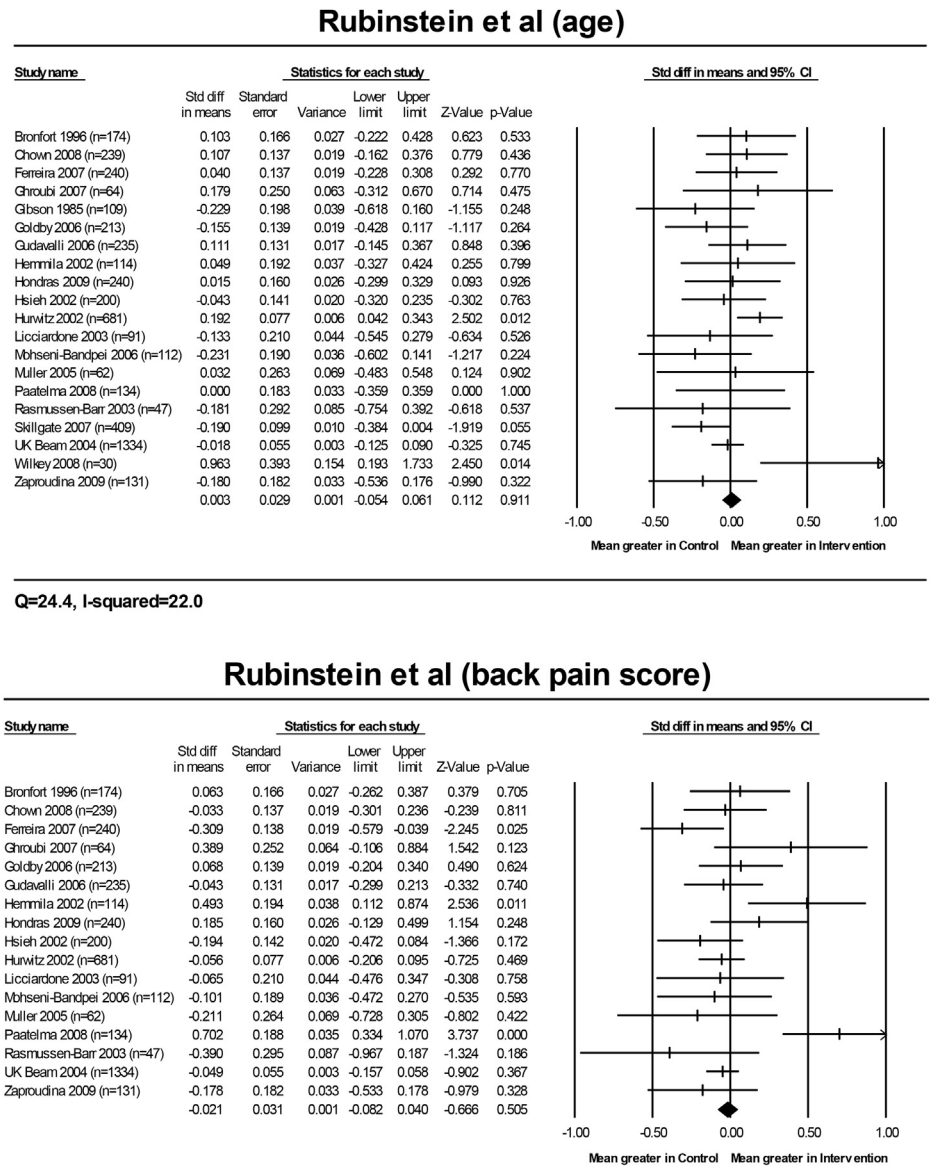


Fig. 3. Forest plots for the low back pain review.

both of our examples, we found that age did not generate as much heterogeneity as other baseline covariates, given only two systematic reviews, this may be due to chance and would require further investigation in a wider range of studies.

Meta-regression of the hip protector review showed that allocation concealment is a statistically significant predictor of the heterogeneity observed in both the age and body mass data. Meta-regression analysis for the back pain review did not suggest that allocation concealment explained the heterogeneity observed; however, half of the articles had an unclear risk of bias. It is, therefore, possible that issues with the reporting quality masked the results. Nevertheless, the consequence of these findings is that the conclusions from both of these reviews are based on potentially biased results and should be treated with caution.

It was put forward, by a reviewer of this article, that publication bias is an alternative explanation for the heterogeneity observed in the systematic reviews by the following argument: imagine a population of trials with an average baseline imbalance in a particular confounding variable of zero. Because of chance, some of the trials will have an imbalance in this variable in one direction, whereas others will have an imbalance in the other direction. If the respective variable is a treatment effect moderator, one form of the imbalance may lead to more positive results. If only these trials are published and available for a systematic review, the component set of trials will exhibit baseline

imbalance and the reason will be due to publication bias rather than ''subverted'' or inadequate allocation.

However, we believe that publication bias is not really a reasonable explanation as if trials imbalanced in a particular direction are more likely to get published because they are positive, then this is likely to reduce heterogeneity but increase the baseline imbalance, which was not observed here. A reviewer should expect to see no heterogeneity and no statistically significant differences in confounders between the groups. If a review showed a significant difference in baseline measure of outcome, say, but no heterogeneity, the conclusion would be that a proportion of the included trials have allocation subversion all favoring the same treatment arm. However, if there were simply significant heterogeneity but no overall difference in the baseline covariate, this does not imply that because subversion is operating in both directions that the review's results are believable. It may mean that the trials are in significant imbalance in a particular direction that favors an unknown covariate. In truly randomized trials, this unknown or unmeasured covariate will be balanced across a group of trials, but when heterogeneity is present, we cannot be confident that this holds true.

A statistically significant pooled baseline imbalance was seen in body mass in the hip protector review. Although the difference was small, any imbalance on an observable variable and the presence of heterogeneity is a marker of inadequate allocation concealment. The magnitude of the difference is and is not important. Obviously, a large baseline imbalance is likely to be more worrying, but because we could not know which factor a trial's randomization is subverted on, we can consider only a justifiable proxy, here, age and a prognostic variable for each review. A much bigger ''true'' imbalance in either an unknown or unmeasured confounder may be present, which would undermine the review. Indeed, our review is likely to underestimate imbalances in randomization if the clinician is allocating patients based on an unmeasured or unreported covariate, which correlates only weakly with age or other reported prognostic variables. Consequently, any observed baseline heterogeneity is, in our opinion, a cause for concern.

It is a commonly held belief that the chance of apparent baseline imbalance in prognostic factors is higher for small sample sizes than for large ones. Senn [18] cites this as one of his seven myths of randomization in clinical trials. Indeed, we included the sample size of the component RCTs as an independent variable in the meta-regressions, and it was not seen to statistically significantly predict heterogeneity in any of them.

We would recommend that the techniques described here are routinely implemented by other systematic reviewers. This supports the findings from previous work we have done that shows heterogeneity in age is important in identifying reviews containing poorly conducted trials [9].

There are a number of limitations to our review. We had a relatively small sample size for the analysis of body mass due to the exclusion of a number of RCTs for which data could not be extracted. To perform the meta-regression analysis, 10 or more trials are required for the test to be reliable and so the results should be interpreted with caution [16].

We retained the judgements made by the Cochrane reviewers on the adequacy of the reported allocation concealment method. Using a different risk of bias tool such as the Jadad scale [19] may have resulted in different judgements being made on the allocation concealment adequacy.

5. Conclusion

Age was reported in more of the component RCTs than our chosen prognostic factor and so is useful to examine. We would recommend that age and another relevant prognostic factor are investigated simultaneously if possible when exploring baseline heterogeneity within systematic reviews. Age is known to be a variable on which allocation within RCTs have been subverted [20]. If there is a lack of available data to perform an analysis on a prognostic factor specific to the review, then age can be considered. However, prognostic factors specific to a systematic review may be more sensitive in detecting heterogeneity than age if age is not an important covariate; for example, if the RCT only includes patients within a very narrow age range, then the subversion may be more likely to be based on a pertinent prognostic factor. It is important that baseline imbalances are identified so that results from systematic reviews with such imbalances are treated with the appropriate caution. When baseline imbalances have been identified, it will allow reviewers to perform sensitivity analyses and publish systematic reviews with potentially more reliable conclusions.

Acknowledgments

The authors would like to thank Belen Corbacho Martin for her help with translating French language manuscripts.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2014.09.023.

References

[1] Egger M, Smith GD, Altman D. Systematic reviews in health care: meta-analysis in context. London: John Wiley & Sons; 2008.
[2] Schulz K, A DG, Moher D. Letter to the Editor: allocation concealment in clinical trials. JAMA 2002;288:2406−9.
[3] Schulz KF. Subverting randomization in controlled trials. JAMA 1995;274:1456−8.
[4] Schulz KF. Randomised trials, human nature, and reporting guidelines. Lancet 1996;348:596−8.

[5] Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. JAMA 1995;273:408.
[6] Torgerson DJ, Torgerson CJ. Designing randomised trials. Basingstoke: Palgrave MacMillan; 2008.
[7] Trowman R, Dumville JC, Hahn S, Torgerson DJ. A systematic review of the effects of calcium supplementation on body weight. Br J Nutr 2006;95(06):1033−8.
[8] Trowman R, Dumville JC, Torgerson DJ, Cranny G. The impact of trial baseline imbalances should be considered in systematic reviews: a methodological case study. J Clin Epidemiol 2007;60:1229−33.
[9] Clark L, Fairhurst C, Hewitt CE, Birks Y, Brabyn S, Cockayne S, et al. A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups. J Clin Epidemiol 2014;67:1016−24.
[10] Ebell MH, Call M, Shinholser J. Effectiveness of oseltamivir in adults: a meta-analysis of published and unpublished clinical trials. J Fam Pract 2013;30:125−33.
[11] Ebell MH. Methodological concerns about studies on oseltamivir for flu. BMJ 2013;347:f7148.
[12] Rubinstein SM, van Middelkoop M, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for chronic low-back pain. Cochrane Database Syst Rev 2011;CD008112.
[13] Gillespie WJ, Gillespie LD, Parker MJ. Hip protectors for preventing hip fractures in older people. Cochrane Database Syst Rev 2010; CD001255.
[14] Hozo S, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol 2005;5:13.
[15] Donner A, Klar N. Issues in the meta-analysis of cluster randomized trials. Stat Med 2002;21:2971−80.
[16] Deeks JJ, Higgins JPT, editors. AD. Chapter 9: analysing data and undertaking meta-analyses. In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions. The Cochrane Collaboration; 2011. Version 5.1.0 (updated March 2011). Available at. www.cochrane-handbook.org. 2011. Accessed November 29, 2013.
[17] Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539−58.
[18] Senn S. Seven myths of randomisation in clinical trials. Stat Med 2013;32:1439−50.
[19] Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996;17:1−12.
[20] Kennedy A, Grant A. Subversion of allocation in a randomised controlled trial. Control Clin Trials 1997;18:S77−8.

Appendix 1

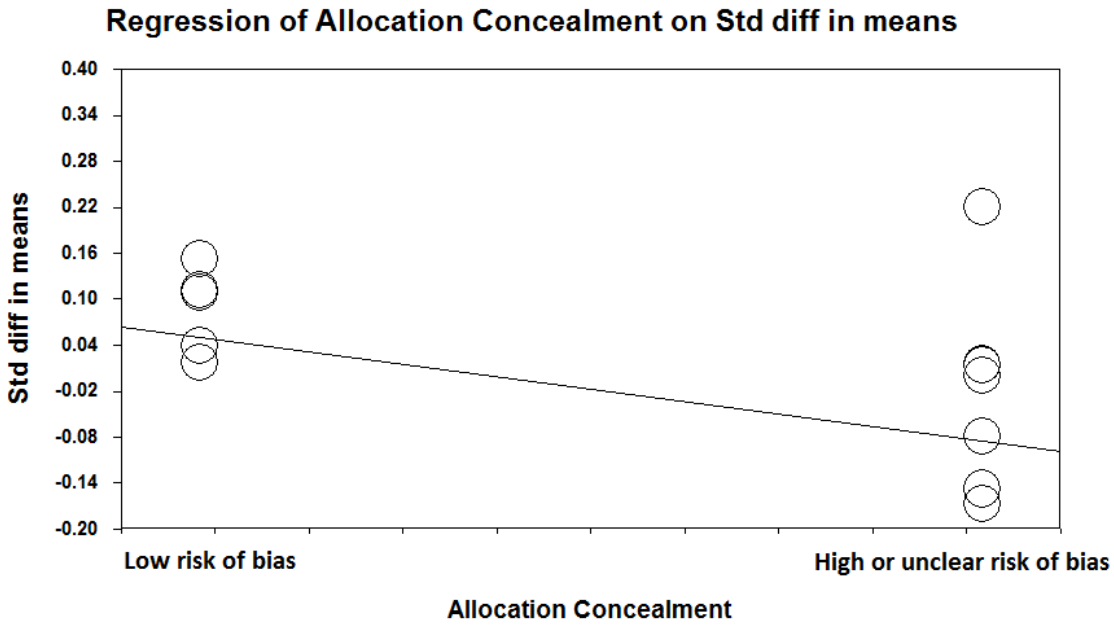
1a.

Meta-regression bubble plot and corresponding statistics for hip protectors review for age (low versus high or unclear risk of bias).

Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	-0.13621	0.04027	-0.21514	-0.05729	-3.38253	0.00072
Intercept	0.18696	0.06081	0.06778	0.30613	3.07461	0.00211
Tau-squared	0.00142					

	Q	df	p-value
Model	11.44150	1.00000	0.00072
Residual	12.29133	10.00000	0.26603
Total	23.73283	11.00000	0.01391

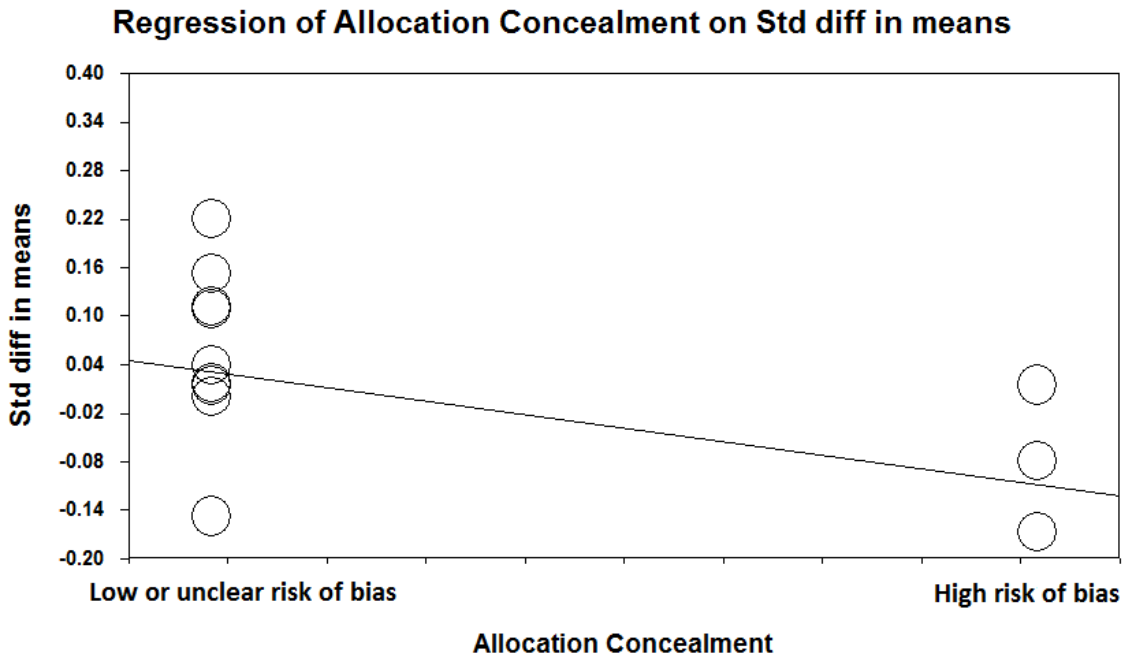


Meta-regression bubble plot and corresponding statistics for hip protectors review for age (low or unclear versus high risk of bias).

Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	-0.14007	0.04452	-0.22732	-0.05281	-3.14627	0.00165
Intercept	0.17146	0.06023	0.05341	0.28950	2.84677	0.00442
Tau-squared	0.00230					

	Q	df	p-value
Model	9.89905	1.00000	0.00165
Residual	13.83379	10.00000	0.18071
Total	23.73283	11.00000	0.01391



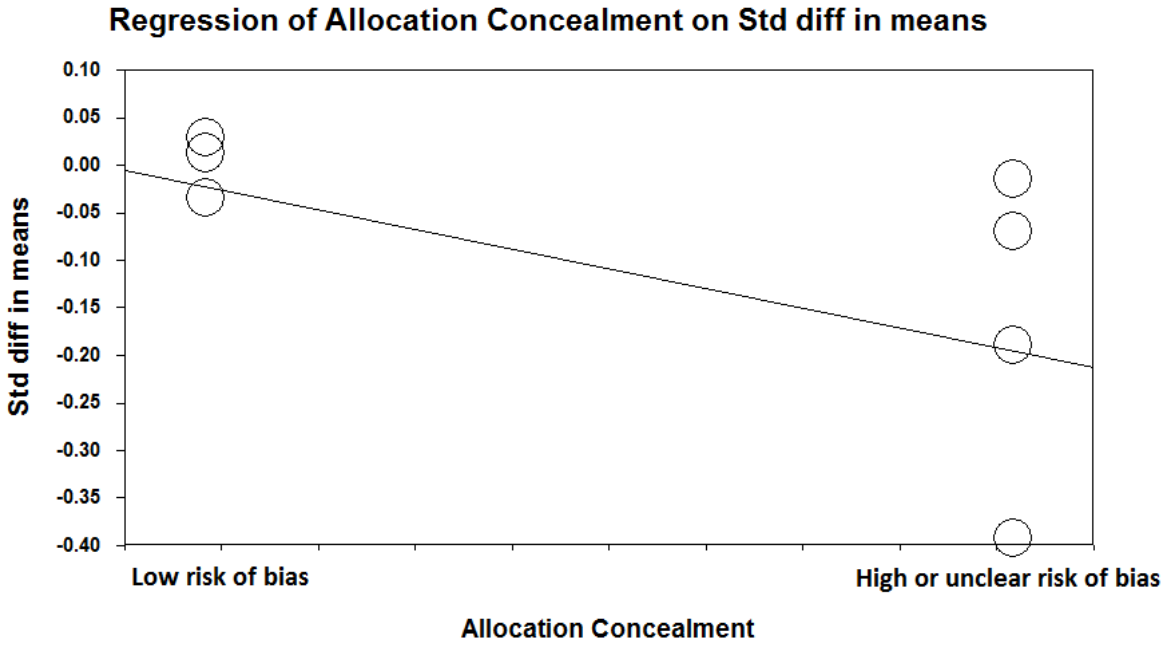
1b.

Meta-regression bubble plot and corresponding statistics for hip protectors review for body mass (low versus high or unclear risk of bias).

Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	-0.17283	0.04591	-0.26281	-0.08285	-3.76447	0.00017
Intercept	0.14978	0.06840	0.01573	0.28384	2.18996	0.02853
Tau-squared	0.00680					

	Q	df	p-value
Model	14.17127	1.00000	0.00017
Residual	11.21759	5.00000	0.04723
Total	25.38885	6.00000	0.00029

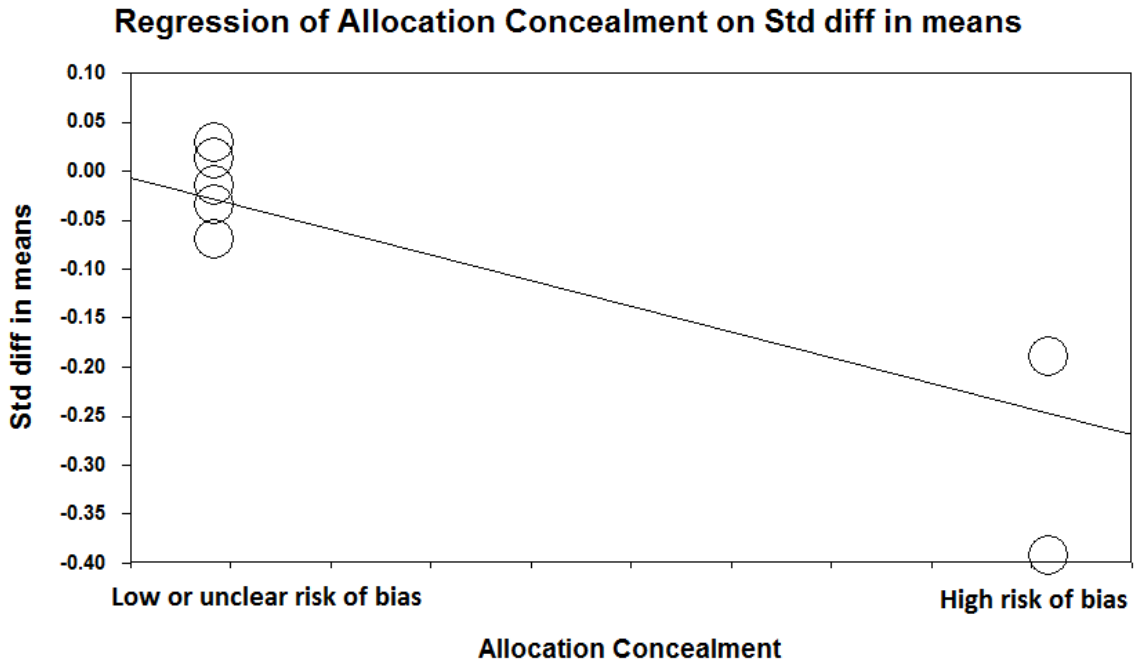


Meta-regression bubble plot and corresponding statistics for hip protectors review for body mass (low or unclear versus high risk of bias).

Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	-0.21860	0.04947	-0.31556	-0.12165	-4.41894	0.00001
Intercept	0.18970	0.06789	0.05663	0.32277	2.79400	0.00521
Tau-squared	0.00090					

	Q	df	p-value
Model	19.52702	1.00000	0.00001
Residual	5.86184	5.00000	0.31990
Total	25.38885	6.00000	0.00029



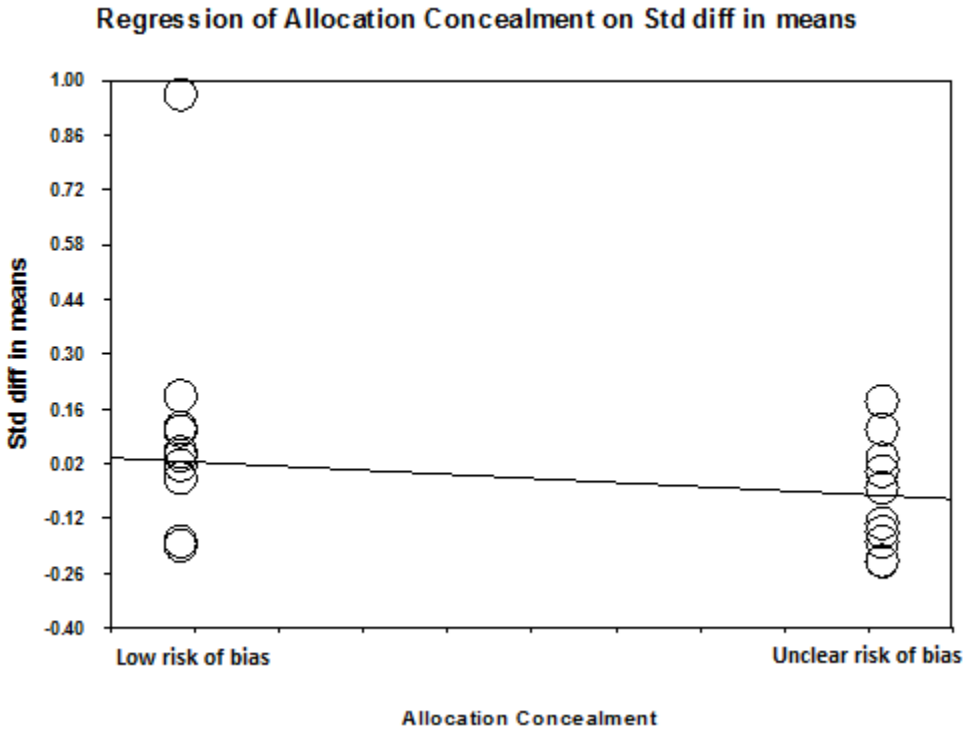
1c.

Meta-regression bubble plot and corresponding statistics for low back pain review for age.

Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	-0.04367	0.03344	-0.10921	0.02187	-1.30604	0.19154
Intercept	0.06978	0.05878	-0.04542	0.18499	1.18718	0.23516
Tau-squared	0.00498					

	Q	df	p-value
Model	1.70573	1.00000	0.19154
Residual	22.64853	18.00000	0.20446
Total	24.35426	19.00000	0.18291



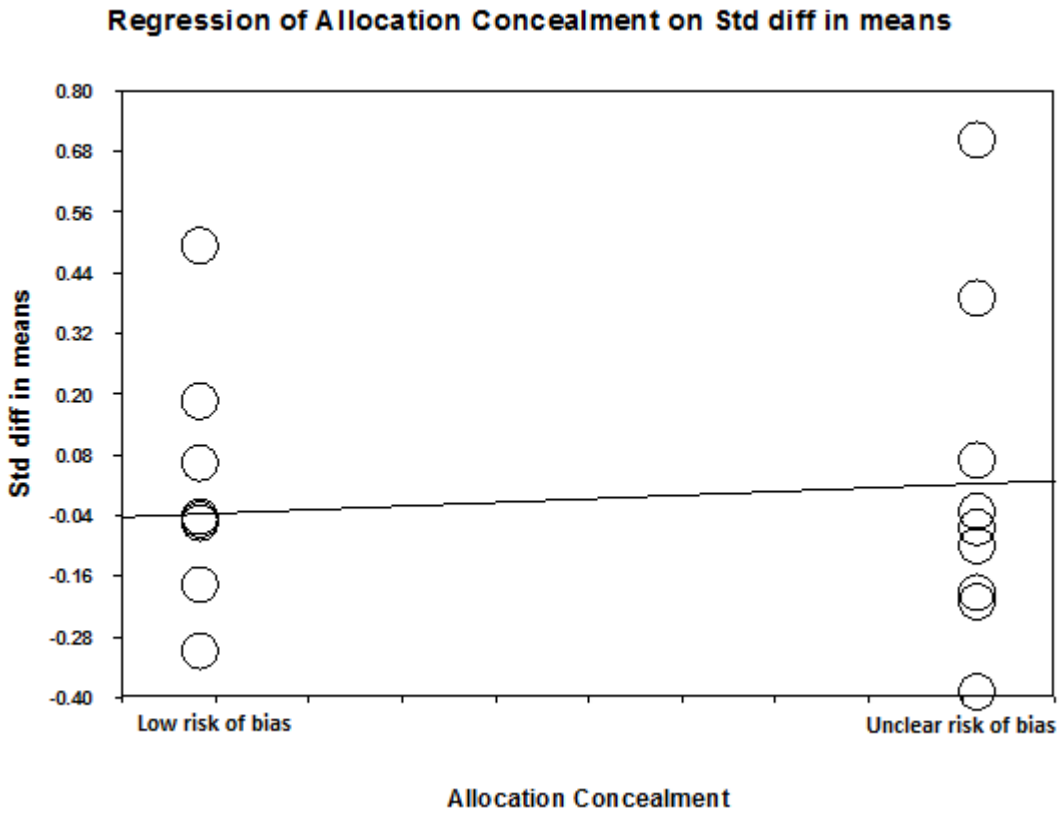
1d.

Meta-regression bubble plot and corresponding statistics for low back pain review for back pain score.

Fixed effect regression

	Point estimate	Standard error	Lower limit	Upper limit	Z-value	p-Value
Slope	0.02995	0.03525	-0.03914	0.09904	0.84964	0.39552
Intercept	-0.06693	0.06261	-0.18966	0.05579	-1.06896	0.28509
Tau-squared	0.02618					

	Q	df	p-value
Model	0.72189	1.00000	0.39552
Residual	35.46640	15.00000	0.00211
Total	36.18829	16.00000	0.00272



Appendix 5: paper 3

Allocation concealment in randomised controlled trials: are we getting better?

Reference

Clark, L., Fairhurst, C. and Torgerson, D.J., 2016. Allocation concealment in randomised controlled trials: are we getting better?. BMJ, 355.

BMJ 2016;355:i5663 doi: 10.1136/bmj.i5663 (Published 17 November 2016)

ANALYSIS

Allocation concealment in randomised controlled trials: are we getting better?

Laura Clark and colleagues assess the allocation concealment methods in a sample of randomised controlled trial publications

Laura Clark *research fellow*, Caroline Fairhurst *research fellow*, David J Torgerson *director*

York Trials Unit, University of York, York YO10 5DD, UK

A robust randomised controlled trial (RCT) must use allocation concealment—that is, separate the act of randomisation from the person recruiting participants. Poor randomisation methods cause exaggerated treatment effects, are open to subversion by researchers or clinicians, and have a knock-on effect on systematic reviews.[1-3]

The CONSORT statement, which leading medical journals endorse, states that the method of allocation (comprising sequence generation, allocation concealment mechanism, and implementation) should be clearly described.[4] Allocation concealment is dependent on the method of sequence generation as well as the concealment mechanism.

Almost a fifth of trials published in major medical journals in 2002 used inadequate concealment, and a quarter failed to describe how the allocation was concealed.[2] Here we examine a sample of RCTs published in 2015 to see whether the situation has improved.

Defining inadequate allocation concealment

We searched four high impact medical journals (*The BMJ, Journal of the American Medical Association (JAMA);* the *Lancet*, and the *New England Journal of Medicine (NEJM))* and found 79 RCTs published between June and August 2015. We extracted and judged their mechanism for allocation concealment, taking into consideration the study design, sequence generation method, and allocation concealment mechanism. We defined an inadequate process as one that used envelopes as the method of allocation concealment (box 1) or used stratified block randomisation by site with small block sizes as the sequence generation method (box 2), except in double blind trials. If insufficient detail was provided in the paper, we checked the protocol or emailed the authors.

Fifteen trials were poorly randomised

Twenty seven (34%) of the RCTs were placebo controlled double blind trials, in which allocation is generally well concealed; participants are assigned a number corresponding to a packet of drugs, and only the pharmacist has access to the unblinding codes. One of these trials used envelopes, but as the pharmacist opened the envelopes after the clinician enrolled the participant it was deemed adequate. We judged these trials, and 22 (28%) of the remaining trials, as adequate. We initially found that 13 trials (16%) had used a randomisation method that put them at risk of bias and 17 trials (22%) contained insufficient detail to determine whether the method of allocation concealment was adequate. We received more information from the authors of nine (53%) of these trials, two of which were found to be inadequate, giving a total of 15 (19%) trials with inadequate concealment (table 1⇓). Seven trials used envelopes to allocate participants, seven trials used small block sizes and/or stratified by site, and one used both small blocks and envelopes.

We noted two inconsistencies with the use of block randomisation. The trial by Senn et al had an imbalance of 12 participants between the randomised groups; the largest possible imbalance for a block size of two stratified by three centres is three. Correspondence with the authors confirmed that 10 cases were misallocated to the intervention group.[26] The authors said that a combination of technical and human errors accounted for the imbalance—"the laptop froze during randomisation, the server went down temporarily, research assistants inadvertently practiced on a live site, and participants went to the wrong session." They reassigned the 10 participants to their originally assigned groups and found no change in benefit for the intervention.

The trial by Cox and colleagues[27] said that patients were allocated 2:1 with blocks of four, but with a 2:1 ratio the block size should be divisible by three. Correspondence with the author confirmed that the statement in the paper was a mistake, and an erratum has been published in the *Lancet*.

More rigour is needed

Our findings—that 19% of trials described inadequate methods of allocation concealment and 22% failed to report their

Correspondence to: laura.clark@york.ac.ukAccepted 18 October 2016

Box 1: Sequentially numbered opaque sealed envelopes (SNOSE)

Envelopes containing the treatment allocation are opened by the recruiting clinician on participant enrolment. To be robust, the envelopes should be truly opaque, sequentially numbered, and opened in the correct order. The clinician should not open the envelope in advance and should ensure that the envelope seal has not been broken. Even in these circumstances we cannot guarantee that envelopes have not been opened in advance to allow strategic scheduling of patient appointments to match the recruiter's preferred allocation. A surgical trial found that three of five surgeons had opened envelopes in order to subvert the randomisation.[5] Trials that use SNOSE are more likely to show a statistically significant treatment effect than trials that use more secure allocation methods, such as web based or telephone randomisation.[2] Despite this, the Cochrane handbook of systematic reviews says that trials that use SNOSE have a low risk of bias.[6] In practice, the bias risk is only lessened or eliminated when the people with access to the envelopes are distinct from those recruiting participants to the trial.

Box 2: Block randomisation

Most trials use restricted randomisation methods to generate the allocation sequence, such as stratification,[7] which requires the use of block allocation within each strata. In this method a limited sequence of allocations are repeated: for example, a block size of four with two treatments (A and B) has six potential blocks of sequences (AABB, ABAB, BBAA, BABA, ABBA, and BAAB). Even if a robust allocation concealment mechanism is used (such as central web based randomisation), subversion of the allocation is possible. For example, stratifying by site and using a small fixed block size makes the allocation sequence predictable[7]—in a two arm trial using randomisation stratified by site and a fixed block size of four, every fourth allocation can be accurately predicted, and the third allocation in a block can be predicted a third of the time, if one keeps a record of the previous allocations given to patients. Only simple randomisation avoids the problem.[8] This is not just a theoretical concern. In a trial of supplemental oxygen for retinopathy of prematurity, clinicians tried to avoid recruiting and allocating patients to the control group, which they judged as being undesirable for some patients.[9] An RCT of rehabilitation for patients with fractured neck of femur failed because, despite using telephone randomisation, the block size of six was deciphered part way through the trial, which led to prediction and selection bias.[10] The use of larger or variable block sizes or avoiding stratifying by site can minimise this problem. Stratification is only beneficial for small sample sizes or slowly recruiting trials or if treatment logistics demand some predictability of treatment volume (for example, surgical treatments). For many, if not most, trials simple randomisation is preferred (if n>100).[7]

randomisation methods clearly—are similar to those found in 2002 (18% and 26% were inadequate and unclear, respectively). The sample size of the 2002 study was much larger (n=234) than here (n=79).

Despite the known inadequacies of the use of envelopes for treatment allocation, we found their use in at least nine trials. Regardless of whether the allocation code is computer generated or the investigators are blinded, the person enrolling the patient can potentially open an envelope in advance. In certain scenarios, such as trials conducted in remote areas, the use of SNOSE may be the only feasible approach—if so, envelopes should be sequentially numbered and it should be stated clearly that a person who is separate from the assessment and recruitment of the patient opens the envelope, as described in the CONSORT statement (item 10).[4] Rigour might be further enhanced by writing patients' details on the outside of envelopes that contain carbon paper, so that the patients' names are transferred to the sheet of paper with the allocation on before the envelope is opened.

Researchers and funders must try harder (box 3). Minimisation is an alternative to block randomisation, which allocates participants to the trial arm that best maintains balance across specified stratifying factors.[28] Minimisation is a dynamic mode of random allocation—instead of using an allocation list that is generated before the trial begins, a participant's allocation depends on their characteristics and, crucially, the characteristics of the participants already enrolled. Allocation by minimisation is much more difficult to predict than stratified, block randomisation, particularly if a random element is introduced (that is, instead of using minimisation to completely determine the allocations, a pre-specified probability that the treatment will be chosen is used for each randomisation).

Leading medical journals should specify that, at a specified point in the future, trials using small block sizes or SNOSE will not be accepted for publication, except for in extenuating circumstances, such as in emergency medicine or based on the location of the trial (for example, in remote areas that don't have access to the internet or a telephone system). If these methods are used the details should be described explicitly and transparently. Meanwhile, journals could request that completed trials that used blocked randomisation perform a Berger-Exner test[29] to assess for selection bias. This statistical test looks for a relationship between baseline variables and the position of the participant in the block. In the absence of subversion no relation should be found between these factors. This is not onerous for the authors and would demonstrate that their trial was conducted with robust methodology. Refusal to perform the test would preclude publication. Some trials[30 31] that use block randomisation do report checking for subversion, but this statistical approach is rarely used.

Journals should ask a methodologist to carefully review the reported methods of randomisation before any trial is published or even sent for review. If those planning future trials felt that publication in a high impact journal was closed to them if their randomisation system was not robust they might be more focused at the design stage. It would also lead to an awareness of, and education in, the importance of robust randomisation methods and clearer reporting. Although journals endorse CONSORT, they are not enforcing it properly. We propose that, for randomisation at least, an editor or reviewer is mandated to ensure that the text is compliant with CONSORT.

Contributors and sources: DT conceived of the original idea, LC identified trials, and CF, DT, and LC extracted data from the trials. DT jointly wrote the first draft with LC, and CF contributed to revising the draft. LC is the guarantor.

Competing interests: We have read and understood the BMJ policy on declaration of interests and declare that we have no competing interests.

We thank Doug Altman for his helpful comments.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Savović J, Jones HE, Altman DG, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med* 2012;157:429-38. doi:10.7326/0003-4819-157-6-201209180-00537 pmid:22945832.
2 Hewitt C, Hahn S, Torgerson DJ, Watson J, Bland JM. Adequacy and reporting of allocation concealment: review of recent trials published in four general medical journals. *BMJ* 2005;330:1057-8. doi:10.1136/bmj.38413.576713.AE pmid:15760970.
3 Clark L, Fairhurst C, Hewitt CE, et al. A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups. *J Clin Epidemiol* 2014;67:1016-24. doi:10.1016/j.jclinepi.2014.04.007 pmid:24909873.
4 Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869. doi:10.1136/bmj.c869 pmid:20332511.
5 Kennedy A, Grant A. Subversion of allocation in a randomised controlled trial. *Control Clin Trials* 1997;18(Supplement 3):S77-S88doi:10.1016/S0197-2456(97)91044-8.
6 Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. Version 5.10. 2011. http://handbook.cochrane.org/front_page.htm).

Box 3: Advice for researchers and funders

Advice for researchers

Consult with a methodologist when designing an RCT

Use simple randomisation for larger sample sizes and consider using minimisation if stratification is required

When calculating costings for an RCT, factor in funding for third party randomisation services

If SNOSE are the only practical method to conceal the randomisation sequence ensure that their preparation and execution is methodologically rigorous—state who prepared the envelopes, what (if any) additional security measures were in place, and who opened the envelopes in the publication of the RCT

Explicitly state randomisation and allocation concealment methods in protocol and publication. If there is a reason why the information cannot be provided in the publication (for example, the word limit) refer to the section of the protocol in which a reader can find it, or supply the information in a supplementary document

Advice for funders

Ensure a methodologist is involved in designing and reporting the RCT

Do not fund RCTs using SNOSE unless justified as the only available option

Do not fund research that uses a block size of less than six with site stratification

Key messages

Good allocation concealment is vital for robust randomised trials

38% of the trials in our sample of those published in major medical journals did not report good allocation concealment methods

Journals should insist that sealed envelopes and other weak concealment methods are no longer used

7 Hewitt CE, Torgerson DJ. Is restricted randomisation necessary? *BMJ* 2006;332:1506-8. doi:10.1136/bmj.332.7556.1506 pmid:16793819.
8 Zhao W. A better alternative to stratified permuted block design for subject randomization in clinical trials. *Stat Med* 2014;33:5239-48. doi:10.1002/sim.6266 pmid:25043719.
9 Engel RR, Oden NL, Cohen GR, Phelps DL. STOP-ROP Multicentre Study Group. Influence of prior assignment on refusal rates in a trial of a supplemental oxygen for retinopathy of prematurity. *Paediatr Perinat Epidemiol* 2006;20:348-59. doi:10.1111/j.1365-3016.2006.00721.x pmid:16879508.
10 Turner J, Russell D, Russell I, et al. Baseline imbalance in a randomised trial: a cautionary tale. *Care of the Critically Ill* 2006;22:1-4.
11 Smith JE, Rockett M, Creanor S, et al. PASTIES Research Team. Pain solutions in the emergency setting (PASTIES)—patient controlled analgesia versus routine care in emergency department patients with non-traumatic abdominal pain: randomised trial. *BMJ* 2015;350:h3147. doi:10.1136/bmj.h3147 pmid:26094712.
12 Smith JE, Rockett M, S SC, et al. PASTIES Research Team. Pain solutions in the emergency setting (PASTIES)—patient controlled analgesia versus routine care in emergency department patients with pain from traumatic injuries: randomised trial. *BMJ* 2015;350:h2988. doi:10.1136/bmj.h2988 pmid:26094763.
13 Detollenaere RJ, den Boon J, Stekelenburg J, et al. Sacrospinous hysteropexy versus vaginal hysterectomy with suspension of the uterosacral ligaments in women with uterine prolapse stage 2 or higher: multicentre randomised non-inferiority trial. *BMJ* 2015;351:h3717.
14 Salminen P, Paajanen H, Rautio T, et al. Antibiotic therapy vs appendectomy for treatment of uncomplicated acute appendicitis. *JAMA* 2015;313:2340-8. doi:10.1001/jama.2015.6154 pmid:26080338.
15 Stéphan F, Barrucand B, Petit P, et al. BiPOP Study Group. High-flow oxygen vs noninvasive positive airway pressure in hypoxemic patients after cardiothoracic surgery: A randomized clinical trial. *JAMA* 2015;313:2331-9. doi:10.1001/jama.2015.5213 pmid:25980660.
16 Polusny MA, Erbes CR, Thuras P, et al. Mindfulness-based stress reduction for posttraumatic stress disorder among veterans: a randomized clinical trial. *JAMA* 2015;314:456-65. doi:10.1001/jama.2015.8361. pmid:26241597.
17 Rhodes KV, Rodgers M, Sommers M, et al. Brief motivational intervention for intimate partner violence and heavy drinking in the emergency department: a randomized clinical trial. *JAMA* 2015;314:466-77. doi:10.1001/jama.2015.8369 pmid:26241598.
18 Klingberg-Allvin M, Cleeve A, Atuhairwe S, et al. Comparison of treatment of incomplete abortion with misoprostol by physicians and midwives at district level in Uganda: a randomised controlled equivalence trial. *Lancet* 2015;385:2392-8. doi:10.1016/S0140-6736(14)61935-8 pmid:25817472.
19 Broekhuijsen K, van Baaren GJ, van Pampus MG, et al. HYPITAT-II study group. Immediate delivery versus expectant monitoring for hypertensive disorders of pregnancy between 34 and 37 weeks of gestation (HYPITAT-II): an open-label, randomised controlled trial. *Lancet* 2015;385:2492-501. doi:10.1016/S0140-6736(14)61998-X pmid:25817374.
20 Ardehali A, Esmailian F, Deng M, et al. PROCEED II trial investigators. Ex-vivo perfusion of donor hearts for human heart transplantation (PROCEED II): a prospective, open-label, multicentre, randomised non-inferiority trial. *Lancet* 2015;385:2577-84. doi:10.1016/S0140-6736(15)60261-6 pmid:25888086.
21 Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *Lancet* 2015;385:2255-63. doi:10.1016/S0140-6736(15)60461-5 pmid:25771249.
22 Barone MA, Widmer M, Arrowsmith S, et al. Breakdown of simple female genital fistula repair after 7 day versus 14 day postoperative bladder catheterisation: a randomised, controlled, open-label, non-inferiority trial. *Lancet* 2015;386:56-62. doi:10.1016/S0140-6736(14)62337-0 pmid:25911172.
23 Arabi YM, Aldawood AS, Haddad SH, et al. PermiT Trial Group. Permissive underfeeding or standard enteral feeding in critically ill adults. *N Engl J Med* 2015;372:2398-408. doi:10.1056/NEJMoa1502826 pmid:25992505.
24 Senn CY, Eliasziw M, Barata PC, et al. Efficacy of a sexual assault resistance program for university women. *N Engl J Med* 2015;372:2326-35. doi:10.1056/NEJMsa1411131 pmid:26061837.
25 Chagpar AB, Killelea BK, Tsangaris TN, et al. A randomized, controlled trial of cavity shave margins in breast cancer. *N Engl J Med* 2015;373:503-10. doi:10.1056/NEJMoa1504473 pmid:26028131.
26 Torgerson DJ. Efficacy of a sexual assault resistance program for university women. *N Engl J Med* 2015;373:1375-6. doi:10.1056/NEJMc1509345 pmid:26422735.
27 Cox TM, Drelichman G, Cravo R, et al. Eliglustat compared with imiglucerase in patients with Gaucher's disease type 1 stabilised on enzyme replacement therapy: a phase 3, randomised, open-label, non-inferiority trial. *Lancet* 2015;385:2355-62. doi:10.1016/S0140-6736(14)61841-9 pmid:25819691.
28 Altman DG, Bland JM. Treatment allocation by minimisation. *BMJ* 2005;330:843. doi:10.1136/bmj.330.7495.843 pmid:15817555.
29 Berger VW, Exner DV. Detecting selection bias in randomized clinical trials. *Control Clin Trials* 1999;20:319-27. doi:10.1016/S0197-2456(99)00014-8 pmid:10440559.
30 Gerdesmeyer L, Wagenpfeil S, Haake M, et al. Extracorporeal shock wave therapy for the treatment of chronic calcifying tendonitis of the rotator cuff: a randomized controlled trial. *JAMA* 2003;290:2573-80. doi:10.1001/jama.290.19.2573 pmid:14625334.
31 Comer C, Redmond AC, Bird HA, Hensor EMA, Conaghan PG. A home exercise programme is no more beneficial than advice and education for people with neurogenic claudication: results from a randomised controlled trial. *PLoS One* 2013;8:e72878. doi:10.1371/journal.pone.0072878.

Table

Table 1| Trials with inadequate concealment

Study	Randomisation/allocation concealment details	Rationale
The BMJ		
Smith et al, 2015*†[11]	Participants were randomised using a secure web based randomisation system to receive either intervention or routine care. Mixture of block sizes of two and four was used, which was kept secret from recruiting staff	Using a mixture of blocks of two and four leads to an 88% chance of prediction, using the technique of assuming the next allocation was the opposite of the previous one
Smith et al, 2015*†[12]	As above	As above
Detollenare et al, 2015[13]	The women were randomly allocated in a 1:1 ratio using web based, computer generated randomisation tables in blocks of four, stratified by hospital and stage of uterine prolapse	Stratification by centre using blocks of four leads to predictability
JAMA		
Salimen et al, 2015[14]	Patients were randomised using a closed envelope method, in a 1:1 ratio. There were 610 opaque, sealed, and sequentially numbered envelopes. The envelopes were shuffled and distributed to each participating hospital. The surgeon on duty in each hospital opened a consecutively numbered envelope that contained group assignment for the patient. Most of the treating surgeons were not part of the core study team and provided care as they did in their normal clinical practice	Use of envelope opened by recruiting clinician. Report implies that sequentially numbered envelopes were shuffled, which would break numerical sequencing, but the authors later clarified that shuffling was performed before sequential numbering
Stéphan et al, 2015[15]	Randomisation was conducted in blocks of two or four, regardless of entry criteria, with opaque envelopes, using a single computer generated random number sequence for all centres. Attending physicians randomly assigned patients in a 1:1 ratio to one of two groups	Envelopes used with small block sizes with clinicians recruiting patients opening the envelopes. 88% chance of prediction
Polusny et al, 2015[16]	Randomisation was conducted in blocks of four. A restricted electronic randomisation chart was provided to the study coordinator by the statistician	Use of envelopes is unclear. Protocol says that minimisation and envelopes would be used. The paper says blocks of four. Response from authors did not clarify whether envelopes were used. Inadequate due to small block size.
Rhodes et al, 2015[17]	The project coordinator prepared opaque envelopes that were indistinguishable from each other and thick enough so that their contents are not legible from the outside. The research assistant returned to the project office to draw the next sequential envelope from the file, beginning with #1001 for the first participant	Advanced opening of envelopes increases risk of bias
The *Lancet*		
Klingberg-Allvin et al, 2015[18]	Women were randomised to intervention or standard care (1:1) in blocks of 12, stratified by study site, using a computer random number generator. Sequentially numbered, opaque, sealed envelopes, each containing a random allocation, were prepared at the coordinating centre and later opened in consecutive order by the research assistants after obtaining written consent	Use of envelopes is inadequate. It is unclear whether the research assistants prepared and opened the envelopes to recruit participants
Broekhuijsen et al, 2015[19]	Participants were randomised (1:1) with a web based system by random permuted blocks with variable block size (range 2–4), stratified by centre	88% chance of prediction
Ardehali et al, 2015[20]	Participants were randomly assigned (1:1) to receive intervention or standard care. An independent biostatistician prepared sealed and masked randomisation envelopes, which were assigned to each trial site. Participants, investigators, and medical staff were not masked to group allocation	Use of envelopes is inadequate. It is unclear who opened the envelopes—if recruiting clinician then increased risk of bias.
Ngandu et al, 2015[21]	Participants were randomised to intervention or control in a 1:1 ratio. Computer generated allocation was done in blocks of four (two individuals randomly allocated to each group) at each site after baseline by the study nurse	Small block size with centre stratification using an open design allows the risk of bias due to prediction of allocation
Barone et al, 2015[22]	The allocation sequence was computer generated centrally at WHO, and enrolment and randomisation was done by a research assistant based at each study site. Randomisation was in a 1:1 ratio, stratified by country, and restricted with randomly varying block sizes of 4–6. We concealed allocation through sealed opaque envelopes. Randomisation envelopes were opened by study staff just before random assignment	Envelopes opened by staff who were recruiting participants
NEJM		
Arabi et al, 2015[23]	Patients were randomly assigned to intervention or standard care with the use of opaque, sealed, sequentially numbered envelopes. The randomisation list was computer generated. Randomisation was performed in random permuted blocks and was stratified according to centre	Use of envelopes increases risk of bias by advanced opening of envelopes. No description of block size.
Senn et al, 2015[24]	At baseline participants completed a computerised survey, underwent randomisation, and immediately attended their first intervention or control session. Randomisation was performed in permuted blocks of two using an online tool, stratified by site	A block size of two allows a 75% chance of prediction. A block size of two with stratification by three centres should mean, at worst, an imbalance of three participants between groups.

Table 1 (continued)

Study	Randomisation/allocation concealment details	Rationale
Chagpar et al, 2015[25]	Patients were enrolled after written informed consent was obtained, with stratification into one of two groups based on cancer stage. In each stratum, patients were randomly assigned in a 1:1 ratio to intervention or control. Sealed randomisation envelopes were assigned on the basis of a randomisation list generated a priori at the Yale Center for Analytical Sciences. Study personnel were unaware of the assignments. The sealed envelopes were opened intraoperatively after the surgeon completed the partial mastectomy.	Use of envelopes increases risk of bias by advanced opening of envelopes.

*Both trials were conducted at the same time by the same authors.

†Originally unclear but after email correspondence became inadequate.

Appendix 6: paper 4

Envelope use and reporting in randomised controlled trials: a guide for researchers

Appendices for paper 4:

Appendix 1: included studies in paper 4

Reference

Clark, L., Dean, A., Mitchell, A. and Torgerson, D.J., 2021. Envelope use and reporting in randomised controlled trials: A guide for researchers. Research Methods in Medicine & Health Sciences, 2(1), pp.2-11

Original Article

RESEARCH METHODS in MEDICINE & HEALTH SCIENCES

Research Methods in Medicine & Health Sciences
2021, Vol. 2(1) 2–11

DOI: 10.1177/2632084320957204
journals.sagepub.com/home/rmm

Envelope use and reporting in randomised controlled trials: A guide for researchers

Laura Clark, Alexandra Dean, Alex Mitchell and David J Torgerson

Abstract

Introduction: To produce robust evidence RCTs need to be rigorously conducted as poorly performed studies introduce bias and can mislead clinicians and policy makers. Poor allocation concealment has the largest single impact on bias in RCTs than other methodological aspects. Envelopes are frequently used as a method of allocation concealment and can be associated with increased risk of bias. This paper aims to review envelope use in RCTs published in 2017–2018 and create a guide as a reference for researchers when planning and publishing RCTs when using envelopes as an allocation concealment method.

Methods: RCTs that used envelopes as a form of allocation concealment that were published in BMJ, JAMA, NEJM and The Lancet in 2017 and 2018 were identified and methodological data on their envelope use extracted and authors were contacted to ascertain reasons for using envelopes in their research.

Results: 338 RCTs were identified that were published in 2017 and 2018. 8% (n = 29) of the RCTs published used envelopes as an allocation concealment method. 24.1% (n = 7) of studies reported envelope studies robustly with all required methodological information stated to enable an assessment of quality. Budget was the most frequent reason given for envelope use (41.7%).

Discussion: Only 24% of published RCTs, that used envelopes, contained robust methodological information to enable the reader to judge whether the randomisation and allocation concealment method was adequate.

Conclusion: RCTs are not reporting envelope use well. RCTs using envelopes should be designed and reported clearly ensuring all necessary methodological information is included.

Keywords

Validity, reliability, bias, evidence-based medicine, methods and methodology, planning the research, designing a randomised blinded trial, randomised trials, clinical trials, meta-analysis

Introduction

Randomised Controlled Trials (RCTs) are considered to be the gold standard in assessing the effectiveness of interventions. To produce robust evidence RCTs need to be rigorously conducted as poorly performed studies introduce bias and can mislead clinicians and policy makers. Probably the single design element associated with biased findings in trials is poor or absent allocation concealment.[1,2]

Allocation concealment

Allocation concealment is defined as the method used to conceal the randomisation sequence from all study personnel until after the patient has been recruited into the study. This stops the randomisation sequence being subverted and the study having a high risk of bias. It has been shown that having an inadequate allocation concealment method can exaggerate the effect size by 41%.[1,2] There are multiple ways that the randomisation sequence can be concealed, such as web-based or

Department of Health Sciences, University of York, York, UK

Corresponding author:
Laura Clark, Department of Health Sciences, University of York, ARRC Building, York, YO10 5DD, UK.
Email: laura.clark@york.ac.uk

telephone systems. Traditionally, before web and telephone systems were available envelopes were used. The use of sealed envelopes as a method still lingers on as a concealment method for a significant proportion of RCTs. For instance, Yelland et al found in 2015 that 9% of RCTs employed sealed envelopes as a method of concealment.[3]

Advantages and disadvantages of envelope use as a method of allocation concealment

There are significant disadvantages to using envelopes for allocation concealment. They can be opened in advance for example,[4] trans illumination can determine the allocation[5–8] such methods allow subversion of the randomisation. On the other hand they are relatively cheap and logistically practical in remote areas that are internet or telephone free or in emergency medicine situations.

In this paper we aim to describe the types of trials that continue to use envelopes and the quality of the envelope concealment used and to provide advice on their safer usage.

Methods

RCTs published in BMJ, JAMA, NEJM and The Lancet in 2017 and 2018 were identified. Two reviewers extracted data from each paper on the randomisation and allocation concealment methods. Those RCTs that used envelopes to conceal the randomisation sequence were identified.

Envelope concealment

We used the approach described by Doig and Simpson[9] to define high quality envelope concealment. There are three areas that were assessed as follows:

1. If the person who created the envelope was stated

Best practice for the use of envelopes in RCTs would be that a randomisation sequence would be generated and personal not involved in the RCT would create the envelopes for the RCT.

2. Whether the envelopes had an additional security measure.

Envelopes should have an additional security measure rather than just being closed and be opaque and sequentially numbered. This order can then be checked and anomalies will be identified if the randomisation sequence has been violated. Other additional security measures include the person who has created the envelope signing the back of the envelope when sealed so it is obvious if it has been tampered with. Inserting foil and or carbon paper into the envelope prevents trans-illumination and the carbon paper allows an additional audit trail as the participants name and date of recruitment can be written on the envelope at the point of recruitment before the envelope has been opened and the carbon paper prints this information to the allocation insert. The envelopes should be kept securely and not with the research team who are responsible for recruiting participants into the study.

3. If the person who opened the envelope (recruited participants) was stated

The person who created the envelope should not be the same person who recruits participants to prevent the ordering of participants into one treatment arm or another.

We identified these and other quality factors and extracted them from each RCT that used envelopes: the envelope description, whether who created the envelope and who opened the envelope was stated. We emailed each corresponding author of the RCTs that used envelopes as an allocation concealment method and inquired as to why they have chosen envelopes as an allocation concealment method.

Results

A total of 338 RCTs were identified that were published in 2017 and 2018. 7.5% of RCTs published in 2017 used envelopes and 9.5% in 2018. Combined, 8.6% (n = 29) of the RCTs published in 2017 and 2018 used envelopes as an allocation concealment method. We emailed each author of the RCTs using envelopes and received responses from 12 (41%) of them. Figure 1 shows the flow of studies.

Table 1 shows the description from the paper of each envelope trial with the necessary quality factors and trial setting. 24.1% (n = 7) of studies that use envelopes for allocation concealment, reported envelope use robustly with all required methodological information stated to enable an assessment of quality. 44.8% (n = 13) reported who created the envelope, 44.8% (n = 13) reported who opened the envelope and 62.1% (n = 18) reported the envelope description adequately.

Table 2 shows the author stated reasons for envelope use, it was found that the most frequent reason given was budget (41.7%).

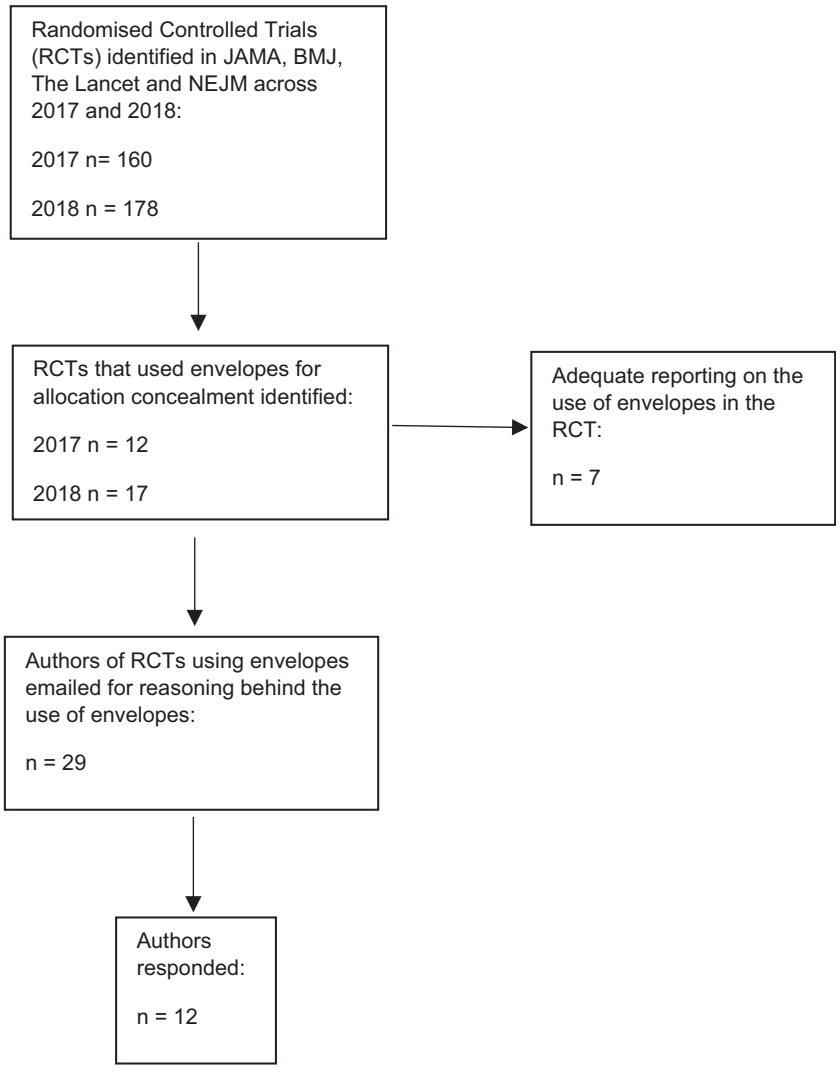


Figure 1. Flow of trials through study.

Discussion

Envelope use is similar to 2015 where Yelland found, in a similar group of journals, that 9% of RCTs used envelopes as a form of allocation concealment [2]. Reasons for envelope use were all appropriate for the trial design and setting that the research was being conducted.

It was disappointing that only 24% of the published RCTs in this sample contained robust methodological information to enable the reader to judge the RCT as adequate and low risk of bias when assessing the randomisation and allocation concealment methods. This therefore means that 76% of RCTs cannot enter into systematic reviews with a low risk of bias, there will be a higher level of uncertainty of the validity of the systematic review. All resources used to perform this research is wasted as the published report is not clear.

Many envelopes were not described as having any additional security measures. They were simply stated as being 'sealed', they may have been sequentially numbered however this was not stated so it cannot be assumed. Only one study (Boden – see Appendix 1 for a list of all included studies) stated that they used very secure envelopes with the addition of foil to wrap the allocation cards within the envelopes. Foil prevents the trans-illumination of the envelopes and further protects an RCT from subversion.[1,9] It has been found that RCTs employing the use of envelopes without additional security measures are associated with an exaggerated effect size.[10] Results from insecure envelopes will be treated with caution by policy makers – thus highlighting the important of using secure envelopes and ensuring that if secure envelopes were used the details are reported comprehensively to enable policy makers to have confidence in the reported

Table 1. Quality factors and setting reported in included RCTs.

Author	Envelope description	Additional security measures stated on envelope	Person who created envelope stated	Who opened envelope stated	Envelope descriptions and methods adequately reported	Setting
Andrews	Sealed opaque envelopes	Non reported	No	No	No	Emergency department
Smits	Closed, opaque envelopes	Non reported	No	No	No	Hospital
Landoni	sealed, opaque, sequentially numbered envelopes	Sequentially numbered	No	No	No	Operating theatre or ICU.
Dwivedi	sequentially numbered, sealed, opaque envelopes	Sequentially numbered	By persons not involved in the trial	No	No	Referral center for epilepsy surgery in northern India, New Delhi
Kulkarni	sealed, opaque envelopes,	Non reported	no	No – 'opened in the preoperative holding area just before the patient entered the operating room'	No	Operating theatre in Uganda
Kaufman	Opaque envelopes concealing the allocation, within sealed individual study packs.	Study packs were kept available from a locked study box from which they could only be taken sequentially.	Yes: independent statistician	No	No	A tertiary paediatric emergency department (Melbourne)
Patel	sealed in sequential numbered opaque envelopes	sequential numbered	Independent support staff	trained health assistants	Yes	Primary health centres in Goa
Nadkarni	sequential numbered opaque sealed envelopes	Sequential numbered	No	trained health assistants based	No	Primary health centres in Goa
Brockman	sequentially assigned sealed randomisation envelopes,	Sequentially assigned	Person independent of the research team and who had no further role in the trial	Envelopes were opened immediately before induction of anaesthesia by the attending anaesthetist	Yes	Department of Anaesthesia and Pain Management, Perth
Chan	sealed opaque envelopes	Non reported	independent staff member assigned the treatments	no	No	Prince of Wales Hospital of The Chinese University of Hong Kong

(continued)

Table 1. Continued.

Author	Envelope description	Additional security measures stated on envelope	Person who created envelope stated	Who opened envelope stated	Envelope descriptions and methods adequately reported	Setting
Mundle	sequentially numbered, sealed, opaque envelope	sequentially numbered	The envelopes were generated by staff at Gynuity Health Projects	Research staff	Yes	Two public hospitals in India
Stocker[a]	sequentially numbered sealed opaque envelopes.	Sequentially numbered	no	no	No	Hospital - Neonates born >34 weeks who had suspected early-onset sepsis in the first 72 h of life and who required antibiotic therapy
Molloy	sequentially drawing sealed envelopes	Sequential	No	Yes: trial pharmacist and clinician	No	9 African hospitals
Franklin	Sequentially numbered, sealed, opaque envelopes	Sequentially numbered, sealed, opaque envelopes	No	No	No	Emergency departments and general pediatric inpatient units in 17 tertiary and regional hospitals in Australia and New Zealand
Boden	sequentially numbered sealed opaque envelopes containing allocation cards wrapped in aluminium foil.	Sequentially numbered, sealed opaque envelopes. Foil used to wrap allocation cards. Patient details were marked on envelopes to record that randomisation was in order of recruitment.	Yes: independent administrator	Yes: physiotherapist	Yes	Multidisciplinary preadmission clinics at three tertiary public hospitals in Australia and New Zealand.
Firanescu	sealed randomisation envelope	Non reported	No	No	No	Four community hospitals in the Netherlands
Mason	opaque, sequentially numbered, sealed envelopes	Opaque, sequentially numbered	No	No	No	Recruitment from workplaces, social media platforms, and schools in Birmingham UK

(continued)

Table 1. Continued.

Author	Envelope description	Additional security measures stated on envelope	Person who created envelope stated	Who opened envelope stated	Envelope descriptions and methods adequately reported	Setting
Salimen	opaque, sealed, and sequentially numbered randomization envelopes were shuffled and then distributed to each participating hospital	opaque, sealed, and sequentially numbered randomization envelopes	No	Yes: surgeon	No	Three hospitals in Finland
Peterli	sealed envelopes	Non reported	No	No	No	Four bariatric centres in Switzerland
Jabre	sealed envelopes	Non reported	No – implied that it is not the same person who opened the envelope	No	No	20 prehospital emergency medical services (EMS) centers: 15 in France and 5 in Belgium.
Labhart	sealed, sequentially numbered, opaque envelopes.	Sealed Sequentially numbered	Yes: 'a separate person not involved in the study'	Yes: study nurse	Yes	Six health care facilities in northern Lesotho.
Huttner	Opaque sealed envelopes	Non reported	No	No	No	hospital units and outpatient clinics in Switzerland, Poland and Israel.
Driver	sequentially numbered, opaque envelopes	sequentially numbered, opaque envelopes	No	Yes: research associate	No	Emergency department at Hennepin County Medical Center, anurban,
Fossat	sealed, opaque, and numbered envelopes	sealed, opaque, and numbered envelopes	Yes: a clinical research assistant	Yes: investigator	Yes	ICU (hospital) in France
Cooper	Sealed opaque envelopes and permuted variable block sizes (2 and 4).	Sealed opaque envelopes and permuted variable block sizes (2 and 4).	No	No	No	Patients both out-of-hospital and in emergency departments in Australia, New Zealand, France, Switzerland, Saudi Arabia, and Qatar

(continued)

Table 1. Continued.

Author	Envelope description	Additional security measures stated on envelope	Person who created envelope stated	Who opened envelope stated	Envelope descriptions and methods adequately reported	Setting
Montaigne	The code sequence was computer generated and kept in sealed envelopes at a central location by non-medical staff not involved in the study.	Non reported	No	Yes: staff cardiologists	No	Hospital in France
Farquhar	Allocations were concealed in sequentially numbered, sealed, opaque envelopes	sequentially numbered, sealed, opaque envelopes	Yes: an independent statistician	Yes: study coordinator	Yes	Fertility clinics in New Zealand
Heinemann	Each study site received sealed envelopes with the respective group allocation. After successful completion of the baseline phase, the respective envelope was opened. .	Non reported	Not explicit but states 'Randomisation was done centrally at the study coordinating centre by staff who were not involved with recruitment or treatment of study participants'	No	No	Diabetes practices in Germany
Blumberger	The randomisation tables were used by staff outside the study team to produce opaque, sealed envelopes, labelled with a participant specific randomisation identification number and containing a treatment allocation	Non reported	Yes: staff outside the study team	Yes: study staff	No	Three hospitals in Germany

[a]Two types of randomisation – SNOSE and drawing cards at random in different centres.

Table 2. Author stated reasons for envelope use.

Author	Author stated reasons for envelope use														
	Budget	Can randomise quickly	Removes issues of equipment malfunction	Maintain optimal blindness	Geographical Location	Unreliable internet and phone signal	Pilot of SMS service highlighted error, SNOSE chosen instead	Rapid randomisation needed due to setting and intervention[a]	Size of study	No telephone or online system available at start of study recruitment	Ease of use	Setting	Practical	Pragmatic within study design
Bellomo	●	●	●											
Patel and Nadkarni				●			●							
Weeks					●	●								
van Herk								●	●					
Salminen	●									●	●			
Jabre								●				●		
Huttner	●												●	
Driver								●						
Farquhar	●													
Heinemann													●	●
Blumberger	●													

[a]One study stated within 1-2 minutes and another <12 hour.

results. The use of foil with sequentially numbered sealed opaque envelopes in the opinion of the authors of this paper should be the gold standard way to set up an envelope for use within an RCT.

One point of interest is there are two RCTs in this sample that describe the same methodology from the large scale RCT but are reporting different results. Interestingly one study was deemed methodologically robust (Patel) and one was not (Nadkarni) as they did not report who created the envelope. This small omission has resulted in a study not being classed as having robust methodology.

There are a variety of valid reasons why envelopes are used within RCTs. We would urge researchers to carefully consider their research budget and assess whether they should allocate additional funds to cover the cost of a more robust and secure randomisation and allocation concealment method if their research design allows. Envelopes are inexpensive but if not executed and published robustly the entire research cannot contribute meaningfully to the evidence base.

After analysing the data gained in this research we can see that there is still a two-fold issue with envelope use. Envelopes are not being prepared in a rigorous manner with additional security measures and they are not reported in a transparent robust way ensuring all methodological information is provided. There is unclear information given to ascertain whether the envelope had additional security measures and whether the person creating the envelope is separate to the person who opens the envelope at the point of recruitment.

Future recommendations

Figure 2 shows the recommendation we have to the following when performing research with envelopes as an allocation concealment method.

It is also pertinent to discuss the evolving nature of technology and allocation concealment methods. For rapid randomisation envelopes are a sensible choice, there are however apps being created that can rapidly randomise participants. These are also a relatively inexpensive method of randomisation and allocation concealment and may be used more widely in the future. Even with the use of innovative technology methodology will still need to be published thoroughly for methodological quality judgements to be made.

Moving forward we would urge authors to plan for and create secure envelopes when using envelopes as a form of allocation concealment are the only option for their RCT and to write their research transparently to include all the methodological information stated in Figure 2. Journals should

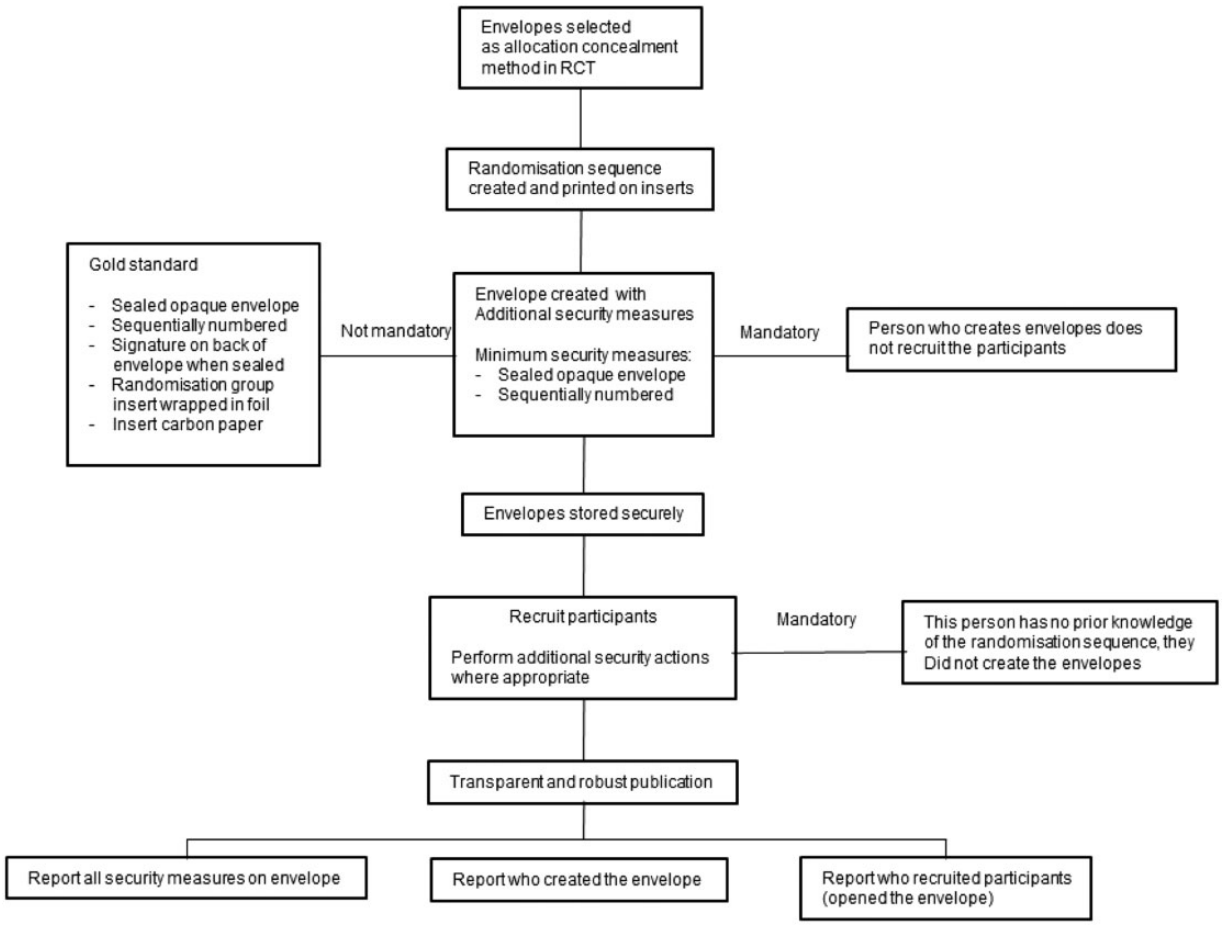


Figure 2. Flow diagram stating methodological steps to creating a robust envelope as an allocation concealment method in an RCT and essential methodological information to be reported in the publication of an RCT.

ensure that any RCT published that uses envelopes as a form of allocation concealment should be reported robustly.

Conclusions

Allocation concealment methods are one of the most influential methodological factors on the validity of an RCT. Envelopes can be used as a robust method of allocation concealment. However, they are the most insecure method associated with subverting an RCT. If they are used within a research design they should be created robustly and reported clearly ensuring all necessary methodological information is included.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Laura Clark https://orcid.org/0000-0001-9227-5447
Alexandra Dean https://orcid.org/0000-0003-4711-5609
Alex Mitchell https://orcid.org/0000-0001-9311-2092
David J Torgerson https://orcid.org/0000-0002-1667-4275

References

1. Swingler GH and Zwarenstein M. An effectiveness trial of a diagnostic test in a busy outpatients department in a developing country Issues around allocation concealment and envelope randomization. *Journal of clinical epidemiology*t 53(7): 702–706.
2. Schulz KF, Chalmers I, Hayes RJ, et al. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; 273: 408–412.

3. Yelland LN, Kahan BC, Dent E, et al. Prevalence and reporting of recruitment, randomisation and treatment errors in clinical trials: a systematic review. *Clin Trials* 2018; 15: 278–285.
4. Kennedy AD, Torgerson DJ, Campbell MK, et al. Subversion of allocation concealment in a randomised controlled trial: a historical case study. *Trials* 2017; 18: 204.
5. Schulz K. Unbiased research and the human spirit: the challenges of randomized controlled trials. *CMAJ* 1995; 153: 783–786.
6. Schulz KF. Subverting randomization in controlled trials. *JAMA* 1995; 274: 1456–1458.
7. Schulz KF. Randomised trials, human nature, and reporting guidelines. *Lancet* 1996; 348: 596–598.
8. Schulz KF, Chalmers I, Hayes RJ, et al. Empirical evidence of bias. *JAMA* 1995; 273: 408.
9. Doig GS and Simpson F. Randomization and allocation concealment: a practical guide for researchers. *J Crit Care* 2005; 20: 187–191.
10. Herbison P, Hay-Smith J and Gillespie WJ. Different methods of allocation to groups in randomized trials are associated with different levels of bias. A meta-epidemiological study. *J Clin Epidemiol* 2011; 64: 1070–1075.

Appendix 1: included studies

ANDREWS, B., SEMLER, M. W., MUCHEMWA, L., KELLY, P., LAKHI, S., HEIMBURGER, D. C., MABULA, C., BWALYA, M. & BERNARD, G. R. 2017. Effect of an Early Resuscitation Protocol on In-hospital Mortality Among Adults With Sepsis and Hypotension: A Randomized Clinical TrialEffects of an Early Resuscitation Protocol on Sepsis Mortality in ZambiaEffects of an Early Resuscitation Protocol on Sepsis Mortality in Zambia. *JAMA,* 318**,** 1233-1240.

BLUMBERGER, D. M., VILA-RODRIGUEZ, F., THORPE, K. E., FEFFER, K., NODA, Y., GIACOBBE, P., KNYAHNYTSKA, Y., KENNEDY, S. H., LAM, R. W. & DASKALAKIS, Z. J. 2018. Effectiveness of theta burst versus high-frequency repetitive transcranial magnetic stimulation in patients with depression (THREE-D): a randomised non-inferiority trial. *The Lancet,* 391**,** 1683-1692.

BODEN, I., SKINNER, E. H., BROWNING, L., REEVE, J., ANDERSON, L., HILL, C., ROBERTSON, I. K., STORY, D. & DENEHY, L. 2018. Preoperative physiotherapy for the prevention of respiratory complications after upper abdominal surgery: pragmatic, double blinded, multicentre randomised controlled trial. *BMJ,* 360**,** j5916.

CHAN, F. K., CHING, J. Y., TSE, Y. K., LAM, K., WONG, G. L., NG, S. C., LEE, V., AU, K. W., CHEONG, P. K. & SUEN, B. Y. 2017. Gastrointestinal safety of celecoxib versus naproxen in patients with cardiothrombotic diseases and arthritis after upper gastrointestinal bleeding (CONCERN): an industry-independent, double-blind, double-dummy, randomised trial. *The Lancet,* 389**,** 2375-2382.

COOPER, D. J., NICHOL, A. D., BAILEY, M., BERNARD, S., CAMERON, P. A., PILI-FLOURY, S., FORBES, A., GANTNER, D., HIGGINS, A. M., HUET, O., KASZA, J., MURRAY, L., NEWBY, L., PRESNEILL, J. J., RASHFORD, S., ROSENFELD, J. V., STEPHENSON, M., VALLANCE, S., VARMA, D., WEBB, S. A. R., TRAPANI, T., MCARTHUR, C., INVESTIGATORS, F. T. P. T. & GROUP, T. A. C. T. 2018. Effect of Early Sustained Prophylactic Hypothermia on Neurologic Outcomes Among Patients With Severe Traumatic Brain Injury: The POLAR Randomized Clinical TrialEffect of Early Sustained Prophylactic Hypothermia on Neurologic Outcomes in Patients With Severe TBIEffect of Early Sustained Prophylactic Hypothermia on Neurologic Outcomes in Patients With Severe TBI. *JAMA,* 320**,** 2211-2220.

DRAKE-BROCKMAN, T. F., RAMGOLAM, A., ZHANG, G., HALL, G. L. & VON UNGERN-STERNBERG, B. S. 2017. The effect of endotracheal tubes versus laryngeal mask airways on perioperative respiratory adverse events in infants: a randomised controlled trial. *The Lancet,* 389**,** 701-708.

DRIVER, B. E., PREKKER, M. E., KLEIN, L. R., REARDON, R. F., MINER, J. R., FAGERSTROM, E. T., CLEGHORN, M. R., MCGILL, J. W. & COLE, J. B. 2018. Effect of Use of a Bougie vs Endotracheal Tube and Stylet on First-Attempt Intubation Success Among Patients With Difficult Airways Undergoing Emergency Intubation: A Randomized Clinical TrialComparison of Intubation Techniques in Patients With Difficult AirwaysComparison of Intubation Techniques in Patients With Difficult Airways. *JAMA,* 319**,** 2179-2189.

DWIVEDI, R., RAMANUJAM, B., CHANDRA, P. S., SAPRA, S., GULATI, S., KALAIVANI, M., GARG, A., BAL, C. S., TRIPATHI, M., DWIVEDI, S. N., SAGAR, R., SARKAR, C. & TRIPATHI, M. 2017. Surgery for Drug-Resistant Epilepsy in Children. *New England Journal of Medicine,* 377**,** 1639-1647.

FARQUHAR, C. M., LIU, E., ARMSTRONG, S., ARROLL, N., LENSEN, S. & BROWN, J. 2018. Intrauterine insemination with ovarian stimulation versus expectant management for unexplained infertility (TUI): a pragmatic, open-label, randomised, controlled, two-centre trial. *The Lancet,* 391**,** 441-450.

FIRANESCU, C. E., DE VRIES, J., LODDER, P., VENMANS, A., SCHOEMAKER, M. C., SMEET, A. J., DONGA, E., JUTTMANN, J. R., KLAZEN, C. A. H., ELGERSMA, O. E. H., JANSEN, F. H., TIELBEEK, A. V., BOUKRAB, I., SCHONENBERG, K., VAN ROOIJ, W. J. J., HIRSCH, J. A. & LOHLE, P. N. M. 2018. Vertebroplasty versus sham procedure for painful acute osteoporotic vertebral compression fractures (VERTOS IV): randomised sham controlled clinical trial. *BMJ,* 361**,** k1551.

FOSSAT, G., BAUDIN, F., COURTES, L., BOBET, S., DUPONT, A., BRETAGNOL, A., BENZEKRI-LEFÈVRE, D., KAMEL, T., MULLER, G., BERCAULT, N., BARBIER, F., RUNGE, I., NAY, M.-A., SKARZYNSKI, M., MATHONNET, A. & BOULAIN, T. 2018. Effect of In-Bed Leg Cycling and Electrical Stimulation of the Quadriceps on Global Muscle Strength in Critically Ill Adults: A Randomized Clinical TrialIn-Bed Leg Cycling and Electrical Stimulation of the Quadriceps on Global Muscle Strength in Critically Ill AdultsIn-Bed Leg Cycling and Electrical Stimulation of the Quadriceps on Global Muscle Strength in Critically Ill Adults. *JAMA,* 320**,** 368-378.

FRANKLIN, D., BABL, F. E., SCHLAPBACH, L. J., OAKLEY, E., CRAIG, S., NEUTZE, J., FURYK, J., FRASER, J. F., JONES, M., WHITTY, J. A., DALZIEL, S. R. & SCHIBLER, A. 2018. A Randomized Trial of High-Flow Oxygen Therapy in Infants with Bronchiolitis. *New England Journal of Medicine,* 378**,** 1121-1131.

HEINEMANN, L., FRECKMANN, G., EHRMANN, D., FABER-HEINEMANN, G., GUERRA, S., WALDENMAIER, D. & HERMANNS, N. 2018. Real-time continuous glucose monitoring in adults with type 1 diabetes and impaired hypoglycaemia awareness or severe hypoglycaemia treated with multiple daily insulin injections (HypoDE): a multicentre, randomised controlled trial. *The Lancet,* 391**,** 1367-1377.

HUTTNER, A., KOWALCZYK, A., TURJEMAN, A., BABICH, T., BROSSIER, C., ELIAKIM-RAZ, N., KOSIEK, K., MARTINEZ DE TEJADA, B., ROUX, X., SHIBER, S., THEURETZBACHER, U., VON DACH, E., YAHAV, D., LEIBOVICI, L., GODYCKI-ĆWIRKO, M., MOUTON, J. W. & HARBARTH, S. 2018. Effect of 5-Day Nitrofurantoin vs Single-Dose Fosfomycin on Clinical Resolution of Uncomplicated Lower Urinary Tract Infection in Women: A Randomized Clinical TrialEffect of Nitrofurantoin vs Fosfomycin on Uncomplicated Lower UTI in WomenEffect of Nitrofurantoin vs Fosfomycin on Uncomplicated Lower UTI in Women. *JAMA,* 319**,** 1781-1789.

JABRE, P., PENALOZA, A., PINERO, D., DUCHATEAU, F.-X., BORRON, S. W., JAVAUDIN, F., RICHARD, O., DE LONGUEVILLE, D., BOUILLEAU, G., DEVAUD, M.-L., HEIDET, M., LEJEUNE, C., FAUROUX, S., GREINGOR, J.-L., MANARA, A., HUBERT, J.-C., GUIHARD, B., VERMYLEN, O., LIEVENS, P., AUFFRET, Y., MAISONDIEU, C., HUET, S., CLAESSENS, B., LAPOSTOLLE, F., JAVAUD, N., REUTER, P.-G., BAKER, E., VICAUT, E. & ADNET, F. 2018. Effect of Bag-Mask Ventilation vs Endotracheal Intubation During Cardiopulmonary Resuscitation on Neurological Outcome After Out-of-Hospital Cardiorespiratory Arrest: A Randomized Clinical TrialEffect on Neurologic Outcomes of Bag-Mask vs Mechanical Ventilation During OHCAEffect on Neurologic Outcomes of Bag-Mask vs Mechanical Ventilation During OHCA. *JAMA,* 319**,** 779-787.

KAUFMAN, J., FITZPATRICK, P., TOSIF, S., HOPPER, S. M., DONATH, S. M., BRYANT, P. A. & BABL, F. E. 2017. Faster clean catch urine collection (Quick-Wee method) from infants: randomised controlled trial. *BMJ,* 357**,** j1341.

KULKARNI, A. V., SCHIFF, S. J., MBABAZI-KABACHELOR, E., MUGAMBA, J., SSENYONGA, P., DONNELLY, R., LEVENBACH, J., MONGA, V., PETERSON, M., MACDONALD, M., CHERUKURI, V. & WARF, B. C. 2017. Endoscopic Treatment versus Shunting for Infant Hydrocephalus in Uganda. *New England Journal of Medicine,* 377**,** 2456-2464.

LABHARDT, N. D., RINGERA, I., LEJONE, T. I., KLIMKAIT, T., MUHAIRWE, J., AMSTUTZ, A. & GLASS, T. R. 2018. Effect of Offering Same-Day ART vs Usual Health Facility Referral During Home-Based HIV Testing on Linkage to Care and Viral Suppression Among Adults With HIV in Lesotho: The CASCADE Randomized Clinical TrialEffect on Viral Suppression of Same-Day ART vs Care Referral After Home HIV Testing in LesothoEffect on Viral Suppression of Same-Day ART vs Care Referral After Home HIV Testing in Lesotho. *JAMA,* 319**,** 1103-1112.

LANDONI, G., LOMIVOROTOV, V. V., ALVARO, G., LOBREGLIO, R., PISANO, A., GUARRACINO, F., CALABRÒ, M. G., GRIGORYEV, E. V., LIKHVANTSEV, V. V., SALGADO-FILHO, M. F., BIANCHI, A., PASYUGA, V. V., BAIOCCHI, M., PAPPALARDO, F., MONACO, F., BOBOSHKO, V. A., ABUBAKIROV, M. N., AMANTEA, B., LEMBO, R., BRAZZI, L., VERNIERO, L., BERTINI, P., SCANDROGLIO, A. M., BOVE, T., BELLETTI, A., MICHIENZI, M. G., SHUKEVICH, D. L., ZABELINA, T. S., BELLOMO, R. & ZANGRILLO, A. 2017. Levosimendan for Hemodynamic Support after Cardiac Surgery. *New England Journal of Medicine,* 376**,** 2021-2031.

MASON, F., FARLEY, A., PALLAN, M., SITCH, A., EASTER, C. & DALEY, A. J. 2018. Effectiveness of a brief behavioural intervention to prevent weight gain over the Christmas holiday period: randomised controlled trial. *BMJ,* 363**,** k4867.

MOLLOY, S. F., KANYAMA, C., HEYDERMAN, R. S., LOYSE, A., KOUANFACK, C., CHANDA, D., MFINANGA, S., TEMFACK, E., LAKHI, S., LESIKARI, S., CHAN, A. K., STONE, N., KALATA, N., KARUNAHARAN, N., GASKELL, K., PEIRSE, M., ELLIS, J., CHAWINGA, C., LONTSI, S., NDONG, J.-G., BRIGHT, P., LUPIYA, D., CHEN, T., BRADLEY, J., ADAMS, J., VAN DER HORST, C., VAN OOSTERHOUT, J. J., SINI, V., MAPOURE, Y. N., MWABA, P., BICANIC, T., LALLOO, D. G., WANG, D., HOSSEINIPOUR, M. C., LORTHOLARY, O., JAFFAR, S. & HARRISON, T. S. 2018. Antifungal Combinations for Treatment of Cryptococcal Meningitis in Africa. *New England Journal of Medicine,* 378**,** 1004-1017.

MONTAIGNE, D., MARECHAL, X., MODINE, T., COISNE, A., MOUTON, S., FAYAD, G., NINNI, S., KLEIN, C., ORTMANS, S. & SEUNES, C. 2018. Daytime variation of perioperative myocardial injury in cardiac surgery and its prevention by Rev-Erbα antagonism: a single-centre propensity-matched cohort study and a randomised study. *The Lancet,* 391**,** 59-69.

MUNDLE, S., BRACKEN, H., KHEDIKAR, V., MULIK, J., FARAGHER, B., EASTERLING, T., LEIGH, S., GRANBY, P., HAYCOX, A. & TURNER, M. A. 2017. Foley catheterisation versus oral misoprostol for induction of labour in hypertensive women in India (INFORM): a multicentre, open-label, randomised controlled trial. *The Lancet,* 390**,** 669-680.

NADKARNI, A., WEOBONG, B., WEISS, H. A., MCCAMBRIDGE, J., BHAT, B., KATTI, B., MURTHY, P., KING, M., MCDAID, D. & PARK, A.-L. 2017. Counselling for Alcohol Problems (CAP), a lay counsellor-delivered brief psychological treatment for harmful drinking in men, in primary care in India: a randomised controlled trial. *The Lancet,* 389**,** 186-195.

PATEL, V., WEOBONG, B., WEISS, H. A., ANAND, A., BHAT, B., KATTI, B., DIMIDJIAN, S., ARAYA, R., HOLLON, S. D. & KING, M. 2017. The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: a randomised controlled trial. *The Lancet,* 389**,** 176-185.

PETERLI, R., WÖLNERHANSSEN, B. K., PETERS, T., VETTER, D., KRÖLL, D., BORBÉLY, Y., SCHULTES, B., BEGLINGER, C., DREWE, J., SCHIESSER, M., NETT, P. & BUETER, M. 2018. Effect of Laparoscopic Sleeve Gastrectomy vs Laparoscopic Roux-en-Y Gastric Bypass on Weight Loss in Patients With Morbid Obesity: The SM-BOSS Randomized Clinical TrialEffect of Sleeve Gastrectomy vs Roux-en-Y

Gastric Bypass on Morbid ObesityEffect of Sleeve Gastrectomy vs Roux-en-Y Gastric Bypass on Morbid Obesity. *JAMA,* 319**,** 255-265.

SALMINEN, P., HELMIÖ, M., OVASKA, J., JUUTI, A., LEIVONEN, M., PEROMAA-HAAVISTO, P., HURME, S., SOINIO, M., NUUTILA, P. & VICTORZON, M. 2018. Effect of Laparoscopic Sleeve Gastrectomy vs Laparoscopic Roux-en-Y Gastric Bypass on Weight Loss at 5 Years Among Patients With Morbid Obesity: The SLEEVEPASS Randomized Clinical TrialFive-Year Outcomes of Laparoscopic Sleeve Gastrectomy vs Gastric Bypass for Morbid ObesityFive-Year Outcomes of Laparoscopic Sleeve Gastrectomy vs Gastric Bypass for Morbid Obesity. *JAMA,* 319**,** 241-254.

SMITS, P. C., ABDEL-WAHAB, M., NEUMANN, F.-J., BOXMA-DE KLERK, B. M., LUNDE, K., SCHOTBORGH, C. E., PIROTH, Z., HORAK, D., WLODARCZAK, A., ONG, P. J., HAMBRECHT, R., ANGERÅS, O., RICHARDT, G. & OMEROVIC, E. 2017. Fractional Flow Reserve–Guided Multivessel Angioplasty in Myocardial Infarction. *New England Journal of Medicine,* 376**,** 1234-1244.

STOCKER, M., VAN HERK, W., EL HELOU, S., DUTTA, S., FONTANA, M. S., SCHUERMAN, F. A., VAN DEN TOOREN-DE, R. K., WIERINGA, J. W., JANOTA, J. & VAN DER MEER-KAPPELLE, L. H. 2017. Procalcitonin-guided decision making for duration of antibiotic therapy in neonates with suspected early-onset sepsis: a multicentre, randomised controlled trial (NeoPIns). *The Lancet,* 390**,** 871-881.

Appendix 7: paper 5

A review found small variable blocking schemes may not protect against selection bias in randomized controlled trials

Appendices for paper 5:

Appendix A: included studies in age analysis (web only material)

Reference

Clark, L., Burke, L., Carr, R.M., Coleman, E., Roberts, G. and Torgerson, D.J., 2022. A review found small variable blocking schemes may not protect against selection bias in randomized controlled trials. Journal of Clinical Epidemiology, 141, pp.90-98.

Journal of Clinical Epidemiology 141 (2022) 90–98

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

A review found small variable blocking schemes may not protect against selection bias in randomized controlled trials

Laura Clark*, Lauren Burke, Rachel Margaret Carr, Elizabeth Coleman, Gareth Roberts, David J. Torgerson

York Trials Unit, Department of Health Sciences, University of York, York, YO10 5DD, United Kingdom

Accepted 7 September 2021; Available online 11 September 2021

Abstract

Objective: Blocking is associated with prediction of the allocation sequence and subversion. This paper explores if blocking strategies lead to an increase in baseline age heterogeneity (a marker for potential subversion) and, whether the use of blocking is changing over time.

Study Design and Settings: The British Medical Journal, Journal of the American Medical Association, The Lancet and the New England Journal of Medicine were hand searched to identify open RCTs published in January between 2001 and 2020. To explore heterogeneity of baseline age meta-analyses were performed on trials implementing blocking, minimization, and simple randomization.

Results: One hundred seventy-nine open RCTs were identified: nine (5.0%) undertook simple randomization, 104 (58.1%) blocking, 25 (13.9%) minimization, and one (0.6%) both. Baseline age heterogeneity of 24% (P= 0.02) was observed in all trials implementing blocking, 62% (P = 0.001) in trials implementing a fixed block of four, 40% (P = 0.07) implementing variable blocks including a 2 and 0% for both simple randomization and minimization. Small block sizes are implemented in modern trials.

Conclusion: Variable block sizes including two are associated with subversion and should not be implemented. If center only stratification is necessary, it should be used alongside larger blocking schemes. Authors should consider alternative methods to restrict randomization.

Key Words: Research design; Bias; Allocation concealment; Randomization; Randomized controlled trials; Methodology

What is new?

Key findings

- Recently published trials are observed to be implementing blocking including a block size of two and stratifying by center.
- Increased heterogeneity of baseline age – an indication of subversion – is observed in trials implementing a variable blocking scheme including a block size of two.
- Avoiding small fixed block sizes and using large variable blocking schemes are recommended to safeguard against prediction of allocation sequence in randomized controlled trials.

What this study adds to what was known

- Variable block sizes that include a block size of two are not a safeguard against subversion as they are associated with moderate heterogeneity of baseline age, both when center stratification has been performed and without.

What is the implication and what should change now?

- If blocking is to be implemented it should be done so with larger blocking schemes that do not include a block size of two.

1. Introduction

Randomized controlled trials (RCTs) are considered the gold standard to assess a difference of effect between treatment groups [1,2]. They need to be designed to minimize bias, including selection bias, which can ensue if the

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

The authors have no relevant conflicts of interest.

* Corresponding author.
E-mail address: laura.clark@york.ac.uk (L. Clark).

https://doi.org/10.1016/j.jclinepi.2021.09.009
0895-4356/

method used to conceal the allocation sequence is inadequate or predictable [3]. Heterogeneity, in meta-analyses, of baseline variables, such as age, is an indicator that some of the included trials have been subverted and therefore impacted by selection bias [4–6]. Such heterogeneity may be as a result of using 'restricted randomization'. Because of perceived problems with 'simple' or unrestricted randomisation most RCTs use a form of restricted randomization, which maintains balance within the arms, and for specified covariate, during participant recruitment [7]. Indeed, approximately 90% of trials published in major clinical journals use some form of restriction in their randomization process [7].

The use of restricted randomization in RCTs has long been implemented [8]. The reasons for restricted randomization include the following: for small trials ($n < 100$) it creates numerically balanced treatment groups which improve statistical efficiency. For larger trials it also leads to balance on stratifying variables, such as treatment center, and where trials are slow to recruit it also avoids chance temporal effects [9]. It can also avoid a streak of the same allocation occurring in a row, which may be useful for logistical reasons (e.g., planning treatment slots) [10]. However, restricted randomization often increases the risk of allocation prediction, which in turn increases the risk of successful allocation subversion [11]. Double blinding through the use of placebos reduces the risk of prediction; however, many trials cannot use placebos and for these 'open' trials restricted randomization may significantly increase the risk of allocation prediction [7].

The most common method of restricting randomization is by using block randomization [11]. Block randomization occurs when the allocation sequences are repeated within a fixed block length, giving equal number of each allocation in the block. For example, a block size of four, which is commonly used, with two treatments A and B has six block sequences (i.e., ABAB, ABBA etc). By using block randomization, we can be sure that within each strata the allocation will never be imbalanced by more than half the block size (e.g., two participants for a block size of four).

Concern regarding selection bias in association with blocking has been reported for many years [12–15]. If the block size is known and a record of allocations is kept then for a block size of four the fourth allocation is always 100% predictable and for two of the six possible blocks the last two allocations are always predictable (i.e., AABB, BBAA) [7]. It can also be possible to work out the block size by keeping a log, for instance, Brown et al. [16] found that 16% of surveyed researchers admitted to keeping a log of previous allocations while recruiting participants. Consequently, it is often recommended that small fixed block sizes, such as a block size of four, are not implemented and variable block sizes are used to reduce predictability [17]. In addition, it has been recommended to use simple randomization for larger trials [18–22] and deal with chance imbalances with statistical adjustment at the analysis stage [7, 23–25]. This technique was implemented recently in the RECOVERY trial which adjusted for a chance imbalance in age [26]. Alternatively, trialists should avoid blocking by center as any center stratification increases predictability [11].

A common approach that many researchers used to reduce prediction is to have a mixture of small blocks such as two, four, and six in the randomization process [13,25]. Research has shown that varied block sizes does not completely guard against prediction [27,28], and including very small blocks (e.g., two) in a randomization scheme actually increases the risk of prediction compared with using a single, larger, block size. Whether using small mixed block sizes leads to increased subversion is unknown. Hill and Wheatley have demonstrated in a simulation study that if the block size was not previously known correct prediction can occur 66% of the time with a block size of eight by guessing the opposite as to the previous allocation [23]. There is no empirical evidence that the weakness of small blocks (i.e., their potential predictability) has been exploited in research to select participants into one group or another.

In this paper, we explore within a sample of RCTs from high impact journals whether or not using small block sizes leads to an increase in baseline heterogeneity, which is a marker for potential subversion and whether the use of small block sizes is changing over time.

2. Methods

2.1. Methods for screening and collating all data

Pairs of authors (L.C. and G.R., R.M.C. and L.B.) hand searched the electronic table of contents of the British Medical Journal, The Journals of the American Medical Association, New England Journal of Medicine, and The Lancet to identify individually randomized open RCTs published in January in each year from 2001 to 2020, each pair compared their identified trials to ensure accuracy. Crossover trials were excluded because the perceived advantage of subverting randomization would be largely canceled at the point of crossover; placebo and double-blind trials were excluded because subversion is often considered prohibitively difficult in these designs and cluster trials were excluded because baseline imbalance can occur due to recruitment bias. Interim/preliminary and secondary analyses were also excluded, alongside trials that terminated early. Full-text records were screened with consensus meetings used for trials that could not be categorized with existing decision rules. Data was extracted and second checked by a different author. Uncertainties were resolved by discussion between pairs of authors or by deferring to the wider review team. The following information was extracted from all included RCTs: author, year of publication, publication title, trial design, randomization method, allo-

cation ratio, block size, stratification factor(s), and baseline age for each arm.

2.2. Heterogeneity of age

To assess whether using small block sizes was associated with baseline bias we undertook a series of preplanned meta-analyses of age differences between the randomized groups. We chose age, rather than any other prognostic factor, because the mean and standard deviation (SD) by group is commonly reported in most trials. We did not extract any other prognostic variables as we only intended to use age in our analysis. Furthermore, while other prognostic factors might be more powerful than age, they are likely to be correlated with age. A fixed effect meta-analysis of age was performed in RevMan version 5.4 [29] for each set of RCTs on the assumption that there was a common treatment estimate (zero) across the trials. If the null hypothesis is true (i.e., there is no age difference between randomized groups) then we would expect no heterogeneity except by chance. It is robust randomization (i.e., randomization that has not been subverted by prediction of the allocation sequence) that creates treatment groups that differ only by chance. Poor allocation concealment has previously been shown to be associated with increase heterogeneity of baseline age [5,6,30].

In the case of blocking, if an insecure blocking scheme were implemented that allowed accurate prediction of the allocation sequence, subversion can occur on a prognostic variable (here we tested age) and baseline heterogeneity could be observable. We assessed the heterogeneity of age for all trials implementing blocking and then those trials implementing: a fixed block size of four, block sizes of a fixed block of two and variable schemes where the smallest block is a two, fixed blocks of more than two and a variable scheme that is more than two and where the block size was unknown. This was repeated for those trials implementing stratification by center. Fixed blocks of four were assessed individually as these are accepted to be an insecure blocking scheme with a body of evidence around the risk of prediction [17] and from our previous work is the most common single small block that is widely used.

To compare the difference in heterogeneity between trials using blocking – that is associated with prediction and subversion of randomization – we performed the analyses on trials implementing simple randomization and minimization which are both considered to be at less risk of subversion.

Heterogeneity was interpreted in line with Cochrane recommendations [31]. Trials could be included if their age data were presented in a format that could be converted to a mean and SD. Those trials presenting data in categories or as a mean range were excluded from the age analysis and those presenting age as median and IQR or range, were converted to a mean and SD using standard approximation formulas consistent with Cochrane recommendations [31]. When trials had three arms or more the intervention and first reported control arm were analyzed, in the case of equivalence studies the first two trial arms were analyzed.

3. Results

Four hundred eighty-two trials were screened and 179 open RCTs were identified. Of these nine (5.0%) trials undertook simple randomization and 157 implemented restricted randomization (87.7%). When assessing trials implementing restricted randomization, 104 (66.2%) trials used blocking only with one (0.6%) trial using both minimization and blocking, 27 stated they used stratification with no further details provided (17.2%), 25 (15.9%) used minimization only. One study in 2017 that implemented blocking used a separate block size for males and females, both have been recorded in this review where applicable. Figure 1 illustrates the flow of studies.

The median block size observed throughout the review period was six, this ranged from two to 30, there was no trend in average block size over time. Variable blocking schemes were implemented more frequently than fixed (n = 45 and n = 33, 42.9% and 31.4%, respectively). Table 1 presents the block sizes and stratification details used in RCTs implementing blocking from 2001 to 2020. Overall, 28 trials stratified by center only. Small block sizes have been implemented in recently published trials and in this data set center stratification with a block size of two was only observed to be implemented from 2015. When examining the proportion of trials implementing center stratification, we found 33.3% (n = 5) trials with a blocking scheme including a two implemented it, 6.3% (n = 1) for those using a fixed block of four and 29.0% (n = 9) for trials using a blocking scheme of blocks larger than two. Reporting was suboptimal with 45 (42.9%) of trials not reporting a clear block size and 19 (18.1%) trials not reporting the stratification factor(s).

Figure 2 shows observed statistically significant heterogeneity, I^2 = 24% (P= 0.02), when all eligible trials that implement blocking (n = 94) are analyzed together. One included trial was observed to have a large SD [32] following conversion from a 95% confidence interval to a SD, we undertook the analysis with and without this trial and found the results still hold. In Table 2 we present the results of the amount of heterogeneity associated with age for the trials implementing block randomization with different sized blocks, when the block size in unknown, and for trials implementing simple randomization and minimization. See appendix A for a list of included studies in the meta-analyses. Simple randomization and minimization yielded an expected 0% heterogeneity – which is also seen for trials that implemented mixed blocking schemes where the smallest block size was larger than two.

For trials that use a fixed block of four a substantial statistically significant heterogeneity of 62% (P = 0.001)

Table 1. Block size and center stratification implementation from 2001 to 2020

Year	Number open RCTs	Number of trials which used blocking	Block size					Stratification		Stratification by center			
			Fixed size of two	Variable size including a two	Fixed size of four	Greater than two	Unknown	Stratified only by centre	Stratified factors unclear	Block size including two	Fixed block size of four	Block size >2	Unknown block size
2001	4	3				1	2	1					1
2002	11	5	1		2		2		1				
2003	6	4				2	2	2				2	
2004	7	3				1	2	1					1
2005	9	7		1		2	4	1	2				1
2006	6	5		1		1	3	1	1				1
2007	8	4			2		2	2					2
2008	9	7			1	2	4	2	2			1	1
2009	9	7			1	4	2	2				2	
2010	11	4				1	3	2				1	1
2011	8	6		1	1	1	3		2				
2012	4	2			2								
2013	10	4			1	1	2	1	1				1
2014	12	6				1	5	1	2				1
2015	11	6		3	1	1	1	1	1	1			
2016	8	3		1	2			1	1	1			
2017	10	7		1[l]	1	5[l]	1		1				
2018	12	8		3		4	1	3	2	1		1	1
2019	15	10		1	2	3	4	5	2	1	1	1	2
2020	9	4		1		1	2	2	1			1	1
Total	179	105	1	13	16	31	45	28	19	4	1	9	14

[l] Trial in 2017, two blocking schedules by gender (included here, 2–4 and 4–6).

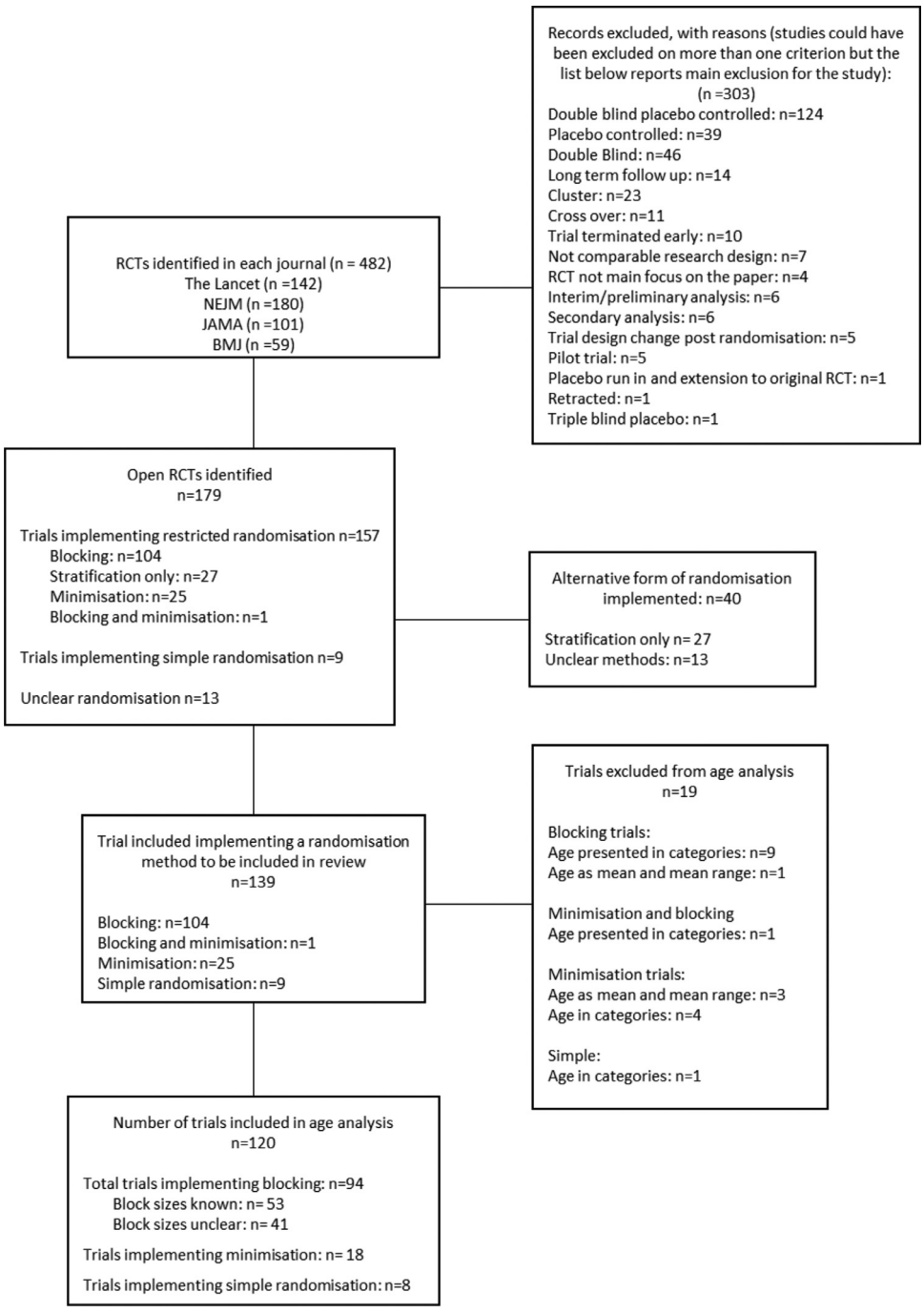


Fig. 1. Flow of studies through study.

was observed. For trials that included a block sizes of two within the randomization schedule moderate heterogeneity was observed (40%), this sample did not include a fixed block of two as there was no age data for the one trial within this dataset that implemented this blocking scheme. When repeating the meta-analyses with only those trials that stratified by center, the heterogeneity increased slightly in trials using a block size of two (now 42%), however, this was not statistically significant. There was zero heterogeneity in mixed blocking schemes with blocks larger than two. There was only one study that used a fixed block size of four that stratified by center; thus no meta-analysis was conducted.

4. Discussion

Methodologists have been warning for some years about the potential dangers of increased allocation prediction when using blocked randomization, especially in conjunction with center stratification [11,20,33]. In this review we

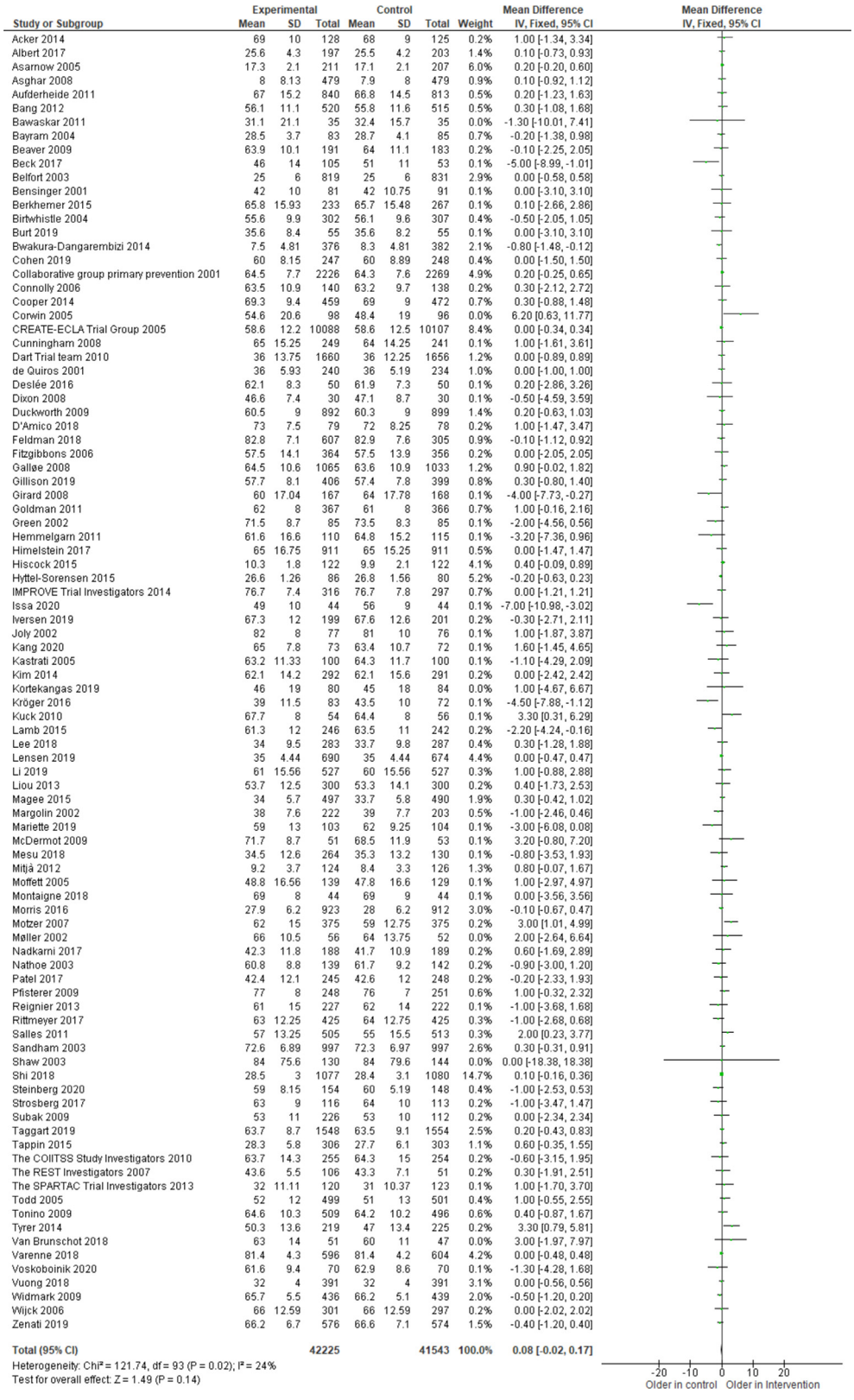


Fig. 2. Meta-analysis of all trials implementing blocking to explore heterogeneity in baseline age.

Table 2. Level of heterogeneity associated with different blocking schemes, with and without center stratification and alternative methods of randomization

	Trials included in MA (n)	X^2	*P* value for heterogeneity	I^2	*P* value of baseline difference
Level of heterogeneity associated with different blocking schemes					
Fixed block of four	14	34.03	0.001	62%	0.29
Variable blocks where the smallest block is a two	13	20.03	0.07	40%	0.41
Fixed blocks of more than two, and variable blocks where the smallest block is greater than a two	27	22.78	0.65	0%	0.95
Block size reported[l]	53	78.16	0.01	33%	0.26
Block size not reported	41	43.83	0.31	9%	0.26
All trials implementing blocking*	94	121.74	0.02	24%	0.14
Level of heterogeneity associated with different blocking schemes stratifying by center					
Fixed block of four	1	-	-	-	-
Variable blocks where the smallest block is a two	4	5.14	0.16	42%	0.78
Fixed blocks of more than two, and variable blocks where the smallest block is greater than a two	7	2.99	0.81	0%	0.68
Block size reported	12	12.03	0.36	9%	0.95
Block size not reported	13	10.30	0.59	0%	0.70
All trials implementing blocking	25	22.44	0.55	0%	0.85
Level of heterogeneity associated with different randomization methods					
Simple randomization	8	6.88	0.44	0%	0.79
Minimization	18	7.33	0.98	0%	0.98

[l] Trial in 2017, two blocking schedules by gender (included here, 2–4 and 4–6), included as a single trial in this meta-analysis.

have shown that using blocking is associated with significant heterogeneity in age and that trials using blocked randomization show an imbalance in age more often than we would expect by chance. A fixed block size of four showed substantial significant heterogeneity. Although it is widely recommended to use variable block sizes to reduce the risk of prediction our review suggests that including a block of size two when using variable block sizes may increase age heterogeneity and should be avoided. We examined whether center stratification led to increased heterogeneity and found that it did not, however, the sample size was small. Simple randomization and minimization showed zero heterogeneity ($I^2 = 0\%$), which is consistent with what is expected with these methods.

Central – or third party – randomization is universally accepted as a secure randomization method and one that should safeguard against subversion [34]; however, if a blocking scheme including a two is implemented then third party safeguards may not be sufficient to ensure secure randomization.

Limitations of this review include missing data, which prevented a full assessment of blocking implementation. When examining the age data some trials were excluded due to the format of reporting age, which could have impacted the observed heterogeneity. Additionally, we were examining heterogeneity of age, whereas selection may have been subverted on a different prognostic variable such as gender: this would lead to underestimation of heterogeneity if baseline variables, other than age, were used to influence treatment allocation. However, performing the analysis on pooled age enabled many trials to be included as it is widely reported, increasing the sample size, and is likely to correlate with the 'true' variable that influenced allocation. Some of the meta-analyses were performed with very few studies which decreases the precision.

It is pertinent to consider that the source of the heterogeneity observed could be due to another type of research misconduct rather than subversion of the randomization schedule. Where participants are selectively excluded at baseline and the baseline heterogeneity analysis is undertaken without these excluded participants who violate intention to treat principles. We find this a less likely explanation than prediction of the randomization sequence and subversion, particularly as the heterogeneity observed in this study was in line with previous research: that fixed blocks of four are at an increased risk of prediction (therefore higher heterogeneity would be expected). Large block sizes – which are harder to predict and less associated with

Table 3. Restricted randomization methods – alternatives to blocking

Randomization type	Description and details	Comments
Minimization	Dynamic form of randomization where participants are allocated according to specific prognostic factors that enables balance to be maintained between groups.	Considered to be more practical and efficient than block randomization. In small trials it has the advantage of producing only a minor difference between groups on variables [41]. Technical issues can lead to imbalance, however, regular checks of this can prevent any problems occurring.
Big stick procedures	Type of 'Maximally tolerated imbalances' (MTI) procedure which uses a 'big stick' to force an allocation sequence back toward balance when it reaches the MTI. Four big stick procedures detailed elsewhere: block urn design [39], Chen's procedure [38], Big stick procedure [37] and the maximal procedure [36].	Can be used incorrectly and result in excessive prediction. The 'big stick' can be invoked when it should not be. Outperforms block randomization at reducing prediction.
Pairwise	Two participants are presented for randomization simultaneously where one is allocated to one arm and the other to the alternative arm. Pairwise randomization is described in detail by Daniels et al. [42].	Not frequently used since first described in 2004. Beneficial to be used when center stratification is required and recruitment is simultaneous [10]. Issues may arise when a suitable pair is not available to randomize.
Merged block randomization	Permuted block randomization with block sequences are merged, to determine the allocation at the point where the sequences merged a coin is tossed and a decision is made on the allocation based on whether it is heads or tails. This novel approach is described in detail in a simulation study by van de Pas [43].	Results in less predictable allocations than block randomization and is a sensible choice for small multicenter clinical trials where the number of participants recruited at each center is anticipated to be small (that can lead to imbalances) [43] Simple and easy to perform in simulations performed.

subversion – and simple randomization and minimization which are considered robust randomization methods have demonstrated the expected low heterogeneity.

There are alternative methods to restrict randomization that are likely to reduce the risk of prediction and should be implemented whenever possible. Dynamic allocation methods, in comparison to blocking provide a more secure method of allocation concealment [35]. We have demonstrated that within this sample, minimization is associated with zero heterogeneity of baseline age. Table 3 summarizes alternative restricted randomization methods to blocking which can be considered but has their own drawbacks. Berger and Odia examined permuted block randomization against maximally tolerated imbalances (MTI) procedures to assess whether it could be classed as a big stick procedure and determined that it was an inferior method to all existing MTI procedures [36–39] and concludes it should not be used within research [40].

Blocking remains the most prevalent way that trials are restricted and will most likely be for some time. If blocking is to be implemented these are our recommendations to ensure it is conducted as methodologically robust as possible:

- The block size should be concealed from all those involved in participant recruitment.
- Fixed block sizes of four should not be implemented.
- A block size of two should not be implemented, even when used within a variable scheme.
- If center stratification is necessary, this research suggests it should be implemented with larger blocking schemes that do not include a block size of two.
- CONSORT (2) needs to be followed to ensure full transparent reporting: the block size and stratification factors need to be reported for a risk of bias assessment to occur when the trial is published.
- Teams using central randomization need to ensure they do not become complacent to the risk of prediction when blocking is used. They need to ensure that small block sizes and stratification by center only is to be avoided if possible and report their methods transparently.

5. Conclusion

There is evidence that blocking is an insecure method of randomization, in addition to fixed blocks of four, variable block sizes including two should probably not be used. Additionally, stratification by center only should, ideally, only be undertaken with larger blocking schemes that do not include a block size of two. Alternative methods are available to restrict randomization which researchers should consider when designing RCTs. Dynamic allocation methods may provide a more secure method to randomize, conceal the allocation, and prevent selection bias.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jclinepi.2021.09.009.

CRediT authorship contribution statement

Laura Clark: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Validation, Supervision, Writing – original draft, Writing – review & editing. **Lauren Burke:** Investigation, Validation, Writing – review & editing. **Rachel Margaret Carr:** Investigation, Validation, Writing – review & editing. **Elizabeth Coleman:** Formal analysis, Data curation, Writing – review & editing. **Gareth Roberts:** Investigation, Validation, Writing – review & editing. **David J. Torgerson:** Conceptualization, Methodology, Writing – review & editing.

References

[1] Sibbald B, Roland M. Understanding controlled trials. Why are randomised controlled trials important? BMJ: Br Med J 1998;316(7126):201.

[2] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010;152(11):726–32.

[3] Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet 2002;359(9306):614–18.

[4] Clark L, Fairhurst C, Cook E, Torgerson DJ. Important outcome predictors showed greater baseline heterogeneity than age in two systematic reviews. J Clin Epidemiol 2015;68(2):175–81.

[5] Clark L, Fairhurst C, Hewitt C, Birks Y, Brabyn S, Cockayne S, et al. Assessing the presence of selection bias in meta–analyses of randomised trials using baseline heterogeneity. Trials 2013;14(1):O96.

[6] Mitchell A, Moe-Byrne T, Cunningham-Burley R, Dean A, Rangan A, Roche J, et al. Poor allocation concealment methods are associated with heterogeneity in age and statistical significance of the primary outcome: review of recent trials published in four general medical journals. J Eval Clin Pract 2019.

[7] Hewitt CE, Torgerson DJ. Is restricted randomisation necessary? BMJ 2006;332(7556):1506–8.

[8] Altman DG, Bland JM. How to randomise. BMJ 1999;319(7211):703–4.

[9] Schultz A, Saville B, Marsh J, Snelling T. An introduction to clinical trial design. Paediatr Respir Rev 2019;32:30–5.

[10] Fairhurst C, Hewitt CE, Torgerson DJ. Using pairwise randomisation to reduce the risk of bias. Res Methods Med Health Sci 2020;1(1):2–6.

[11] Kahan BC, Rehal S, Cro S. Risk of selection bias in randomised trials. Trials 2015;16(1):1–7.

[12] Berger V. Quantifying the magnitude of baseline covariate imbalances resulting from selection bias in randomized clinical trials. Biometr J 2005;47(2):119–27.

[13] Schulz KF. Subverting randomization in controlled trials. JAMA 1995;274(18):1456.

[14] Berger VW, Exner DV. Detecting selection bias in randomized clinical trials* 1. Control Clin Trials 1999;20(4):319–27.

[15] Proschan M. Influence of selection bias on type I error rate under random permuted block designs. Stat Sinica 1994:219–31.

[16] Brown S, Thorpe H, Hawkins K, Brown J. Minimization—reducing predictability for multi-centre trials whilst retaining balance within centre. Stat Med 2005;24(24):3715–27.

[17] Efird J. Blocked randomization with randomly selected block sizes. Int. J Environ Res Public Health 2011;8(1):15–20.

[18] Zhao W. Selection bias, allocation concealment and randomization design in clinical trials. Contemp Clin Trials 2013; 36(1):263–5.

[19] Zhao X. Letter to the Editor: The expression of the allocation concealment for randomised controlled trials needs to be improved. Injury international journal of the care of the injured 2011;42:430–1.

[20] Berger VW. Do not use blocked randomization. Headache: J Head Face Pain 2006;46(2):343.

[21] Schulz KF. Unbiased research and the human spirit: the challenges of randomized controlled trials. CMAJ 1995;153(6):783.

[22] Schulz KF, Grimes DA. Unequal group sizes in randomised trials: guarding against guessing. Lancet North Am Ed 2002;359(9310):966–70.

[23] Hills RK, Gray R, Wheatley K. Balancing treatment allocations by clinician or center in randomized trials allows unacceptable levels of treatment prediction. J Evidence-Based Med 2009;2(3):196–204.

[24] Rosenberger WF, Lachin JM. Randomization in clinical trials: theory and practice. John Wiley & Sons; 2015.

[25] Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. Lancet North Am Ed 2002;359(9305):515–19.

[26] Group RC. Dexamethasone in hospitalized patients with Covid-19—preliminary report. N Engl J Med 2020.

[27] Berger VW. Varying the block size does not conceal the allocation. J Crit Care 2006;2(21):229.

[28] Kennes LN, Cramer E, Hilgers RD, Heussen N. The impact of selection bias on test decisions in randomized clinical trials. Stat Med 2011;30(21):2573–81.

[29] 5 R. Review Manager (RevMan) [Computer program]. Version 5.4.1, The Cochrane Collaboration, 2020. 2020.

[30] Hicks A, Fairhurst C, Torgerson DJ. A simple technique investigating baseline heterogeneity helped to eliminate potential bias in meta-analyses. J Clin Epidemiol 2018;95:55–62.

[31] Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021). Cochrane, 2021. Available at www.training.cochrane.org/handbook.2021., accessed on 26th March 2021

[32] Shaw FE, Bond J, Richardson DA, Dawson P, Steen IN, McKeith IG, et al. Multifactorial intervention after a fall in older people with cognitive impairment and dementia presenting to the accident and emergency department: randomised controlled trial. BMJ 2003;326(7380):73.

[33] Berger V. Selection bias and covariate imbalances in randomized clinical trials. John Wiley & Sons; 2007.

[34] Kennedy AD, Torgerson DJ, Campbell MK, Grant AM. Subversion of allocation concealment in a randomised controlled trial: a historical case study. Trials 2017;18(1):1–6.

[35] Beller EM, Gebski V, Keech AC. Randomisation in clinical trials. Med J Aust 2002;177(10):565–7.

[36] Berger VW, Ivanova A, Deloria, Knoll M. Minimizing predictability while retaining balance through the use of less restrictive randomization procedures. Stat Med 2003;22(19):3017–28.

[37] Soares JF, Jeff Wu C. Some restricted randomization rules in sequential designs. Commun Stat Theory Methods 1983;12(17):2017–34.

[38] Yung-Pin C. Biased coin design with imbalance tolerance. Stochastic Models 1999;15(5):953–75.

[39] Zhao W, Weng Y. Block urn design—a new randomization algorithm for sequential trials with two or more treatments and balanced or unbalanced allocation. Contemp Clin Trials 2011;32(6):953–61.

[40] Berger VW, Odia I. Characterizing permuted block randomization as a big stick procedure. Contemp Clin Trials Commun 2016;2:80–4.

[41] Altman DG, Bland JM. Treatment allocation by minimisation. BMJ 2005;330(7495):843.

[42] Daniels J, Wheatley K, Gray R. Pairwise randomisation to balance within centres without possible foreknowledge of allocation. Controlled Clinical Trials. New York, NY: Elsevier Science Inc; 2003.

[43] van der Pas SL. Merged block randomisation: a novel randomisation procedure for small clinical trials. Clin Trials 2019;16(3):246–52.

Appendix A: included studies in age analysis (web only material)

Acker, M. A., et al. (2014). "Mitral-valve repair versus replacement for severe ischemic mitral regurgitation." N Engl J Med 370: 23-32.

Albert, N., et al. (2017). "Five years of specialised early intervention versus two years of specialised early intervention followed by three years of standard treatment for patients with a first episode psychosis: randomised, superiority, parallel group trial in Denmark (OPUS II)." Bmj 356.

Annane, D., et al. (2010). "Corticosteroid treatment and intensive insulin therapy for septic shock in adults: a randomized controlled trial." Jama 303(4): 341-348.

Asarnow, J. R., et al. (2005). "Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: a randomized controlled trial." Jama 293(3): 311-319.

Asghar, R., et al. (2008). "Chloramphenicol versus ampicillin plus gentamicin for community acquired very severe pneumonia among children aged 2-59 months in low resource settings: multicentre randomised controlled trial (SPEAR study)." Bmj 336(7635): 80-84.

Aufderheide, T. P., et al. (2011). "Standard cardiopulmonary resuscitation versus active compression-decompression cardiopulmonary resuscitation with augmentation of negative intrathoracic pressure for out-of-hospital cardiac arrest: a randomised trial." The Lancet 377(9762): 301-311.

Bakker, J. J., et al. (2010). "Outcomes after internal versus external tocodynamometry for monitoring labor." New England Journal of Medicine 362(4): 306-313.

Bang, Y.-J., et al. (2012). "Adjuvant capecitabine and oxaliplatin for gastric cancer after D2 gastrectomy (CLASSIC): a phase 3 open-label, randomised controlled trial." The Lancet 379(9813): 315-321.

Bawaskar, H. S. and P. H. Bawaskar (2011). "Efficacy and safety of scorpion antivenom plus prazosin compared with prazosin alone for venomous scorpion (Mesobuthus tamulus) sting: randomised open label clinical trial." Bmj 342.

Bayram, N., et al. (2004). "Using an electrocautery strategy or recombinant follicle stimulating hormone to induce ovulation in polycystic ovary syndrome: randomised controlled trial." Bmj 328(7433): 192.

Beaver, K., et al. (2009). "Comparing hospital and telephone follow-up after treatment for breast cancer: randomised equivalence trial." Bmj 338.

Beck, R. W., et al. (2017). "Effect of continuous glucose monitoring on glycemic control in adults with type 1 diabetes using insulin injections: the DIAMOND randomized clinical trial." Jama 317(4): 371-378.

Belfort, M. A., et al. (2003). "A comparison of magnesium sulfate and nimodipine for the prevention of eclampsia." New England Journal of Medicine 348(4): 304-311.

Bensdorp, A., et al. (2015). "Prevention of multiple pregnancies in couples with unexplained or mild male subfertility: randomised controlled trial of in vitro fertilisation with single embryo transfer or in vitro fertilisation in modified natural cycle compared with intrauterine insemination with controlled ovarian hyperstimulation." Bmj 350.

Bensinger, W. I., et al. (2001). "Transplantation of bone marrow as compared with peripheral-blood cells from HLA-identical relatives in patients with hematologic cancers." New England Journal of Medicine 344(3): 175-181.

Berkhemer, O. A., et al. (2015). "A randomized trial of intraarterial treatment for acute ischemic stroke." N Engl J Med 372: 11-20.

Bhatnagar, R., et al. (2020). "Effect of thoracoscopic talc poudrage vs talc slurry via chest tube on pleurodesis failure rate among patients with malignant pleural effusions: a randomized clinical trial." Jama 323(1): 60-69.

Birtwhistle, R. V., et al. (2004). "Randomised equivalence trial comparing three month and six month follow up of patients with hypertension by family practitioners." Bmj 328(7433): 204.

Borghi, L., et al. (2002). "Comparison of two diets for the prevention of recurrent stones in idiopathic hypercalciuria." New England Journal of Medicine 346(2): 77-84.

Brown, C. T., et al. (2007). "Self management for men with lower urinary tract symptoms: randomised controlled trial." Bmj 334(7583): 25.

Burt, R. K., et al. (2019). "Effect of nonmyeloablative hematopoietic stem cell transplantation vs continued disease-modifying therapy on disease progression in patients with relapsing-remitting multiple sclerosis: a randomized clinical trial." Jama 321(2): 165-174.

Bwakura-Dangarembizi, M., et al. (2014). "A randomized trial of prolonged co-trimoxazole in HIV-infected children in Africa." New England Journal of Medicine 370(1): 41-53.

Carlos, M., et al. (2016). "Patient choice in opt-in, active choice, and opt-out HIV screening: randomized clinical trial." Bmj 352.

Christensen, H., et al. (2004). "Delivering interventions for depression by using the internet: randomised controlled trial." Bmj 328(7434): 265.

Cohen, E. E., et al. (2019). "Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study." The Lancet 393(10167): 156-167.

Connolly, S. J., et al. (2006). "Comparison of β-blockers, amiodarone plus β-blockers, or sotalol for prevention of shocks from implantable cardioverter defibrillators: the OPTIC study: a randomized trial." Jama 295(2): 165-171.

Cooper, C. J., et al. (2014). "Stenting and medical therapy for atherosclerotic renal-artery stenosis." New England Journal of Medicine 370(1): 13-22.

Corwin, P., et al. (2005). "Randomised controlled trial of intravenous antibiotic treatment for cellulitis at home compared with hospital." Bmj 330(7483): 129.

Cunningham, D., et al. (2008). "Capecitabine and oxaliplatin for advanced esophagogastric cancer." New England Journal of Medicine 358(1): 36-46.

D'Amico, A. V., et al. (2008). "Androgen suppression and radiation vs radiation alone for prostate cancer: a randomized trial." Jama 299(3): 289-295.

de Quiros, J. C. L. B., et al. (2001). "A randomized trial of the discontinuation of primary and secondary prophylaxis against Pneumocystis carinii pneumonia after highly active antiretroviral therapy in patients with HIV infection." New England Journal of Medicine 344(3): 159-167.

Deslée, G., et al. (2016). "Lung volume reduction coil treatment vs usual care in patients with severe emphysema: the REVOLENS randomized clinical trial." Jama 315(2): 175-184.

Dixon, J. B., et al. (2008). "Adjustable gastric banding and conventional therapy for type 2 diabetes: a randomized controlled trial." Jama 299(3): 316-323.

Duckworth, W., et al. (2009). "Glucose control and vascular complications in veterans with type 2 diabetes." New England Journal of Medicine 360(2): 129-139.

Feldman, T. E., et al. (2018). "Effect of mechanically expanded vs self-expanding transcatheter aortic valve replacement on mortality and major adverse clinical events in high-risk patients with aortic stenosis: the REPRISE III randomized clinical trial." Jama 319(1): 27-37.

Fitzgibbons, R. J., et al. (2006). "Watchful waiting vs repair of inguinal hernia in minimally symptomatic men: a randomized clinical trial." Jama 295(3): 285-292.

Galløe, A. M., et al. (2008). "Comparison of paclitaxel-and sirolimus-eluting stents in everyday clinical practice: the SORT OUT II randomized trial." Jama 299(4): 409-416.

Garry, R., et al. (2004). "The eVALuate study: two parallel randomised trials, one comparing laparoscopic with abdominal hysterectomy, the other comparing laparoscopic with vaginal hysterectomy." Bmj 328(7432): 129.

Geddes, J. R., et al. (2010). "Lithium plus valproate combination therapy versus monotherapy for relapse prevention in bipolar I disorder (BALANCE): a randomised open-label trial." Lancet (London, England) 375(9712): 385-395.

Gillison, M. L., et al. (2019). "Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial." The Lancet 393(10166): 40-50.

Girard, T. D., et al. (2008). "Efficacy and safety of a paired sedation and ventilator weaning protocol for mechanically ventilated patients in intensive care (Awakening and Breathing Controlled trial): a randomised controlled trial." The Lancet 371(9607): 126-134.

Glazener, C. M., et al. (2017). "Mesh, graft, or standard repair for women having primary transvaginal anterior or posterior compartment prolapse surgery: two parallel-group, multicentre, randomised, controlled trials (PROSPECT)." The Lancet 389(10067): 381-392.

Goldman, S., et al. (2011). "Radial artery grafts vs saphenous vein grafts in coronary artery bypass surgery: a randomized trial." Jama 305(2): 167-174.

Green, J., et al. (2002). "Physiotherapy for patients with mobility problems more than 1 year after stroke: a randomised controlled trial." The Lancet 359(9302): 199-203.

Gupta, J., et al. (2013). "Levonorgestrel intrauterine system versus medical therapy for menorrhagia." New England Journal of Medicine 368(2): 128-137.

Hagen, S., et al. (2017). "Pelvic floor muscle training for secondary prevention of pelvic organ prolapse (PREVPROL): a multicentre randomised controlled trial." The Lancet 389(10067): 393-402.

Hemmelgarn, B. R., et al. (2011). "Prevention of dialysis catheter malfunction with recombinant tissue plasminogen activator." New England Journal of Medicine 364(4): 303-312.

Himelstein, A. L., et al. (2017). "Effect of longer-interval vs standard dosing of zoledronic acid on skeletal events in patients with bone metastases: a randomized clinical trial." Jama 317(1): 48-58.

Hiscock, H., et al. (2015). "Impact of a behavioural sleep intervention on symptoms and sleep in children with attention deficit hyperactivity disorder, and parental mental health: randomised controlled trial." Bmj 350.

Hyttel-Sorensen, S., et al. (2015). "Cerebral near infrared spectroscopy oximetry in extremely preterm infants: phase II randomised clinical trial." Bmj 350.

Investigators, I. T. (2014). "Endovascular or open repair strategy for ruptured abdominal aortic aneurysm: 30 day outcomes from IMPROVE randomised trial." Bmj 348.

Investigators, R. (2007). "Uterine-artery embolization versus surgery for symptomatic uterine fibroids." New England Journal of Medicine 356(4): 360-370.

Investigators, S. T. (2013). "Short-course antiretroviral therapy in primary HIV infection." New England Journal of Medicine 368(3): 207-217.

Issa, Y., et al. (2020). "Effect of early surgery vs endoscopy-first approach on pain in patients with chronic pancreatitis: the ESCAPE randomized clinical trial." Jama 323(3): 237-247.

Iversen, K., et al. (2019). "Partial oral versus intravenous antibiotic treatment of endocarditis." New England Journal of Medicine 380(5): 415-424.

Joly, P., et al. (2002). "A comparison of oral and topical corticosteroids in patients with bullous pemphigoid." New England Journal of Medicine 346(5): 321-327.

Kang, D.-H., et al. (2020). "Early surgery or conservative care for asymptomatic aortic stenosis." New England Journal of Medicine 382(2): 111-119.

Kastrati, A., et al. (2005). "Sirolimus-eluting stent or paclitaxel-eluting stent vs balloon angioplasty for prevention of recurrences in patients with coronary in-stent restenosis: a randomized controlled trial." Jama 293(2): 165-171.

Kedhi, E., et al. (2010). "Second-generation everolimus-eluting and paclitaxel-eluting stents in real-life practice (COMPARE): a randomised trial." The Lancet 375(9710): 201-209.

Kim, F., et al. (2014). "Effect of prehospital induction of mild hypothermia on survival and neurological status among adults with cardiac arrest: a randomized clinical trial." Jama 311(1): 45-52.

Kitchener, H., et al. (2008). "Efficacy of systematic pelvic lymphadenectomy in endometrial cancer (MRC ASTEC trial): a randomised study." Lancet (London, England) 373(9658): 125-136.

Kortekangas, T., et al. (2019). "Three week versus six week immobilisation for stable Weber B type ankle fractures: randomised, multicentre, non-inferiority clinical trial." Bmj 364.

Kröger, N., et al. (2016). "Antilymphocyte globulin for prevention of chronic graft-versus-host disease." New England Journal of Medicine 374(1): 43-53.

Kuck, K.-H., et al. (2010). "Catheter ablation of stable ventricular tachycardia before defibrillator implantation in patients with coronary heart disease (VTACH): a multicentre randomised controlled trial." The Lancet 375(9708): 31-40.

Lamb, S. E., et al. (2015). "Exercises to improve function of the rheumatoid hand (SARAH): a randomised controlled trial." The Lancet 385(9966): 421-429.

Lee, J. D., et al. (2018). "Comparative effectiveness of extended-release naltrexone versus buprenorphine-naloxone for opioid relapse prevention (X: BOT): a multicentre, open-label, randomised controlled trial." The Lancet 391(10118): 309-318.

Lensen, S., et al. (2019). "A randomized trial of endometrial scratching before in vitro fertilization." New England Journal of Medicine 380(4): 325-334.

Li, H.-K., et al. (2019). "Oral versus intravenous antibiotics for bone and joint infection." New England Journal of Medicine 380(5): 425-436.

Liou, J.-M., et al. (2013). "Sequential versus triple therapy for the first-line treatment of Helicobacter pylori: a multicentre, open-label, randomised trial." The Lancet 381(9862): 205-213.

Magee, L. A., et al. (2015). "Less-tight versus tight control of hypertension in pregnancy." New England Journal of Medicine 372(5): 407-417.

Margolin, A., et al. (2002). "Acupuncture for the treatment of cocaine addiction: a randomized controlled trial." Jama 287(1): 55-63.

Mariette, C., et al. (2019). "Hybrid minimally invasive esophagectomy for esophageal cancer." New England Journal of Medicine 380(2): 152-162.

McDermott, M. M., et al. (2009). "Treadmill exercise and resistance training in patients with peripheral arterial disease with and without intermittent claudication: a randomized controlled trial." Jama 301(2): 165-174.

Mehanna, H., et al. (2019). "Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial." The Lancet 393(10166): 51-60.

Mehta, S. R., et al. (2005). "Effect of glucose-insulin-potassium infusion on mortality in patients with acute ST-segment elevation myocardial infarction: the CREATE-ECLA randomized controlled trial." Jama 293(4): 437-446.

Mendelow, A. D., et al. (2005). "Early surgery versus initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas in the International Surgical Trial in Intracerebral Haemorrhage (STICH): a randomised trial." The Lancet 365(9457): 387-397.

Mesu, V. K. B. K., et al. (2018). "Oral fexinidazole for late-stage African Trypanosoma brucei gambiense trypanosomiasis: a pivotal multicentre, randomised, non-inferiority trial." The Lancet 391(10116): 144-154.

Mitjà, O., et al. (2012). "Single-dose azithromycin versus benzathine benzylpenicillin for treatment of yaws in children in Papua New Guinea: an open-label, non-inferiority, randomised trial." The Lancet 379(9813): 342-347.

Moffett, J. A. K., et al. (2005). "Randomised trial of a brief physiotherapy intervention compared with usual physiotherapy for neck pain patients: outcomes and patients' preference." Bmj 330(7482): 75.

Møller, A. M., et al. (2002). "Effect of preoperative smoking intervention on postoperative complications: a randomised clinical trial." The Lancet 359(9301): 114-117.

Montaigne, D., et al. (2018). "Daytime variation of perioperative myocardial injury in cardiac surgery and its prevention by Rev-Erbα antagonism: a single-centre propensity-matched cohort study and a randomised study." The Lancet 391(10115): 59-69.

Morris, J. M., et al. (2016). "Immediate delivery compared with expectant management after preterm pre-labour rupture of the membranes close to term (PPROMT trial): a randomised controlled trial." The Lancet 387(10017): 444-452.

Motzer, R. J., et al. (2007). "Sunitinib versus interferon alfa in metastatic renal-cell carcinoma." New England Journal of Medicine 356(2): 115-124.

Nadkarni, A., et al. (2017). "Counselling for Alcohol Problems (CAP), a lay counsellor-delivered brief psychological treatment for harmful drinking in men, in primary care in India: a randomised controlled trial." The Lancet 389(10065): 186-195.

Nathoe, H. M., et al. (2003). "A comparison of on-pump and off-pump coronary bypass surgery in low-risk patients." New England Journal of Medicine 348(5): 394-402.

O'Shaughnessy, J., et al. (2011). "Iniparib plus chemotherapy in metastatic triple-negative breast cancer." New England Journal of Medicine 364(3): 205-214.

Parsons, J. K., et al. (2020). "Effect of a behavioral intervention to increase vegetable consumption on cancer progression among men with early-stage prostate cancer: the MEAL randomized clinical trial." Jama 323(2): 140-148.

Patel, V., et al. (2017). "The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: a randomised controlled trial." The Lancet 389(10065): 176-185.

Pfisterer, M., et al. (2009). "BNP-guided vs symptom-guided heart failure therapy: the Trial of Intensified vs Standard Medical Therapy in Elderly Patients With Congestive Heart Failure (TIME-CHF) randomized trial." Jama 301(4): 383-392.

Primrose, J. N., et al. (2014). "Effect of 3 to 5 years of scheduled CEA and CT follow-up to detect recurrence of colorectal cancer: the FACS randomized clinical trial." Jama 311(3): 263-270.

Reignier, J., et al. (2013). "Effect of not monitoring residual gastric volume on risk of ventilator-associated pneumonia in adults receiving mechanical ventilation and early enteral feeding: a randomized controlled trial." Jama 309(3): 249-256.

Rittmeyer, A., et al. (2017). "Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial." The Lancet 389(10066): 255-265.

Robling, M., et al. (2016). "Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): a pragmatic randomised controlled trial." The Lancet 387(10014): 146-155.

Roncaglioni, M. C. and C. G. o. t. P. P. Project (2001). "Low-dose aspirin and vitamin E in people at cardiovascular risk: a randomised trial in general practice." The Lancet 357(9250): 89-95.

Salisbury, C., et al. (2013). "Effectiveness of PhysioDirect telephone assessment and advice services for patients with musculoskeletal problems: pragmatic randomised controlled trial." Bmj 346.

Salles, G., et al. (2011). "Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial." The Lancet 377(9759): 42-51.

Sandham, J. D., et al. (2003). "A randomized, controlled trial of the use of pulmonary-artery catheters in high-risk surgical patients." New England Journal of Medicine 348(1): 5-14.

Schiffl, H., et al. (2002). "Daily hemodialysis and the outcome of acute renal failure." New England Journal of Medicine 346(5): 305-310.

Shaw, F. E., et al. (2003). "Multifactorial intervention after a fall in older people with cognitive impairment and dementia presenting to the accident and emergency department: randomised controlled trial." Bmj 326(7380): 73.

Shi, Y., et al. (2018). "Transfer of fresh versus frozen embryos in ovulatory women." New England Journal of Medicine 378(2): 126-136.

Slotman, B. J., et al. (2015). "Use of thoracic radiotherapy for extensive stage small-cell lung cancer: a phase 3 randomised controlled trial." The Lancet 385(9962): 36-42.

Steinberg, J. S., et al. (2020). "Effect of renal denervation and catheter ablation vs catheter ablation alone on atrial fibrillation recurrence among patients with paroxysmal atrial fibrillation and hypertension: the ERADICATE-AF randomized clinical trial." Jama 323(3): 248-255.

Strosberg, J., et al. (2017). "Phase 3 trial of 177Lu-Dotatate for midgut neuroendocrine tumors." New England Journal of Medicine 376(2): 125-135.

Subak, L. L., et al. (2009). "Weight loss to treat urinary incontinence in overweight and obese women." New England Journal of Medicine 360(5): 481-490.

Taggart, D. P., et al. (2019). "Bilateral versus single internal-thoracic-artery grafts at 10 years." New England Journal of Medicine 380(5): 437-446.

Tappin, D., et al. (2015). "Financial incentives for smoking cessation in pregnancy: randomised controlled trial." Bmj 350.

Team, D. T. (2010). "Routine versus clinically driven laboratory monitoring of HIV antiretroviral therapy in Africa (DART): a randomised non-inferiority trial." The Lancet 375(9709): 123-131.

Todd, M. M., et al. (2005). "Mild intraoperative hypothermia during surgery for intracranial aneurysm." New England Journal of Medicine 352(2): 135-145.

Tonino, P. A., et al. (2009). "Fractional flow reserve versus angiography for guiding percutaneous coronary intervention." New England Journal of Medicine 360(3): 213-224.

Tyrer, P., et al. (2014). "Clinical and cost-effectiveness of cognitive behaviour therapy for health anxiety in medical patients: a multicentre randomised controlled trial." The Lancet 383(9913): 219-225.

van Brunschot, S., et al. (2018). "Endoscopic or surgical step-up approach for infected necrotising pancreatitis: a multicentre randomised trial." The Lancet 391(10115): 51-58.

Van Driel, W. J., et al. (2018). "Hyperthermic intraperitoneal chemotherapy in ovarian cancer." New England Journal of Medicine 378(3): 230-240.

Van Wijck, A. J., et al. (2006). "The PINE study of epidural steroids and local anaesthetics to prevent postherpetic neuralgia: a randomised controlled trial." The Lancet 367(9506): 219-224.

Varenne, O., et al. (2018). "Drug-eluting stents in elderly patients with coronary artery disease (SENIOR): a randomised single-blind trial." The Lancet 391(10115): 41-50.

Voskoboinik, A., et al. (2020). "Alcohol abstinence in drinkers with atrial fibrillation." New England Journal of Medicine 382(1): 20-28.

Vuong, L. N., et al. (2018). "IVF transfer of fresh or frozen embryos in women without polycystic ovaries." New England Journal of Medicine 378(2): 137-147.

Widmark, A., et al. (2009). "Endocrine treatment, with or without radiotherapy, in locally advanced prostate cancer (SPCG-7/SFUO-3): an open randomised phase III trial." The Lancet 373(9660): 301-308.

Zenati, M. A., et al. (2019). "Randomized trial of endoscopic or open vein-graft harvesting for coronary-artery bypass." New England Journal of Medicine 380(2): 132-141.

Zheng, M.-X., et al. (2018). "Trial of contralateral seventh cervical nerve transfer for spastic arm paralysis." New England Journal of Medicine 378(1): 22-34.

Zipfel, S., et al. (2014). "Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): randomised controlled trial." The Lancet 383(9912): 127-137.