

Delaying and withholding interventions: ethics and the stepped wedge trial

Ariella Binik^{1,2}

¹Department of Philosophy,
McMaster University, Hamilton,
Ontario, Canada

²Institute on Ethics & Policy
for Innovation, McMaster
University, Hamilton, Ontario,
Canada

Correspondence to

Dr Ariella Binik, Department
of Philosophy, McMaster
University, Hamilton, ON
L8S4K1, Canada;
binika@mcmaster.ca

Received 4 September 2018

Revised 12 June 2019

Accepted 1 July 2019

Published Online First

24 July 2019

ABSTRACT

Ethics has been identified as a central reason for choosing the stepped wedge trial over other kinds of trial designs. The potential advantage of the stepped wedge design is that it provides all arms of the trial with the active intervention over the course of the study. Some groups receive it later than others, but the study intervention is not withheld from any group. This feature of the stepped wedge design seems particularly ethically advantageous in two instances: (1) when the study intervention appears especially likely to be effective and (2) when the consequences of not receiving the intervention may be dire. But despite an increase in the use of the stepped wedge design and appeals to its ethical superiority as the motivation for its selection, there has been limited attention to the stepped wedge trial in the ethics literature. In the following, I examine whether there are persuasive ethical reasons to prefer or to require a stepped wedge trial. I argue that while the stepped wedge design is ethically permissible, it is not morally superior to other kinds of trials. To this end, I examine the ethical justification for providing, withholding, and delaying interventions in research.

The stepped wedge trial has been increasing in popularity.^{1–4} It is used in research on HIV, cancers, social policy and criminal justice,⁴ and has been identified as well suited for testing experimental vaccines during emerging epidemics.^{5–7} In the stepped wedge design, clusters crossover from control to intervention sequentially until all clusters are exposed to the intervention.⁴ This means that the intervention is administered to different groups at different times, but the aim is to provide all groups with the active intervention over the course of the trial.

This distinct feature of the stepped wedge trial is thought to be ethically advantageous.^{1,3,8} In fact, ethics is cited in roughly 40% of these studies,⁹ making it the most commonly cited reason for choosing the stepped wedge design.^{9,10} The ethical advantage may be understood as follows: by providing all trial arms with the study intervention at some point during the study, this design promotes equity and avoids ethical challenges associated with withholding treatments, inferior standards of care, and placebo controls. The stepped wedge design seems particularly ethically advantageous in two instances: (1) when an active intervention is thought to be very likely to be effective and (2) when the consequences of not receiving the intervention may be

dire. But despite the increased use of the stepped wedge trial and appeals to its ethical superiority, there has been limited critical attention to this study design.¹

In what follows, I examine whether there are persuasive reasons to think that the stepped wedge trial offers ethical advantages over other trial designs. I argue (1) that delaying an intervention is not a moral solution to the challenge of withholding effective interventions. (2) I then argue that the ethical tension the stepped wedge design aims to address can be resolved without providing the active intervention to all research participants. It can be resolved by appealing to the moral principle of clinical equipoise. And (3) I argue that clinical equipoise is not necessarily disrupted when the consequences of not receiving the study intervention may be dire. Taken together, these arguments suggest that while the stepped wedge design is ethically permissible, and may be supported by socio-political or logistical reasons, it offers no ethical advantage over parallel cluster trials or individually randomised controlled trials.

BACKGROUND

As with other cluster randomised controlled trials, the stepped wedge trial enrolls groups, such as villages or hospitals to take part in a study. But while a parallel cluster trial randomises groups to receive either the active intervention or a control, the stepped wedge design randomises clusters according to the time at which they will receive the intervention. Clusters begin in a control position and then cross over to the active intervention at regular intervals with all clusters receiving the intervention over the course of a trial¹¹ (table 1). The design differs from the cluster crossover trials because in the stepped wedge, the switch to intervention happens uniquely in one direction (from control to intervention).¹¹

Investigators often choose a stepped wedge trial for logistical reasons, such as the infeasibility of introducing an intervention into multiple clusters at the same time.^{3,4} For instance, a stepped wedge may be the best approach for a study in which one team of investigators delivers an educational intervention (such as a campaign to prevent bullying) to participating youth centres one at a time. Another logistical

ⁱ The ethical design, conduct, and implications of cluster randomised trials have recently been analysed,^{40,41} and guidance has been developed,⁴¹ but unique aspects of the stepped wedge design are under-explored.

ⁱⁱ The stepped wedge design may randomise individuals, rather than clusters, but this is rare.¹⁵



© Author(s) (or their employer(s)) 2019. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Binik A. *J Med Ethics* 2019;**45**:662–667.

Table 1 Features of the stepped wedge design

	Individually randomised controlled trial	Cluster randomised controlled trial	Stepped wedge cluster randomised controlled trial
Who is randomised?	Individuals	Clusters (groups)	Clusters (groups)
What is randomised?	Experimental and control interventions	Experimental and control intervention	Timing of experimental intervention

reason to prefer a staged implementation is that it may allow for a more efficient use of resources. For instance, a steady allocation of funding and staff may cost less or be easier to manage than designs implementing the intervention in half of participating clusters at a trial's outset.² Others have cited socio-political motivations for selecting the stepped wedge, including the likelihood that it may facilitate recruitment¹² or prove more appealing to stakeholders.^{2 13}

Perhaps the most prominent reason for choosing the stepped wedge design is that it is ethically advantageous.^{1 3 8 14} There are two situations in which the stepped wedge design may seem particularly ethically beneficial: (1) when there are strong convictions that the intervention being studied is effective. I call this the 'Likely Efficacy' ethical motivation. (2) When the standard of care is ineffective (or there is none) and the consequences of not receiving the study intervention may be dire. I call this the 'Dire Circumstances' ethical motivation. In both situations, the idea is that the stepped wedge design is ethically preferable because all clusters are expected to receive the intervention.

Examining the ethics of the stepped wedge design is significant for several reasons. If this design offers ethical benefits, then it may be worth increasing its use further, or using it in particularly challenging situations. For instance, prominent discussions focus on whether the stepped wedge design should be used to conduct research during public health emergencies,^{7 15–20} and this question merits careful consideration. But if it is not ethically preferable, then the use of the stepped wedge trial should be carefully considered given challenges associated with the design. For instance, the stepped wedge design often requires a longer duration than a parallel design to reach the same statistical power,¹⁴ and may complicate efforts to maintain blinding or to prevent contamination. Further, the design adds logistical complications, may require larger numbers of subjects and measurements, may be vulnerable to drop-out or under-recruitment, and induces confounding by time.¹¹ Overall, the stepped wedge design necessitates additional effort and methodological support,¹⁵ⁱⁱⁱ which suggests that the design should not be selected without good reasons. The central question considered in the following is whether ethics is a good reason to select the stepped wedge trial.

ETHICAL MOTIVATIONS FOR SELECTING THE STEPPED WEDGE DESIGN

Likely efficacy

The stepped wedge design is often selected because of a strong conviction that the intervention being studied is likely to be effective.⁹ That is, when the balance of evidence appears to have

tipped in favour of the study intervention, there seem to be good ethical reasons to provide it to all clusters. Consider the following example:

Communication skills in maternity care in Syria

Research indicates the importance of good communication on women's satisfaction levels with their maternity care. In public hospitals at the community level in Syria, no communication skills training was available for medical residents. To address this, researchers designed a stepped wedge study that implemented a new skills training package to all resident doctors in four teaching hospitals in Damascus and aimed to determine the effect of this training on the satisfaction levels of women with their care.²¹ The study randomized four tertiary care teaching maternity hospitals in Damascus to receive the intervention at two months intervals.

At the outset of this trial, there was no communication skills training available for medical residents, patients expressed dissatisfaction with some elements of their care, and evidence existed linking communication skills training to improved patient satisfaction levels. This suggests that the study intervention is likely to be beneficial, or at least more effective than no intervention. In part for these reasons, the stepped wedge design was identified as a 'key strength of the study'.²¹

Dire consequences

A second instance in which the stepped wedge may seem ethically advantageous is when the consequences of not receiving the active intervention include the potential for significant harm (eg, when the standard of care is ineffective and risks are high). For instance, some have endorsed alternative trial designs as the ethical choice for research during pandemics¹⁸ or when there is 'a very bleak prognosis'.²² Consider the following example of a stepped wedge trial in which participants face significant risks as the result of difficult background circumstances:

Ready to Use Therapeutic Food

This trial aimed to compare a home-based therapy using ready-to-use therapeutic food (RUTF) with standard therapy in treating malnourished children in Malawi. In Malawi, childhood malnutrition is common and the standard therapy has poor recovery rates. The new RUTF therapy had demonstrated success in pilot studies and researchers designed a stepped wedge study to implement the RUTF and to examine whether it offered an improvement over standard therapy for childhood malnutrition.¹³ This trial included children in seven nutritional rehabilitation units (NRUs) located in small towns and rural areas of Southern Malawi. Two NRUs were allocated to intervention at the outset, with an additional NRU beginning participation every three weeks thereafter.¹³

One might argue that given that the potential outcomes of childhood malnutrition are severe, it would be ethically advantageous to provide all participants with the intervention. That is, the stepped wedge design could be understood as addressing a concern about depriving participants of the active intervention when the stakes are high.

THE ETHICAL CHALLENGE AND THE STEPPED WEDGE AS A SOLUTION

The underlying challenge in the 'likely efficacy' and 'dire circumstances' motivation for choosing the stepped wedge design is similar. The concern is that depriving some research participants of the active intervention is unjust. This ethical challenge is perhaps best described as a problem of equipoise. Equipoise is a moral principle that can loosely be defined as uncertainty or

ⁱⁱⁱ See Dousseau and Grady for a comprehensive analysis of the disadvantages associated with the stepped wedge design.¹⁵

disagreement. Equipoise is often recognised as playing a fundamental role in the ethical justification for research.^{23–25} That is, in order for it to be ethically permissible to randomise subjects to any arm of a trial, there must be a state of equipoise. The existence of equipoise helps to ensure that participants in all arms of a trial receive competent medical care.²⁶

The ethical problem that the stepped wedge design aims to resolve is the ethical challenge of research in the absence of equipoise. In instances in which evidence appears to favour the intervention and when the consequences of not receiving the active intervention are likely to be dire, the concern is that equipoise has already been disrupted and it is unfair to deprive any participants of the active intervention. The idea that the stepped wedge design mitigates ethical concerns about research without equipoise appears often. For example, commentators write:

...a stepped wedge design mitigates the ethical dilemma of non-treatment...⁵

A central tenet of parallel or crossover RCTs is that there must be equipoise, that is, a genuine uncertainty of whether one intervention is better than another. Where there is no equipoise, it may be unethical to randomize patients...The stepped wedge design addresses this concern....¹

... a stepped wedge design is considered advantageous when compared to a traditional parallel design. First, if there is a prior belief that the intervention will do more good than harm, rather than a prior belief of equipoise, it may be unethical to withhold the intervention...¹⁴

According to these quotes, the ethical challenge is that the balance of uncertainty has tipped in favour of the intervention and the study is no longer in equipoise. In the absence of equipoise, it would be unethical to withhold the active intervention from any research participants. But the idea is that the stepped wedge design resolves this problem by providing all control groups with the intervention during the study.

In the following sections, I examine the ethical challenge of unfair deprivation and the proposed solution of using a stepped wedge design. I argue for three claims: first, that delaying an intervention that ought to be provided to all research participants is impermissible. The stepped wedge design cannot resolve the problem of research without equipoise. Second, that a prominent interpretation of equipoise—clinical equipoise—helps to clarify why equipoise often does exist when there is some evidence about the efficacy of an intervention. Third, there may be no ethical obligation to provide all research participants with the active intervention even when potential outcomes may be dire. Taken together, these arguments suggest that the stepped wedge is permissible, but not ethically advantageous.

EXAMINING THE PROPOSED SOLUTION

Delaying interventions

The claim that providing all clusters with a likely beneficial intervention in stages is an ethically preferable alternative to withholding interventions is appealing. It draws on the idea that if all groups ultimately receive the active intervention, then no one has experienced unfair deprivation. Understood in this way, the stepped wedge design seems to resolve the ethical challenge of research in the absence of equipoise. But I would like to suggest two reasons this solution is not persuasive.

First, this solution assesses the ethics of randomisation at the wrong point of the trial. The proposed justification suggests that at the end of the trial, one could look back and say that all participants were treated fairly since they received the experimental

intervention. But this reasoning neglects that the ethics of a trial should be established at its outset.²⁷ For a research protocol to be ethically permissible, it must offer participants a reasonable balance between risks and potential benefits,^{28–31} which requires fair treatment for all trial participants. If there is no uncertainty about the relative merits of the trial arms at the outset, then participants should not be randomised to an inferior group. That is, if evidence in favour of the active intervention is strong enough to suggest that all participants should receive it at the beginning of the trial, then there is no clear reason why an ethics committee reviewing the protocol prospectively should accept a justification that depends on control groups crossing over to the active intervention at a later point (perhaps months or years later) in the trial. This reasoning seems to depend on a retroactive consideration of a trial's harms and benefits.

Second, delaying the provision of an intervention that should be provided to all participants does not resolve the ethical challenge of withholding interventions. There is no ethical principle that explains why it might be permissible to deprive any participants of an effective intervention even temporarily.^{3 32} And if it is impermissible to withhold the intervention in the first place, then it is far from clear why depriving it for a shorter period of time would be ethical. That is, a phased implementation would provide the intervention to additional participating clusters at a later point in the trial, but it is not clear that it would resolve any initial unfairness in distributing interventions across trial arms.^{3 32iv}

Are there persuasive arguments that may be constructed or drawn on to justify delayed interventions? I will consider and reject two possibilities. One might defend a delayed intervention by arguing that a temporary delay is a reasonable option if not an ideal one. This argument emphasises the idea that a delayed intervention is preferable to not providing the intervention at all. But the problem with this defence can be seen by considering the ready-to-use therapeutic food trial. In this example, the defence would be that delaying the ready-to-use therapeutic food (RUTF) in some clusters is less than ideal, but preferable to not receiving it at all. This kind of defence would be unpersuasive because it neglects the harm that may occur to participants during the delay. If children were to suffer from complications of malnutrition before their cluster crosses over to the active intervention, then a delayed intervention would not be a preferable alternative.^v The suggestion is not that there is an ethical challenge with this research, but instead that it is difficult to justify delays to interventions that should be provided in the first instance (and that a more successful ethical justification should appeal to other criteria).

Could a modified defense of delaying interventions resolve this challenge? A modified defense might argue that delaying an intervention is ethically permissible because it is a reasonable if not an ideal option provided that the temporary loss of the intervention is unlikely to produce serious or irreversible harm. This modification would retain the idea that receiving an intervention late is an improvement over not offering it at all and ensure that delays are only deemed ethically permissible when they do not permit dire consequences such as a high risk

^{iv} See Rid and Miller,⁴² Eyal and Lipsitch,⁷ and Lipsitch and Eyal¹⁶ for similar arguments about the impermissibility of delaying interventions in the context of public health emergencies.^{7 16 42}

^v See Rid and Miller for an analysis of delaying interventions in a ring vaccination trial design (which also involved delayed vaccination).⁴²

of mortality. Commentators have appealed to similar ideas in the context of debates over the ethical permissibility of withholding interventions in placebo-controlled trials.^{33 34vi} This modified defence of delaying interventions would justify the communication skills trial. Given that the pregnant women are unlikely to suffer irreversible physical harm if their doctors receive the communication skills intervention at a later point in the trial, the delay in this trial may be understood as a defensible option that is unlikely to result in lasting harm.

The modified defence is more persuasive than the first, but also problematic. In the context of the stepped wedge design, it seems to rely on the idea that a delayed intervention is a 'less bad option' and that this is a reasonable justification. But ethical obligations to research participants and efforts to maintain public trust in research should aim at a higher standard. It may not be possible to offer all trial participants the best possible treatment at all times, but researchers' and the state's duties of care to trial participants depend on research participants receiving competent and fair treatment.³⁵ Ensuring that there has been no irreversible harm falls short of the mark. That is, the morally relevant question should not be whether the research participants have suffered serious or irreversible harm, but rather how should we treat all trial participants fairly.^{vii}

Another problem for the modified defence is that it would not permit valuable and justifiable studies, including the childhood malnutrition trial. If the active intervention in this trial turns out to be effective, then delaying it could result in significant harm. It follows that it would not be ethically defensible according to the modified argument for delayed interventions. One might accept this outcome and argue that the stepped wedge design does not render trials involving subjects in high risk circumstances ethically permissible. But I would like to suggest that the trials described above are ethically defensible and highly valuable. Ethical questions associated with withholding interventions are better addressed by considering when it is permissible to randomise to the standard of care rather than by appealing to delayed intervention as an ethical solution.

A DIFFERENT SOLUTION: CLINICAL EQUIPOISE

I've argued that there is no persuasive argument justifying an ethical delay for an intervention that is morally owed to a research participant. But I am not suggesting that the stepped wedge design is unethical. My suggestion is that delaying an active intervention for some research participants may be permissible for a different reason: it is ethical to delay an intervention or to randomise participants to a standard of care control group when

clinical equipoise exists, and clinical equipoise may continue to exist when there is some evidence about the benefit of one intervention over another.

Clinical equipoise, perhaps the most prominent interpretation of equipoise, refers to a state of uncertainty or disagreement in the community of expert practitioners about the relative merit of different interventions.³⁶ It helps to determine the permissibility of randomisation and to ensure that participants in all arms of a trial receive fair and competent treatment. An important feature of clinical equipoise is that it is sufficiently robust to permit randomisation when there exists some evidence in favour of a study intervention. It is disrupted only when this evidence is strong enough to change opinion within the community of expert practitioners. I will elaborate by arguing that clinical equipoise existed at the outset of both the communication skills trial and the RUTF trial. The implications of this will be that the stepped wedge design was permissible but not ethically required since there was no initial challenge of research in the absence of equipoise.

Likely efficacy

A careful look at the evidence reveals that clinical equipoise existed at the outset of the communication skills trial. Evidence in favour of the intervention included existing challenges in doctor–patient communication, a lack of communication skills training for medical residents providing labour and delivery care, and research linking health providers' communication skills to higher patient satisfaction levels.²¹ This evidence suggests that the intervention is likely to be beneficial or at least preferable to no intervention, which would promote the status quo.

However, there was also evidence suggesting that a communication skills training package may not be effective. A number of factors in the study hospitals compromised patient satisfaction levels, including overcrowding, stressful environments and discrimination of the basis of socio-economic status. Further, these hospitals had a policy of not administering pain relief, eye-to-eye contact between a male doctor and a patient in labour was considered unacceptable, and patients were not permitted to be accompanied to the hospital by their relatives.²¹ Given this range of challenges, it was not clear whether a communication skills intervention alone targeted the most significant problem or whether broader structural changes were required to improve patient satisfaction levels.²¹ This uncertainty about the benefit of the intervention suggests that clinical equipoise existed, despite some evidence in favour of the study intervention.^{viii}

CLINICAL EQUIPOISE AND DIRE CIRCUMSTANCES

What about stepped wedge trials in which participants face a high risk of harm? It has been suggested that when conventional care is not particularly beneficial and involves a high rate of mortality, then equipoise is undermined even when an experimental intervention appears only marginally promising.¹⁷ The idea is that the prospect of a successful experimental option—however uncertain—is preferable to conventional care when conventional care is known to have poor success. This idea

^{vi} The commentary on placebos focuses on the permissibility of withholding rather than delaying effective interventions and relies on a broader rejection of the moral principle of equipoise. For prominent examples, see Emanuel and Miller,³³ Temple and Ellenberg,³⁴ and Miller and Brody.⁴³

^{vii} An analogy helps to demonstrate the problem with the modified defence. Consider a patient who visits her doctor for advice concerning the treatment of a disorder that is not life-threatening. The doctor orders some tests, which indicate the need for a particular prescription, but the prescription is delayed for several months as a result of misplaced lab results and poor follow-up. The delay causes no irreversible harm, but slows the patient's recovery, which is accompanied by social and economic challenges arising from prolonged illness. No irreversible damage has occurred and receiving the prescription late is better than not receiving it at all. But even if this patient has not been seriously harmed, she has been wronged. Endorsing the modified defence of delayed interventions is comparable to endorsing a poor (but not dire) outcome for a patient.

^{viii} Ultimately, the study found that while the intervention led to slight changes observed in the resident's communication skills, the training package did not achieve an overall improvement in women's satisfaction levels with the doctor patient relationship.²¹ This led the investigators to conclude that despite evidence that communication skills were lacking from the medical curriculum, broader structural changes in the delivery of care may be required to improve satisfaction scores.²¹

emerged prominently in the context of research during public health emergencies.^{ix} The point I wish to make is that clinical equipoise is not necessarily undermined in trials in which the potential outcomes involve high risks.

Clinical equipoise is an evidentiary standard. It follows that it is disturbed by the state of the evidence concerning the merit of various interventions rather than by the potential severity of the outcomes in any trial arms. The idea that equipoise may continue to exist when conventional treatment involves high risk outcomes has recently been endorsed in policy reports³⁷ and in bioethics commentary.^{38–39} For instance, Alex London argues persuasively that the belief that equipoise has been undermined in high risk circumstances is often based on unwarranted assumptions about the potential efficacy of novel interventions, and that most new interventions turn out to be ineffective.³⁸

How should these insights be taken into account in the stepped wedge trial? At the outset of the RUTF trial, clinical equipoise obtained. That is, there was evidence in favour of both the active intervention and the standard of care. The World Health Guidelines recommended a standard therapy for childhood malnutrition. The standard of care therapy had demonstrated success in some areas, but recovery rates remained poor in Malawi.¹³ This may be due to difficulties adhering to the therapy, which requires caretakers to remain with the child in the nutritional rehabilitation unit and then to prepare porridges over an open fire multiple times a day once they return home.¹³ The alternative home based RUTF therapy reduced the length of inpatient treatment, facilitated home treatment, and had demonstrated success in research and teaching hospitals. But it was not clear whether this success would translate to rural health centres and district hospitals.¹³ More generally, there existed evidence in favour of both the standard of care and the experimental intervention and this evidence was not decisive before the results of the trial become available, which suggests that clinical equipoise existed. It is worth emphasising that assessments about clinical equipoise should be made on the basis of the existing evidence, rather than on the basis of potential outcomes.

This analysis suggests that the stepped wedge design does not offer ethical advantages over other trial designs. But the stepped wedge trial is ethically permissible and may be an appropriate design to select for methodological, practical, or political reasons. For instance, the decision to use a stepped wedge in the RUTF trial included considerations of resource constraints, cultural beliefs, and efforts to control for bias introduced by seasonal variations to the severity and type of childhood malnutrition.¹³ These rationales for selecting the stepped wedge design are plausible.

The broad suggestion from this analysis is that ethical reasons do not clearly motivate the selection of the stepped wedge design. The ethics of randomising interventions in the stepped wedge is best addressed by examining whether clinical equipoise obtains, rather than by appealing to delayed intervention as a more ethical alternative. If there is no clinical equipoise, then delaying the intervention does not overcome the ethical challenge. And if there is clinical equipoise, then providing the active intervention to all clusters offers no ethical advantage.

^{ix} Commentators have argued that equipoise does not obtain in Ebola vaccine trials⁷ or in instances when conventional care offers little benefit and mortality is high.²² Some endorsed the stepped wedge as a solution to ethical challenges arising in research during a public health emergency.¹⁸

CONCLUSION

The stepped wedge trial design is increasing in popularity, but has not received sufficient attention. A distinctive feature of this design—that participating clusters are all scheduled to receive the research intervention at some point during the trial—has given rise to the idea that the stepped wedge design is ethically advantageous. I've argued that these claims are unpersuasive. In particular, I argued that delaying an intervention that should be provided is not justifiable. I then examined the foundation of the ethical challenge motivating the selection of the stepped wedge trial—the claim that they resolve the challenge of research without equipoise. I argued that a careful examination reveals that clinical equipoise may continue to exist when an intervention is perceived as likely to be effective or in high risk circumstances. Taken together, these arguments suggest that the stepped wedge is ethically permissible, but not required, and offers no clear ethical advantages over other trial designs.

Acknowledgements : I am grateful to the members of the Ethics of Pragmatic RCTs Research Group for their helpful comments and support (CIHR Operating Grant PJT-153045). I am also grateful to audiences at the Second International Conference on Stepped Wedge Trial Design and at The Ethox Centre for their valuable questions and suggestions.

Contributors I am the sole author of this submission.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- 1 Mdege ND, Man MS, Taylor Nee Brown CA, *et al.* Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *J Clin Epidemiol* 2011;64:936–48.
- 2 Fatemi Y, Jacobson RM. The stepped wedge cluster randomized trial and its potential for child health services research: a narrative review. *Acad Pediatr* 2015;15:128–33.
- 3 Prost A, Binik A, Abubakar I, *et al.* Logistic, ethical, and political dimensions of stepped wedge trials: critical review and case studies. *Trials* 2015;16:351.
- 4 Hemming K, Haines TP, Chilton PJ, *et al.* The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ* 2015;350:h391.
- 5 Piszczek J, Partlow E. Stepped-wedge trial design to evaluate Ebola treatments. *Lancet Infect Dis* 2015;15:762–3.
- 6 Cohen J, Kupferschmidt K. Tough choices ahead in ebola vaccine trials. *ScienceMag* 2014;1–8.
- 7 Eyal N, Lipsitch M. Vaccine testing for emerging infections: the case for individual randomisation. *J Med Ethics* 2017;43:625–31.
- 8 Beard E, Lewis JJ, Copas A, *et al.* Stepped wedge randomised controlled trials: systematic review of studies published between 2010 and 2014. *Trials* 2015;16:353.
- 9 Grayling MJ, Wason JM, Mander AP. Stepped wedge cluster randomized controlled trial designs: a review of reporting quality and design features. *Trials* 2017;18:33.
- 10 Eichner FA, Groenwold RHH, Grobbee DE, *et al.* Systematic review showed that stepped-wedge cluster randomized trials often did not reach their planned sample size. *J Clin Epidemiol* 2019;107:89–100.
- 11 Campbell MJ, Hemming K, Taljaard M. The stepped wedge cluster randomised trial: what it is and when it should be used. *Med J Aust* 2019;210:253–4.
- 12 Gerritsen DL, Smalbrugge M, Teerenstra S, *et al.* Act In case of Depression: the evaluation of a care program to improve the detection and treatment of depression in nursing homes. Study Protocol. *BMC Psychiatry* 2011;11:91.
- 13 Ciliberto MA, Sandige H, Ndekha MJ, *et al.* Comparison of home-based therapy with ready-to-use therapeutic food with standard therapy in the treatment of malnourished Malawian children: a controlled, clinical effectiveness trial. *Am J Clin Nutr* 2005;81:864–70.
- 14 Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol* 2006;6:54.
- 15 Doussau A, Grady C. Deciphering assumptions about stepped wedge designs: the case of Ebola vaccine research. *J Med Ethics* 2016;42:797–804.
- 16 Lipsitch M, Eyal N. Improving vaccine trials in infectious disease emergencies. *Science* 2017;357:153–6.
- 17 Adebamowo C, Bah-Sow O, Binka F, *et al.* Randomised controlled trials for Ebola: practical and ethical issues. *Lancet* 2014;384:1423–4.
- 18 Edwards SJ. Ethics of clinical science in a public health emergency: drug discovery at the bedside. *Am J Bioeth* 2013;13:3–14.

- 19 Diakite I, Mooring EQ, Velásquez GE, *et al.* Novel Ordered Stepped-Wedge Cluster Trial Designs for Detecting Ebola Vaccine Efficacy Using a Spatially Structured Mathematical Model. *PLoS Negl Trop Dis* 2016;10:e0004866.
- 20 Nuffield Council on Bioethics. Call for Evidence for Research in Global Health Emergencies. <http://nuffieldbioethics.org/project/global-health-emergencies/call-evidence>
- 21 Bashour H, Kanaan M, Kharouf M, *et al.* The effect of training doctors in communication skills on women's satisfaction with doctor-woman relationship during labour and delivery: a stepped wedge cluster randomised trial in Damascus. *British Medical Journal Open* 2013;3:e002674.
- 22 Caplan AL, Plunkett C, Levin B. Selecting the right tool for the job. *Am J Bioeth* 2015;15:4–10.
- 23 Hey SP, London AJ, Weijer C, *et al.* Is the concept of clinical equipoise still relevant to research? *BMJ* 2017;359:j5787.
- 24 London AJ, Value S. Clinical Equipoise, and Research in a Public Health Emergency. *Bioethics* 2018.
- 25 Kukla R. Equipoise, uncertainty, and inductive risk in research involving pregnant women. In: Baylis F, Ballantyne A, eds. *Clinical research involving pregnant women*. Switzerland: Springer, 2016:179–96.
- 26 Weijer C, Miller PB, Graham M. The Duty of Care and Equipoise in Randomized Controlled Trials. In: Arras JD, Fenton E, Kukla R, eds. *The Routledge Companion to Bioethics*. New York: Routledge, 2014.
- 27 Beecher HK. Scarce resources and medical advancement. *Daedalus* 1969;98:275–313.
- 28 The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. 1979. DHEW Publication No. (OS) 78-0012. Washington DC.
- 29 World Medical Association. Declaration of Helsinki. Ethical Principles for Medical research Involving Human Subjects. 2013. <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/> (accessed Aug 2 2018).
- 30 CIOMS, Council for International Organizations of Medical Sciences in Collaboration with the World Health Organization. *International ethical guidelines for health-related research involving humans*. Geneva, 2016.
- 31 Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA* 2000;283:2701–11.
- 32 Binik A, Weijer C, McRae AD, *et al.* Does clinical equipoise apply to cluster randomized trials in health research? *Trials* 2011;12:118.
- 33 Emanuel EJ, Miller FG. The ethics of placebo-controlled trials--a middle ground. *N Engl J Med* 2001;345:915–9.
- 34 Temple R, Ellenberg SS. Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 1: ethical and scientific issues. *Ann Intern Med* 2000;133:455–63.
- 35 Miller PB, Weijer C. Trust based obligations of the state and physician-researchers to patient-subjects. *J Med Ethics* 2006;32:542–7.
- 36 Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med* 1987;317:141–5.
- 37 London AJ, Omotade OO, Mello MM, *et al.* Ethics of randomized trials in a public health emergency. *PLoS Negl Trop Dis* 2018;12:e0006313.
- 38 London AJ. Social value, clinical equipoise, and research in a public health emergency. *Bioethics* 2019;33:326–34.
- 39 Mullard A. Parsing clinical success rates. *Nat Rev Drug Discov* 2016;15:447.
- 40 Macklin R. Cluster randomized trials: another look. *Hastings Cent Rep* 2014;44:37–43.
- 41 Weijer C, Grimshaw JM, Taljaard M, *et al.* Ethical issues posed by cluster randomized trials in health research. *Trials* 2011;12:100.
- 42 Rid A, Miller FG. Ethical Rationale for the Ebola "Ring Vaccination" Trial Design. *Am J Public Health* 2016;106:432–5.
- 43 Miller FG, Brody H. A critique of clinical equipoise. Therapeutic misconception in the ethics of clinical trials. *Hastings Cent Rep* 2003;33:19–28.