

Basic principles in designing studies to assess the effects of interventions

Authors

Mike Clarke, Centre for Public Health, Queen's University, Belfast, United Kingdom; Evidence Aid, London, United Kingdom.

Dimuthu Rathnayake, Ministry of Health, Colombo, Sri Lanka.

4.1.1 Learning objectives

To understand key factors to consider when developing a study to assess the effects of an intervention, action or strategy for health emergency and disaster risk management (Health EDRM), including:

1. The importance of reliable and robust estimates of the effects of interventions.
2. Minimizing the risk of bias.
3. The role of randomized trials.
4. Aspects of conducting prospective, comparative studies.

4.1.2 Introduction

This chapter will show how research can provide reliable and robust evidence about the likely effects of different interventions in order to help people choose between alternatives when there is more than one intervention suitable for an individual, or a variety of actions or strategies that are appropriate for a population. To be reliable, this evidence needs to come from studies in which the interventions were compared in ways that minimize the effects of biases (1), such as biases that might arise from using information about a participant's likely outcomes to select who will or will not receive one of the interventions being compared. To be robust, the studies also need to be large enough to minimize the effects of chance.

This chapter outlines how such studies might be carried out in the Health EDRM context and highlights important features for the design, conduct and interpretation of such studies. The various types of research design that might be used to study different areas of importance to Health EDRM are discussed in Chapter 3.5. In this chapter, particular emphasis is placed on a type of comparative effectiveness study called a randomized trial, because this design seeks to minimize bias and generate reliable and robust estimates of the relative effects of interventions. It does this by creating comparison groups that differ only in regard to the interventions being compared. In randomized trials, some of the individuals who join the study are randomly allocated to receive the intervention being tested,



4.1

which might be a new way of protecting people from contaminated water, a treatment for fractures or a treatment to minimize anxiety, for example; this is often called the experimental group. The other participants in the trial would be allocated to an alternative intervention or a control group. Cluster randomized trials are a modified version of this design, when randomization is done at the level of clusters (such as families, villages or hospital wards), rather than at the individual level. These are discussed in Chapter 4.3.

Randomized trials seek to answer research questions about cause and effect in a controlled manner. Their aim is to produce an estimate of the impact or effect of the intervention by comparing the outcomes in the experimental group to those in the control group. The purpose of this is to generate evidence, which can then be used to make assumptions about how the intervention might affect people who are similar to those in the trial.

However, although we focus here on randomized trials, many of their key features discussed below are also applicable to other prospective studies in which individuals are recruited and followed up.

4.1.3 Why do we use randomized trials?

Randomized trials are prospective studies in which eligible participants are randomly allocated to one of the two or more groups that are to be compared, with each group receiving a different intervention. This allows a comparison to be made of how each intervention affects the outcomes that are measured — such as the speed of a person's recovery, their quality of life, or how well they understand information about a disaster-related threat to their health or livelihood. However, for some research questions – on topics such as estimating the proportion of people who have different levels of mental or physical trauma after an earthquake, for example – other study designs would be used; these are discussed elsewhere in this book, such as in Chapter 3.2 for assessing risk factors.

4.1.4 Planning the trial: eligibility criteria

Chapter 3.5 discussed the importance of having a clear research question for a study, including the need to match the research question to the comparison to be made in a randomized trial, using the example of fish oil for PTSD. Case Study 4.1.1 describes a randomized trial undertaken with rescue workers after the Great East Japan Earthquake in 2011. It illustrates both the comparison that was made and the decision about the population to study. The decisions about who to study are set out in the inclusion and exclusion criteria for a trial, which may be broad or narrow, and determine who is and is not eligible for the study (2).

Case Study 4.1.1**The APOP randomized trial of fish oil for attenuating post-traumatic stress disorder (PTSD) symptoms among rescue workers after the Great East Japan earthquake**

The Great East Japan Earthquake and tsunami of 11 March 2011 caused tremendous damage to the north-eastern coast of Japan, leaving 20 000 people dead or missing. Many rescue workers were exposed to traumatic experiences. Researchers decided to investigate whether PTSD symptoms might be attenuated by the use of fish oil. The same researchers had previously shown that PTSD symptoms at 12 weeks after injury were significantly alleviated if patients with physical injury took fish oil. The new study was done among disaster medical assistance team (DMAT) members who were deployed during the acute disaster phase of the earthquake. The randomized trial was approved on 1 April 2011 and started the following day.

After providing informed consent, participants were randomly assigned to one of two groups - one group that received the fish oil supplementation plus psychoeducation, or the other group, which received psychoeducation alone (3). The fish oil was given as seven capsules per day, each containing 320mg of fish oil. 172 rescue workers joined the trial between 2 and 12 April 2011 and were followed up over the next few months.

The primary outcome was measured using the Impact of Event Scale-Revised (IES-R), and this showed no significant difference at 12 weeks between the decline in scores for participants in the fish oil group compared to those in the control group (4).

In an explanatory trial – also known as an efficacy trial – the inclusion criteria might be kept narrow to ensure that the people recruited to the study are all similar to one another. Such a trial would determine whether, in such ideal circumstances, there is a difference between the interventions being compared. Examples of such studies include: randomized trials to compare the speed of onset of pain control when two formulations of an analgesic drug are used in people with specific types of minor injury; a comparison of surgical techniques for managing fractures of the lower leg; or a test of a psychological therapy in school-aged children following a tsunami. In studies of this kind, the participants would be carefully chosen so that they have the characteristics that are felt to be most receptive to the intervention being tested. One rationale for such trials is that, if the experimental intervention is no better than the routine intervention in these “ideal” circumstances, it is unlikely to be better in a much broader population.

However, in health emergencies and when seeking to manage disaster risk, randomized trials are more likely to take the form of an effectiveness or pragmatic trial. This is because a wide range of participants is likely to be recruited, and there would likely be less strict control over the specific elements of the interventions being tested, in order to make the trial as close as possible to routine practice. In effectiveness studies, the eligibility criteria are broad enough to ensure that many of the types of people who are likely to be considered for the intervention in the future are included.



4.1

Such trials might use the “uncertainty principle” to set the eligibility criteria (5), meaning that people would be eligible for a trial if there is sufficient uncertainty about what effects the interventions would have for them. This is also a fair way to allocate interventions when a choice has to be made about who is given or not given the intervention, as is often the case in Health EDRM. When deciding on the eligibility criteria for a trial, and its feasibility, careful consideration is also required of what number of participants will be needed to answer the research question: researchers deal with this when calculating the necessary sample size, which is discussed in Chapter 4.2.

4.1.5 Participant selection and informed consent

The success of any prospective study relies on the cooperation of the people who are participating in it. In medicine, one major difference between treating patients inside or outside a research study is the formal process of informed consent that is likely to be required for the study (see Chapter 6.4); this can be challenging in disaster situations where the intervention has to be administered quickly, there is little time to provide detailed information or no opportunity for a full discussion with potential participants. However, there are several examples of ethically acceptable trials conducted in such difficult circumstances. For instance, the CRASH trial recruited patients with serious head injuries and showed that a widely used treatment, steroids, was not beneficial (6).

The uncertainty principle can also be employed in deciding whether or not a trial is ethical (see Chapters 3.4 and 6.4 for a discussion of the ethics of research). For example, it can be used when considering whether it is ethical to not do a randomized trial. If there is uncertainty about the relative effects of two interventions, and both are available and suitable for the target population, the most ethical approach may be for them to enter a randomized trial. This ensures that participants have a fair chance of receiving the more beneficial intervention (since it will be unknown when they join the trial which this will be) and the data collected should help to resolve uncertainty in the future, as was the case with the aforementioned CRASH trial for people with head injuries (6).

4.1.6 Randomizing participants

The key feature distinguishing randomized trials from other prospective studies is the use of a random process to determine which of the interventions is received by each participant. This process ensures that any differences between the outcomes for those in the randomized groups will be due either to the effects of the interventions being compared, or to the effects of chance.

Randomization can be achieved in a variety of ways, and some methods are described here. The key elements are the use of a random sequence to allocate participants to one of the groups, and ensuring that no-one knows which group a person will be allocated to before they join the trial. If an individual's allocated group is known in advance, this may lead to a different decision being made about whether they join the trial, or to some other form of manipulation, such as delaying their joining until a different allocation is available.

Generating a random sequence

In simple randomization, each participant has the same probability of being allocated to each intervention being tested. This can be achieved using simple physical techniques such as flipping a coin, rolling a dice or drawing lots. It might also be done by shuffling envelopes into which information about the allocation has been placed. Mathematical techniques, using random numbers, can also be used. Simple randomization is completely unpredictable, provided that the allocation for an individual participant is concealed up until the point that they enter the trial. However, the disadvantage of simple randomization is that, particularly in a small trial, it can lead to large, chance imbalances between the groups. For example, if a coin is flipped 100 times, it is likely that at some point in the sequence there will be a consecutive run of 6, 7 or 8 heads or tails. If this occurred in a trial, it could lead to an imbalance in the number of people in the groups, making analysis of the trial difficult. It could also lead to imbalances in participant characteristics between the groups, which might also make the analysis of the trial more difficult.

These potential problems can be overcome by using a technique called blocked randomization, which allows stratification of the allocated interventions (or a more complex, computer-based technique called minimization (7)). Blocked randomization means that after a particular number of participants have been allocated, the numbers in the different intervention groups will be balanced. For example, a block size of four in a trial with two intervention groups guarantees that for each sequence of four people joining the trial, two will be allocated to one group and two to the other group; using that block size for a trial as a whole will therefore ensure that the difference between the number of people in each of the two groups will be no more than two (if, at the start of the final block, two are both allocated to the same group). Similarly, using blocks for different types of people in the trial (for example, young and old, or those living in rural, semi-urban and urban settings) can ensure balance within those groups.

Concealing the random sequence until the participant joins the trial

Allocation concealment is not the same as blinding or masking the intervention, which is discussed below and happens after the person has entered the trial. Allocation concealment takes place earlier, before the person enters the trial. It means that no-one involved in recruiting potential participants can know what they will receive until they have joined the trial. Allocation concealment prevents manipulation that might arise if knowing the allocation leads to a different decision about someone's eligibility or their willingness to join the trial.

One way to implement adequate allocation concealment is to use sealed, opaque, sequentially numbered envelopes, which must be used in the predetermined sequence and cannot be opened to reveal the allocation until the person has entered the trial. Researchers might also use randomization systems in which an online or computer-based system, or a telephone call, is used to first capture data on the participant before their allocation is given.

4.1**4.1.7 Blinding or masking**

In some studies, it is important that the people involved in conducting the trial do not know which intervention a participant is receiving. This is usually called 'blinding' or, particularly when the research is related to eyesight, 'masking', and might be achieved by giving patients in the control group a dummy intervention or placebo. However, adding placebos or blinding to trials can be difficult, because doing so increases the resources needed for the trial and can make the interpretation of the results more difficult because after the trial, in routine practice, those receiving or administering an intervention would know what is being taken or given (8).

There are a number of different people involved in a trial who might be kept blind to the intervention and there are a variety of reasons for doing so. Typically, the participant might be kept blind in order to reduce the risk that they will either report outcomes differently because they know which intervention they are receiving or, through a placebo effect, will actually respond differently simply because of their knowledge of the intervention rather than as a result of the intervention itself. Problems can also arise if participants knowing which intervention group they are in makes them change their behaviour in ways that would not happen outside of the trial.

To illustrate the potential impact of blinding: in a randomized trial of an iron-fortified biscuit for children with iron deficiencies, those who know they are in the control group might try to change their eating habits, while those in the intervention group might change in a different way, perhaps assuming that the biscuits will provide the nutrition that they need. Blinding might be achieved by giving those in the control group a biscuit that is identical in every way except for the ingredient being tested, to act as a placebo.

It might also be important to keep people other than the participant blind to the allocated intervention. This can include those treating and caring for patients in a study and the people measuring outcomes. Keeping the practitioners blind ensures that they are less likely to do other things differently for a patient – just as the participant might modify their behaviour if they know which intervention they have been allocated, practitioners might add extra treatments if they know a patient is in the control group or monitor them more carefully if they are receiving the experimental intervention.

Likewise, if the people assessing the participants' outcomes or collecting data know that someone is receiving the experimental intervention, they might look more closely for side effects. If someone is in the control group, unblinded assessors might be more pessimistic when recording their outcomes. For example, in a trial testing different types of dressing for wounds after surgery, it could be important that the outcome assessor responsible for classifying the level of infection in a wound did not know which dressing was used when they made their assessment. Problems can also arise if the statistician analysing the trial's results is influenced in how they do this by knowing which group is the experimental group. In such circumstances, it would be important to keep them blind to which group is which.

4.1.8 Avoiding publication bias: registering and reporting a study

Even if a researcher is careful to minimize bias when designing and conducting their prospective study, biases can be introduced when they make decisions about reporting its findings. These can lead to problems when the results are used by others. Publication bias arises when the results of a study have an influence over whether it is published. Selective reporting bias can mean that, even though the study is published, some of its findings remain unpublished, while others are given more prominence. Chapter 6.6 describes some of the elements to consider when reporting a research study, and the importance of publishing research in ways that will help people and organizations such as United Nations agencies, NGOs and others involved in Health EDRM to use the findings in their future decision making.

During recent decades, efforts to combat the problems of publication and selective reporting bias have led to the development of prospective registers of research studies (9). Registering the study before the first participant is recruited makes the existence of the study public knowledge in a way that ensures that this could not possibly be influenced by its results. It also requires the researcher to say, in advance, what they are studying. Some journals will not publish the results of trials that have not been prospectively registered. Furthermore, in the context of a sudden-onset disaster, carefully pre-planning the trial, registering and perhaps even publishing its full design in advance, allows a trial to be sitting “on the shelf” ready to be activated. Case Study 4.1.2 presents one such example, where a detailed plan has been prepared for a blinded, randomized trial of regional anaesthesia in earthquake survivors with lower limb trauma.

4.1

Case Study 4.1.2**Plan for a randomized trial of anaesthesia and pain management for patients with lower limb trauma after an earthquake**

After an earthquake, the largest burden of injuries is due to trauma of the legs and feet, and pain management for these patients is a substantial challenge. The Regional Anaesthesia for Painful Injuries after Disasters (RAPID) trial has been designed to evaluate whether regional anaesthesia, either with or without ultrasound guidance, can reduce pain from earthquake-related lower limb injuries in a disaster setting (10). The plan for the trial was prospectively registered in February 2016.

After informed consent has been obtained, study participants will be randomized in a 1:1:1 allocation to standard care (parenteral morphine at 0.1 mg/kg), standard care plus a landmark-guided fascia iliaca compartment block, or standard care plus an ultrasound-guided femoral nerve block. In order to blind participants and healthcare providers who are not part of the research to a patient's allocated group, sham ultrasound activities will be used in the first two groups and a normal saline injection will be given to the first group (the control group). The primary outcome measure will be a standard pain intensity score over the first 24 hours, with secondary outcome measures including analgesic requirements, adverse events, and participant satisfaction.

If the trial shows that regional anaesthesia is effective in a disaster setting, its future use for survivors of earthquake trauma could reduce both their acute suffering and the long-term complications of the injury.

4.1.9 Other types of prospective, comparative study

When it is not feasible to use randomization to allocate individuals to different interventions, there are other methods that can be used. For example, for a research question relating to a comparison of different methods of coordinating the multidimensional response to a disaster, randomly assigning individuals or groups of people to coordinate their actions in very different ways would be likely to lead to chaos. Instead, the new method of coordination could be implemented and then its impact assessed using a "counterfactual" to estimate what might have happened without the intervention in order to decide whether it improved, worsened or made no difference to outcomes. This might also be the case for other interventions; methods for conducting such studies are discussed in Chapter 4.15. To illustrate the planning of such a study, Case Study 4.1.3 describes how the findings from research into a surge of dengue cases at a hospital in Sri Lanka might be used in the evaluation of future changes to hospital strategy and health systems research.

Case Study 4.1.3**Planning an evaluation of strategies that would be implemented in a future health emergency**

Dengue is the most important infectious disease-related public health concern in Sri Lanka. A massive outbreak occurred at the time of the south-western monsoon rains in 2017: approximately 185 000 dengue cases were reported and more than 400 people died (11). The National Institute of Infectious Diseases, as the leading hospital for managing infectious diseases in Sri Lanka, played a major role during the outbreak and researchers there studied the size and effects of the dengue epidemic (12). Their study identified particular challenges and, along with a systematic review (13), has led to proposals for implementation in the future. These include the need for public health systems to use robust systems approaches with sufficiently detailed managerial approaches. It would not be possible to assess the effects of these systems-level strategies in a randomized trial because it would not be feasible to allocate them to some individuals or hospitals, and not to others. However, it would still be useful to know how effective they are. In order to assess this, a prospective study would be put in place to gather outcome measures that could then be compared with the earlier data. This would seek to answer a research question about whether the new systems were an improvement on the old systems, and provide evidence to inform the decision to continue with them or refine them further for future dengue epidemics. However, caution would be needed when deciding whether the comparison of the future epidemic with that in 2017 was a valid comparison of “like with like” in relation to everything except the new strategies. The prospective study would collect information on the dengue cases, the use of hospital resources and outcomes for patients. It would include attendance at the outpatient department, admissions to hospital and bed occupancy before and during the next outbreak, and demand on services such as the haematology laboratories. These data would then be compared with the findings from 2017, with care being taken to ensure that any differences were not merely due to differences in the way in which the data were gathered.

Two other types of prospective study that might be used when randomized trials are not feasible are described below.

Controlled before-after study

In a controlled before-after study, the decision about whether a person will be in the intervention or the control group is not made by the researcher. The outcomes of the people in both groups are measured before and after the intervention is introduced for one of the groups. For example, if some people who lost their homes after a windstorm are given a new type of shelter, their respiratory health would be monitored before and after the delivery of the new shelters, as well as that of a control group of people provided with the usual shelter. One disadvantage of these studies is that they have a high risk of bias because there may be differences between the intervention and control groups. If these differences not only determined whether a person went into the intervention or the control group but also had an effect on their outcomes, it is possible that the



4.1

study's findings might simply arise from these underlying differences rather than from the effects of the intervention.

Interrupted time series

In an interrupted time series design, outcomes are collected at multiple time points, before and after the intervention is introduced. A single setting or group of participants is used, and there is no control group. The effect of the intervention would then be estimated by comparing the trend in the outcomes after its implementation with the trend beforehand. For example, if the level of gender-based violence was holding steady or slowly declining in a displaced person's camp, but declined rapidly after a new strategy was put in place, this would suggest that the new strategy is beneficial.

However, a disadvantage of this design is that if any other features of the camp had changed close to the time that the intervention was introduced, it would not be known whether those changes may have caused (in full or in part) any detected improvement (or conversely, if the new intervention did not appear to have an impact, may have cancelled out what would have been a benefit).

4.1.10 Conclusions

For many centuries, decisions about interventions and policies intended to improve the health of populations were based mostly on personal experience, anecdotal case histories and comparisons of people who had received one intervention with an entirely separate group who had not received it or had received something different. Although these sources of knowledge are still in use today, they are subject to biases which mean that the information they provide may be unreliable.

In recent decades, routine health care and policy making has relied increasingly on randomized trials and systematic reviews (see Chapter 2.6) of these as a source of reliable and robust estimates of the relative effects of different interventions. Provided the trial is sufficiently large, random allocation ensures that any differences in outcomes between groups must be due to the effects of the interventions. This allows future decision makers to have greater confidence in the answer provided by the trial when they are choosing interventions or setting policy.

4.1.11 Key messages

- o **People choosing between different interventions, actions and strategies need reliable and robust evidence on their relative effects.**
- o **Such evidence needs to come from research that has minimized the effects of bias and chance.**
- o **Randomized trials provide a means for testing interventions in such a way that any difference between the outcomes of the participants in the groups being compared are due to the effects of the intervention, or chance.**
- o **Pre-planning a trial, or other prospective study, allows it to be ready to be activated when needed, for example in a sudden-onset disaster.**

4.1.12 Further reading

Clarke M, Allen C, Archer F, Wong D, Eriksson A, Puri J. What evidence is available and what is required, in humanitarian assistance? Scoping Paper 1. New Delhi: International Initiative for Impact Evaluation (3ie). December 2014. https://www.3ieimpact.org/sites/default/files/2019-01/3ie_scoping_paper_1-humanitarian-top.pdf (accessed 9 February 2020).

White H. An introduction to the use of randomized control trials to evaluate development interventions. *Journal of Development Effectiveness*. 2013; 5(1): 30-49.

4.1.13 References

1. Clarke M, Atkinson P, Badenoch D, Chalmers I, Glasziou P, Podolsky S et al. The James Lind Library's Introduction to Fair Tests of Treatments. Oxford, UK: James Lind Library. 2018 <http://www.jameslindlibrary.org/download/18790> (accessed 9 February 2020).
2. Yusuf S, Held P, Teo KK, Toretzky ER. Selection of patients for randomized controlled trials: implications of wide or narrow eligibility criteria. *Statistics in Medicine*. 1990; 9: 73-86.
3. Matsuoka Y, Nishi D, Nakaya N, Sone T, Hamazaki K, Hamazaki T et al. Attenuating posttraumatic distress with omega-3 polyunsaturated fatty acids among disaster medical assistance team members after the Great East Japan Earthquake: the APOP randomized controlled trial. *BMC Psychiatry*. 2011; 11: 132.
4. Nishi D, Koido Y, Nakaya N, Sone T, Noguchi H, Hamazaki K, et al. Fish oil for attenuating posttraumatic stress symptoms among rescue workers after the Great East Japan earthquake: a randomized controlled trial. *Psychotherapy and Psychosomatics*. 2012; 81: 315-7.
5. Peto R, Baigent C. Trials: the next 50 years. Large scale randomized evidence of moderate benefits. *BMJ*. 1998; 317: 1170-1.
6. Edwards P, Arango M, Balica L, Cottingham R, El-Sayed H, Farrell B, et al. Final results of MRC CRASH, a randomized placebo-controlled trial of intravenous corticosteroid in adults with head injury - outcomes at 6 months. *Lancet*. 2005; 365: 1957-9.
7. Treasure T, MacRae KD. Minimization: the platinum standard for trials. *BMJ*. 1998; 317: 362-3.
8. Anand R, Norrie J, Bradley JM, McAuley DF, Clarke M. Fool's gold? Why blinded trials are not always best. *BMJ*. 2020; 368: l6228.
9. Ranawaka UK, de Abrew A, Wimalachandra M, Samaranayake N, Goonaratna C. Ten years of clinical trial registration in a resource-limited setting: Experience of the Sri Lanka clinical trials registry. *Journal of Evidence Based Medicine*. 2018; 11(1): 46-50.
10. Levine AC, Teicher C, Aluisio AR, Wiskel T, Valles P, Trelles M, et al. Regional Anesthesia for Painful Injuries after Disasters (RAPID): study protocol for a randomized controlled trial. *Trials*. 2016; 17(1): 542.
11. Tissera HA, Jayamanne BDW, Raut R, Janaki SMD, Tozan Y, Samaraweera PC, et al. Severe Dengue Epidemic, Sri Lanka, 2017. *Emerging Infectious Diseases* 2020; 26(4): 682-91.
12. Rathnayake D, Wijewickrama A, Aluthge K. Response of the National Institute of Infectious Diseases, Sri Lanka to an unexpected dengue epidemic in 2017. *Ceylon Medical Journal*. 2018; 63(3): 108-112.
13. Rathnayake D, Clarke M, Jayasooriya L. Hospital surge capacity: The importance of better hospital pre-planning to cope with patient surge during dengue epidemics – A systematic review. *International Journal of Healthcare Management*. 2019; 1-8 (first published online 21 November 2019).