

CHAPTER 1

Randomized controlled trials: the basics

What is a randomized controlled trial?

The randomized controlled trial (RCT) is one of the simplest, most powerful, and revolutionary tools of research.^{1,2} In essence, the RCT is a study in which people are allocated 'at random' to receive one of several interventions.

The people who take part in RCTs (the '*study population*') are called '*participants*' or, irritatingly to some people, '*subjects*'. Participants do not have to be patients, as a study can be conducted in healthy volunteers, in relatives of patients, in members of the general public, in communities, or institutions. The people who design and carry out the study and analyze the results are called the '*investigators*'. The interventions are sometimes called '*clinical maneuvers*', and include varied actions such as preventive strategies, diagnostic tests, screening programs, and treatments. For instance, if we are conducting a study in which patients with rheumatoid arthritis who are randomized to receive either ibuprofen or a new drug (let us call it '*perfectafen*') for the relief of pain, we and our colleagues would be the investigators; the participants the patients with rheumatoid arthritis; and the interventions ibuprofen and perfectafen.

Typically, RCTs seek to measure and compare different events called '*outcomes*' that are present or absent after the participants receive the interventions. Because the outcomes are quantified (or measured), RCTs are regarded as '*quantitative*' studies. In our hypothetical RCT comparing ibuprofen and perfectafen, for instance, the investigators could select pain as the main outcome, measuring it in terms of the number of patients who achieve complete relief 1 week after starting treatment.

Because RCTs are used to compare two or more interventions, they are considered '*comparative*' studies. Usually, one of the interventions is regarded as a standard of comparison or '*control*', and the

group of participants who receive it is called the '*control group*'. This is why RCTs are referred to as randomized '*controlled*' trials. The control can be conventional practice, a placebo, or no intervention at all. The other groups are called the '*experimental*' or the '*treatment*' groups. In our example, the experimental group is the group that receives 'perfectafen' (the new treatment) and the control group is the one that receives ibuprofen, the standard treatment. Some trials could compare different doses of the same medication, or different ways of administering the intervention as part of either the experimental or control groups.

RCTs are '*experiments*' because the investigators can influence the number and the type of interventions, as well as the regimen (amount, route, and frequency) with which the interventions are applied to the participants. This is in contrast to other types of studies, called '*observational*', in which the events are not influenced by the investigators. We describe these, briefly, in Chapter 7.

In summary, RCTs are quantitative comparative controlled experiments in which a group of investigators study two or more interventions in a series of individuals who are randomly '*allocated*' (chosen) to receive them.

What does random allocation mean?

Random allocation means that participants are assigned to one of the study groups by chance alone.³ The decision as to which group they will be in is not determined or influenced by the investigators, the clinicians, or the study participants.

Despite its simplicity, the principle of randomization is often misunderstood by clinicians, researchers, journal reviewers, and even journal editors. Methods to allocate participants according to date of birth (odd or even years), the number of their hospital records, the date in which they are invited to participate in the study (odd or even days), or alternately into the different study groups should not be regarded as really generating random allocation sequences. Although if no one cheats, these 'non-random' or 'quasi-random' studies could produce well-balanced groups, knowledge of the group to which a participant is destined can affect the decision about whether to enter him or her into the trial. This could bias the results of the whole trial.⁴

What is the purpose of random allocation?

By allocating the participants randomly, the characteristics of the participants are likely to be similar across groups at the start of the

comparison (also called '*baseline*'). By keeping the groups '*balanced at baseline*' (as similar as possible at the beginning of the study) the investigators will be more able to isolate and quantify the impact of the interventions they are studying, while minimizing effects from other factors that could influence the outcomes (these are called '*confounding factors*').

Either known or unknown factors not related directly to the interventions can influence the outcomes of a study. It is fairly easy to match the groups for possible confounding factors, when we know about them. The groups can be kept balanced without randomization as long as all the possible confounding factors have been measured. For example, if 'perfectafen' is evaluated in a retrospective study, the investigators could select a group of patients who received ibuprofen and who took antacids that would match the proportion of patients who took antacids and received 'perfectafen'. But we cannot match groups for factors about which we are not aware. The value of randomization is that if it is done properly, it reduces the risk of serious imbalance in important unknown as well as known factors that could influence the clinical course of the participants. No other study design allows investigators to balance these unknown factors.

The risk of imbalance among the groups is not abolished completely, even if the allocation is perfectly randomized. There are many types of bias that can influence the composition and characteristics of the study groups, even before a trial begins and long after it is completed. We discuss these biases in Chapter 3.

How can randomization be achieved?

We can generate random sequences of allocation in several different ways. Regardless of the method used, investigators should follow two principles: first, they must define the rules that will govern allocation; and second, they should follow those rules strictly throughout the whole study.

In principle, the simplest methods to generate random sequences of allocation are 'flipping a coin' (for studies with two groups) and 'rolling a die' (for studies with two or more groups), although they are rarely used because they do not leave an audit trail.

Investigators can also use 'random number tables' to generate the sequences. Random number tables contain a series of numbers which occur equally often, and that are arranged in a random (therefore unpredictable) fashion. The numbers usually have two or

more digits. The use of a random number table forces investigators to decide the correspondence between the numbers and the groups (e.g. odd corresponding to group A and even to group B; or numbers from 01 to 33 to group A, from 34 to 66 to group B, and from 67 to 99 to group C). Then they have to select the starting point in the table (i.e. the beginning, the end, or any point in the middle of the table marked by a pencil dropped with the eyes closed) and the direction in which the table will be read (e.g. upward or downward). If the numbers in the table contain more than two digits, the investigators have to select the position of the numbers that will determine allocation. For example, if the table contains numbers with four digits (e.g. 2314, 5781, 6703, 8092), the investigators can choose, for example, the last two digits, or the first two, or the first and third. The crucial point is to first define the procedure, and then, once the procedure is defined, do not modify it at any point during the study.

A similar set of numbers may be generated by a computer that is programmed to do so, or by most scientific calculators. The procedures and rules that the investigators must follow are identical to those described for the random number tables.

Regardless of the method the investigators use to generate random sequences of allocation, the number and characteristics of the participants allocated to each of the study groups will probably differ (although slightly) at any given point during the study.³ To minimize these differences, investigators can use some strategies known as '*restricted* (or *block*) randomization', or '*stratified randomization*'.

Restricted randomization is used to keep the numbers of participants in all the study groups as close as possible. It is achieved by creating 'blocks' of sequences that will ensure that the same number of participants will be allocated to the study groups within each block. For example, in a study with three groups (A, B, and C), the investigators can create six blocks: ABC, ACB, BAC, BCA, CAB, and CBA.

Stratified randomization is used to keep the 'characteristics' of the participants (e.g. age, weight, or functional status) as similar as possible across the study groups. To achieve this, investigators must first identify factors (or 'strata') that are known to be related to the outcome of the study. Once these factors are identified, the next step is to produce a separate block randomization scheme for each factor to ensure that the groups are balanced within each stratum.

On occasion, investigators may not desire the same number of participants in each of the study groups and can decide to allocate

unequal numbers to each group, while preserving the homogeneity of the distribution of the characteristics of the participants across the study groups. This is called '*weighted*' or '*unequal*' randomization. This type of randomization tends to be used by investigators who wish to expose fewer participants to the experimental group because of concerns about unexpected adverse events. In the example of ibuprofen versus perfectafen, the investigators may decide to allocate one patient to perfectafen for every four patients who receive ibuprofen.

Unfortunately, the methods of allocation in studies described as 'randomized' are sometimes poorly reported, and sometimes not reported at all, even when such studies are published in prominent journals.^{5,6} Because of these poor descriptions, it is not possible to determine, on most occasions, whether the investigators used a proper method to generate random sequences of allocation. Also, even when the reports of studies described as randomized provide details of the methods of allocation, it has been shown that 5%–10% do not use methods that generate random sequences.^{7,8} The reporting of randomization and other aspects of RCTs will be discussed in detail in Chapter 5.

What can be randomized in RCTs?

The most frequent unit of allocation in RCTs is individual people, either patients (the commonest) or caregivers (e.g. treating physicians or nurses). But other units can equally well be randomized to answer specific questions.

Sometimes it is more appropriate to randomize groups of people rather than individuals. This is known as '*cluster*' randomization. Examples of these clusters are hospitals, families, and geographic areas. Investigators frequently use this approach when the RCTs are designed to evaluate interventions that may affect more than one individual within a particular group (e.g. RCTs evaluating the effect of a videotape on smoking cessation on prison inmates, or the effects on parents following a policy of early discharge from hospital after childbirth). It is also used when the way in which the participants in one study group are treated or assessed is likely to modify the treatment or assessment of participants in other groups. This phenomenon is known as '*contamination*'. For example, contamination would be present in an RCT comparing the use of a booklet describing strategies to increase patient participation in treatment decisions versus conventional practice, if patients who have received the booklet shared it with patients who did not.

In other cases, investigators may decide to randomize not only individuals or groups of individuals, but also the order in which the measurements are obtained from each participant. For instance, in an RCT evaluating the effects of morphine on cancer pain, the investigators could randomize the order in which analgesia, adverse effects, and quality of life are assessed.

When are randomized trials needed?

Randomized trials are needed to determine the effects of a health care intervention when these effects are not absolutely clear from observational studies. The effects of some health care interventions, such as antibiotics for pneumonia, or Cesarean section for an obstructed labor, are so dramatic that no further testing is required. More often the effects are less dramatic and may be highly influenced by external factors. Small to moderate effects of interventions can be very important, if the health problem is serious or common.

How are RCTs used?

When reading a trial protocol or a report, it is always wise to consider the purpose of the trial. The theoretical purpose of an RCT is to promote health through a better understanding of the benefits or harms of one or more interventions. A well-conceived, well-performed RCT can inform, enhance, and sometimes change clinical practice or policy. Trials can help individual clinicians to guide their practice, and clinical communities to determine or modify practice patterns. They can provide patients and the public with the information to help them choose what they feel to be the best for them as individuals. Government agencies utilize RCTs for approval of drugs or devices. Insurance agencies, private or government, use them to determine which services or procedures warrant insurance coverage. Institutions can use them to make health policy decisions.

RCTs can, of course, also be used for other purposes. They may be carried out for career advancement or purely for curiosity. They may be funded by companies (most often pharmaceutical, but increasingly also the manufacturing of devices) for regulatory and marketing purposes. They also serve as a powerful form of rhetoric to convince skeptics and doubters, or to control trends that could be considered as too expensive or too disruptive.

How are trials managed and overseen?

Major attention is usually given to when and how RCTs are conceived, designed, and analyzed. All too often, however, too little attention is paid to the actual ongoing meticulous management and oversight of a clinical trial.

Ideally, all activities within a trial must be guided by a *'protocol'*, a document that outlines the research question, the rationale for the trial, and the systems that must be set up for recruitment of participants, randomization, data entry, filing, and analysis. These must be clearly established and understood by everyone concerned.

Trials are conducted by research teams led by someone known as the *'principal investigator'*, a person who is able to command the respect of fellow collaborators, other clinicians, and the rest of the trial management team. A key member of this team is the *'trial coordinator'*, the person responsible for the day-to-day management of the trial and who must be able to respond to the problems that inevitably arise. In addition to the principal investigator (usually known as the *'PI'*) and the coordinator, the team often includes research assistants, statisticians, data managers, administrative staff, and, increasingly, computer programmers. This team is responsible for ensuring the highest possible levels of quality during patient recruitment, data collection and analysis, and knowledge transfer.⁹

Collecting information and entering it on a computer is relatively simple. Ensuring that the data are valid and sensible is a complicated and detailed process. This often requires lateral thinking, flexibility, good communication, and a great deal of common sense.

The Internet is now playing a larger role in the management of trials, challenging the traditional roles of (and even the need for) each of the members of the management team. An increasing number of tools now allow posting of protocols on the World Wide Web, self-matching by potential trial participants, automatic computer-generated randomization codes, data entry and analysis, results reporting, and audit. Many of these tools are driven by commercial interests and are undergoing rapid transformation under the impetus for market dominance. Governments and academic groups as well are also starting to support the use of online tools. One of the main challenges in the foreseeable future will be to achieve standardized ways to handle each of the components

of a trial online, to promote economies of scale, and efficient and equitable access and exchange of knowledge worldwide.

Can RCTs answer all questions related to health care interventions?

Although RCTs are considered ‘the best of all research designs’¹⁰ or ‘the most powerful tool in modern clinical research’¹¹ they are by no means a panacea to answer all health care questions. There are many situations in which they are not feasible, necessary, appropriate, or sufficient to help solve important problems.

The term ‘*intervention*’ is widely used in health care, but infrequently defined. On most occasions the term intervention refers to treatment. However, as we discussed at the beginning of this chapter, this term can be, and often is, used in a much wider sense, to include any health care element offered to the study participants that may have an effect on their health status. Examples include preventive strategies, screening programs, diagnostic tests, the setting in which health care is provided, or educational models. Some of these may be difficult or impossible to study with the methodology of an RCT.

Even when RCT evidence is available, it may not be sufficient to provide all the answers that clinicians, patients, or policy makers need.^{12,13} In these cases, we may either require further trials, or use other types of studies to complement the information provided by available RCTs. We discuss other study designs and other types of information, with their advantages and disadvantages, in Chapter 7.

There are many questions for which RCTs are not appropriate. These are usually related to aspects of health care that cannot or should not be influenced by the investigators, such as issues related to the etiology, natural history of diseases, or when the outcomes of interest are adverse effects. It would be unethical and wrong, for instance, to design an RCT in which people would be randomized to smoke or not for decades to compare the prevalence of lung cancer between smokers and non-smokers.

In other circumstances, RCTs may not be worthwhile because of financial constraints, low compliance rates or high drop out rates, or long intervals between the interventions and the outcomes. It would not be possible to carry out an RCT to evaluate the effects of an intervention with very rare outcomes or with effects that take long periods of time to develop. In these cases, other study designs such as case-control studies or cohort studies are more appropriate.

Most RCTs focus on clinical questions and management of disease. Many of the major determinants of health or illness, such as absolute or relative poverty, social class, literacy, transportation or other infrastructure, are not amenable to medical interventions. RCTs can only answer questions for which quantitative results are applicable. A research focus on the types of problems that can be addressed by RCTs can divert our attention and resources from other, equally important health-related problems. Many things that really count cannot be counted.

It follows that before we start reading an RCT, or even searching for one, we should take into account that there are other study designs that may be more appropriate to answer our particular questions. In addition, one RCT in isolation, even when it is appropriate and perfectly designed, is unlikely to provide all the answers we need. We should consider the information provided by a single RCT as an important piece in a puzzle with many empty spaces. This information will have to be assessed and used in conjunction with other types of information (e.g. data from other RCTs or from other study designs, and our own experience), and the values and preferences of the people involved in the decisions, depending on the circumstances in which the decisions are being made.

Our musings

It is very difficult to convey, at the same time, the strengths of RCTs, the value that they have, the risks of over-reliance on them, or their abuse. These concepts swing back and forth as a pendulum. As one of the pioneers of controlled trials, Sir Austin Bradford Hill, put it 'when we think that RCTs can provide all the answers, that doesn't mean just that the pendulum has swung too far, but that it has swung completely off the hook'.¹⁴

Following the celebration of the 50th anniversary of modern trials, several articles drew attention to the way in which these powerful tools had been hijacked by special interest groups, reducing their ability to provide valid, precise, and relevant answers to important questions.^{2,15,16} Since then, these warning calls have been reinforced by highly visible examples of misconduct among funders, policy makers, and researchers, as well as by articles and books by former editors of prominent journals about the current levels of corruption and unethical behavior that exists within the research engine that fuels the drug regulatory process.¹⁷⁻¹⁹

We now feel that there is an increasing polarization of views about RCTs, along a spectrum of views that ranges from those who put trials on the pedestal of the hierarchy of evidence to those who consider RCTs a dangerous distraction. At the dawn of the 21st century, as the complexity of most health care issues increases^{20,21} we have come to realize that trials are valuable sources of knowledge, but not always the most important or even trustworthy ones. One of our greatest challenges will be to learn not only how to carry out scientifically sound and morally ethical RCTs, but why and when to do them.

References

1. Silverman WA, Chalmers I. Sir Austin Bradford Hill: an appreciation. *Controlled Clinical Trials* 1992;13:100–105.
2. Jadad AR, Rennie D. The randomized controlled trial gets a middle-aged checkup. *Journal of American Medical Association* 1998;279(4):319–320.
3. Altman DG. *Practical Statistics for Medical Research*, 2nd edition. London: Chapman & Hall, 2006.
4. Schulz KF. Subverting randomization in controlled trials. *Journal of American Medical Association* 1995;274:1456–1458.
5. Altman DG, Doré CJ. Randomisation and baseline comparisons in clinical trials. *Lancet* 1990;335:149–153.
6. Moher D, Fortin P, Jadad AR, Jüni P, Klassen T, Le Lorier J, Liberati A, Linde K, Penna A. Completeness of reporting of trials in languages other than English: implications for the conduct and reporting of systematic reviews. *Lancet* 1996;347:363–366.
7. Mosteller F, Gilbert JP, McPeck B. Reporting standards and research strategies for controlled trials: agenda for the editor. *Controlled Clinical Trials* 1980;1:37–58.
8. Evans M, Pollock AV. Trials on trial: a review of trials of antibiotic prophylaxis. *Archives of Surgery* 1984;119:109–113.
9. Farrell B. Efficient management of randomized trials: nature or nurture. *British Medical Journal* 2006;317:1236–1239.
10. Norman GR, Streiner DL. *Biostatistics: The Bare Essentials*. Hamilton: Decker, 2000. Smith R. The trouble with medical journals. London: Royal Society of Medicine Press, 2007.
11. Silverman WA. Gnosis and random allotment. *Controlled Clinical Trials* 1981;2:161–166.
12. Naylor CD. Grey zones of clinical practice: some limits to evidence-based medicine. *Lancet* 1995;345:840–842.
13. Freemantle N. Dealing with uncertainty: will science solve the problem of resource allocation in the UK NHS? *Social Sciences and Medicine* 1995;40:1365–1370.

14. Hill AB. Heberden Oration 1965: reflections on the controlled trial. *Annals of Rheumatic Diseases* 1965;25:107–113.
15. Horton R. The clinical trial: deceitful, disputable, unbelievable, unhelpful, and shameful – what next? *Controlled Clinical Trials* 2001;22:593–604.
16. Smith R. Medical journals and pharmaceutical companies: uneasy bed-fellows. *British Medical Journal* 2003;326:1202–1205.
17. Angell M. *The Truth About the Drug Companies: How They Deceive Us and What to Do about It*, 1st edition. New York: Random House. 2004.
18. Kassirer JP. *On the Take: How Medicine's Complicity with Big Business Can Endanger Your Health*. USA: Oxford University Press, 2005.
19. Smith R. *The Trouble with Medical Journals*. Royal Society of Medicine Press. 2007.
20. Plsek PE, Greenhalgh T. The challenge of complexity in health care. *British Medical Journal* 2001;323:625–628.
21. www.healthandeverything.org (accessed December 22, 2006).