



The RCT: a very beautiful technique

The 2009 Janet Kinson Lecture

V Coates*

Introduction

Taking my lead from those who have presented before me, I understand that the ethos of the Janet Kinson Lecture is that 'it enables reflection on areas that are important to the speaker, and in so doing demonstrates their education and learning'.¹ I have worked in diabetes education and research over the last 15 or so years, and the subject chosen for the lecture at the 2009 Diabetes UK Annual Professional Conference was derived from these experiences.

Clinicians are expected to use evidence as a basis for practice whenever possible and randomised controlled trials (RCTs) are held up as the gold standard when designing research to test interventions. While the strengths of the design are not disputed when testing new medications, diabetes management employs a wide array of activities with a behavioural component, and these interventions must also be evidence based. However, there are many challenges when designing studies to investigate interventions with a behavioural component. My presentation focused on some of the issues to be faced when designing RCTs to determine the efficacy of behavioural interventions, and drew on examples from several RCTs in diabetes in which I was involved. The ways in which these issues may compromise external validity will also be considered.

Rationale for topic selection

Traditionally, health care was based on conventional practice, local wisdom and personal preferences, but the emergence of clinical governance

Table 1. A simple experimental design³

	Pre-test	Intervention	Post-test
Experimental group	O ₁	X	O ₂
Control group	O ₁		O ₂

O = observations that are of interest – subscripts are used to distinguish between specific time points.

endorsed the need for evidence-based practice. Three crucial factors in health care delivery today are patient safety, meeting targets and cost effectiveness. Therefore, the need to know whether clinical care is effective and economical was never more important. When the focus of care was on pharmacological treatment there was often a wealth of evidence to draw from, but when thinking about many aspects of nursing practice – for example, patient assessment, education, promotion of self-management or empowerment – the evidence base was much more elusive. My own work has been to promote research to underpin nursing practice and this led to my interest in RCTs and the results that they may generate.

Best evidence

The evidence base for nursing needs to be derived from properly conducted scientific research. While there are many research methods available, experimental methods – and, in particular, RCTs – are generally recognised as having the necessary rigour to inform clinical decision making. It has been said that 'The "gold standard" research method for addressing what works in evidence-informed policy making and practice is the randomised controlled trial'.²

Randomised controlled trials

The basic framework for conducting an experiment is simple (Table 1).³ As described by Pawson and Tilley,⁴ you take two evenly matched groups, measure the specific attributes that are of interest, treat one group with the intervention but not the other, then measure the groups again and compare the changes between the two groups. If the two groups are identical at the start and the only difference is the intervention given to the one group, it can be assumed that any differences must be due to the intervention. 'Lo and behold you have a clear measure of the impact of the program'.⁴

The origins of the experimental method

The first randomised experiments were conducted in agriculture in which the experiments were based on plots of land to which different crops or fertilisers were assigned.⁵ Pocock⁵ points out that this enabled the experimenter to have all units available at one time and to tightly control the situation. Randomised experiments to evaluate new treatments gained increasing credibility and became recognised as 'the only reliable basis for evaluating the efficacy and safety of new treatments'.⁵ Despite the agricultural origins of

Professor Vivien Coates, Professor of Nursing Research, Institute of Nursing Research, University of Ulster; Assistant Director of Nursing for the Western Health and Social Care Trust, Northern Ireland

*Correspondence to: Professor Vivien Coates, Room G243, School of Nursing, University of Ulster, Coleraine Campus, Cromore Road, Coleraine, Co Londonderry BT52 1SA, Northern Ireland;

e-mail: ve.coates@ulster.ac.uk

Received: 19 October 2009

Accepted: 20 October 2009



this design, one of the great champions of this method was Dr Archie Cochrane who established a register of RCTs in health care research to enable health care professionals to access robust research in order to inform their practice. This was eventually developed into The Cochrane Library that provides 'high quality, independent evidence to inform health care decision making'.⁶

The strengths of RCTs are not in question when, for example, testing out a new medication in which one group takes the new product whilst the other group takes the placebo. However, RCT methods can prove difficult to design in the case of complicated behavioural interventions, many of which are at the heart of clinical practice in diabetes. This is not a new challenge – indeed, the title of this Janet Kinson Lecture was based on a quotation from Dr Cochrane who stated: 'The RCT is a very beautiful technique, of wide applicability, but as with everything else there are snags.'⁷ Six areas prone to 'snags' are outlined below.

1. Complex interventions

In the words of Lindsay,⁸ a complex intervention may be characterised by '... actions that are difficult to define and by varied, and difficult to control, contextual factors'.⁸ Such characteristics would often apply to behavioural interventions designed to improve diabetes management. For example: 'Do diabetes specialist nurses (DSNs) make a difference to diabetes outcomes?', 'Are nurse-led clinics as effective as consultant physician-led clinics?', and 'What are the effects of patient education on diabetes management?'

Complex interventions are influenced by multiple related and interdependent variables such as practitioner behaviour, patient behaviour, timing and frequency of behaviour, organisational issues, the setting and location, and local culture.

When conducting an experiment, clearly defined and delineated behaviours usually need to be the focus in order to provide the necessary levels of accuracy of interpretation and measurement. Yet, to do so may require the complexity of diabetes management

and its susceptibility to a range of external influences to be tightly controlled and for other factors to be excluded from the research design. Standardising procedures, a fundamental criterion of the RCT, is virtually impossible when highly complex treatment interventions are involved.

As an example, Glasgow⁹ presents a pyramid of social-ecological perspectives that may both influence and impact upon outcomes relevant to diabetes self-management education. The pyramid includes what he terms formal spheres of influence such as health care systems, work sites and policy actions, as well as informal ones including social, physical and environmental factors. All these factors may influence the behaviour of an individual involved in a clinical trial and, conversely, a trial would need to take into account an array of outcomes, not only the physiological ones that tend to be most readily measured. The inclusion of such an array of variables was not envisaged when the initial experimental design was developed and in which the only item to change was the intervention itself.

2. The influence of context

In an experiment all contextual variables are controlled, but in the health service this is difficult – indeed, it would be an over-simplification to assume that all contextual variables can be controlled in the real world. In my own locality (Northern Ireland) we have just completed a major re-organisation of the health service in which 18 health and social care trusts have been amalgamated into five trusts. Imagine the effect that this would have upon an experiment, if midway through the study the entire organisational structure were to change and with it trust boundaries, local practices and possibly locality of service provision.

Another form of contextual influence is that of policy change. Within the world of diabetes care a good example is that of the new General Medical Services contract and the Quality and Outcomes Framework in which GPs were remunerated according to the extent to which quality targets were achieved. This policy led to an improvement in the average HbA_{1c} across the country. If you were

mid-trial with a study designed to improve HbA_{1c} when this policy was introduced, it would be impossible to say whether differences detected – or, indeed, not detected – between the two groups were due to changes in policy rather than due to the effects of the intervention *per se*.

Another example of the changing context of care was encountered in the middle of an RCT (the ESMON study) regarding the effects of self-monitoring of blood glucose in those newly diagnosed with type 2 diabetes.¹⁰ A major marketing campaign was launched by one of the medical devices companies to promote the use of blood glucose meters on breakfast television. This meant that patients who had been randomly allocated to the control group then were influenced by the advertising campaign and wondered if they should have been given a meter, and some asked if they could switch to the other group. This issue was overcome, but it serves to illustrate that even with a carefully planned study it is not possible to predict the impact that outside influences may have upon the subject under scrutiny.

3. Need for clarity

It is generally recognised that: 'A clear definition, the ability to control outside factors and standardisation of the intervention are cornerstones of the RCT'.¹¹ The ability to define clearly what exactly is being investigated has often proved to be a challenge in behavioural interventions. If we consider some of the issues that are of relevance to diabetes management – such as self-care decision making, self-management, patient empowerment, depression and anxiety, socio-economic position, and quality of life – it can be seen that either the nature of the intervention or the outcomes could be open to interpretation; therefore, they need to be carefully defined in the context of the study. The difficulty of clarifying such variables was illustrated in a literature review by Paterson *et al.*¹² to investigate the assumptions underlying the traditional conceptualisation of self-care decision making. They found that the concept was not static, but that it waxed and waned and was influenced by many factors. Such



abstract concepts need to be clearly defined if the results are to be properly understood and if the intervention is to be replicated.

While this may sound like an obvious point, an analysis of 47 RCTs of complex nursing interventions from America, Australia and Europe, published over a two-year period, found that none of the reports gave sufficient detail about the intervention to allow them to be replicated.⁸

4. A representative sample

When recruiting patients to a study it is important to attain a sample that closely resembles the patient population from which they are drawn. However, all clinical trials have specific inclusion and exclusion criteria. The rationale for this is that individuals must be appropriate for the study and that the impact of the intervention should not be masked or biased by participants having other important attributes that may confound the effects. Consequently, it has been found that studies tend to recruit people who are not too young and not too old, are English speaking, and preferably have no other health care problems.

Following an analysis of research on self-management in type 2 diabetes, it was found that people who were most at risk of complications were also most unlikely to be attracted to or remain in the experimental programme.¹³ Furthermore, samples tend to: comprise well-educated individuals (or at least literate); include those with the desire and ability to participate in the intervention (whatever that might require); and under-represent ethnic minorities. Therefore, RCTs may be conducted on a sample of people who are not typical of the general patient population.

5. Randomisation

Randomisation is, according to the CONSORT statement, 'the process of assigning participants to groups such that each participant has a known and usually an equal chance of being assigned to a given group'.¹⁴

Randomisation is one of the essential criteria in an RCT, yet – as highlighted by Snowden *et al.* – it is known that randomisation can be a difficult concept for potential participants to

understand.¹⁵ The actual nature of the trial is often poorly understood; the random basis of allocation and its rationale are often problematic. Featherstone and Donovan¹⁶ were also interested in patients' perspectives of participation in an RCT. They interviewed 22 men who had been involved in a large RCT and found that most struggled to make sense of aspects of the trial.

The process of randomising is also important for those running the trial. In the past it was acceptable for the researchers involved in the study to conduct the randomisation. However, it was realised that randomisation could be fraught with bias created by those running the trial. According to Schulz: 'Researchers need to realise that, given the opportunity, trial implementers will frequently subvert the intended aims of random assignment.'¹⁷ Thus, it was recommended that steps be taken to ensure that those with an interest in the outcomes of the clinical trial cannot influence the allocation of participants to the experimental or control groups. It is now generally agreed that the business of randomly assigning patients to the different groups should be done at a distance by a third party. Methods by which this can be acceptably achieved are included in the CONSORT statement.¹⁴

Patient preferences

However, in conducting several trials with a behavioural component it became clear to me that randomisation may not always be the answer. If patients have a strong preference for one group or another it can be a snag as it may affect outcome. Although allocated to an intervention group, it does not mean that the individual will undertake the specified behaviours as per protocol and this is particularly challenging if not allocated to the group of their choice (if they had one).

This was our experience when running the ESMON study.¹⁰ Although the nature of the trial was explained and informed consent obtained before participation in the study, it was found that some of the patients did not fully understand the implications of what they were agreeing to. For example, we asked those

in the intervention group to follow a monitoring protocol which necessitated the following:

- Patients asked to self-monitor blood glucose twice a day for four days per week (one fasting and one after main meal).
- Results to be recorded in a diary.
- Training in meter use by the DSN – technique observed.
- Advised on interpretation of results: responses to food/exercise/sick day rules.
- Advised to contact DSN if fasting readings five out of seven days were >10mmol/L or post-prandial >12mmol/L.
- Asked to confirm symptoms of hypoglycaemia with a blood glucose result prior to treatment if possible.

Conversely, those in the control group were asked not to monitor for the duration of the trial which was to last for one year. However, at the end of the study it was explained that those in the intervention group could stop monitoring, if desired, while those in the control group would be shown how to monitor if they wished to do so. Although all was explained by the DSNs who knew the patients, there were still occasions when a patient found they were not in their preferred group and asked if they could withdraw from the study as they actually only wanted to take part if they could be in the other group. Naturally, the patients were free to withdraw but the experience served to illustrate that, even though we thought we were explaining the nature of the study carefully, still there were those who did not grasp the basic concept.

The need for and implications of allowing for patient preference in clinical trials have been debated.² On the one hand the implications of employing RCT methodology for all research questions can be questioned, while on the other is the view that only a fully randomised trial is sufficiently robust.

While views vary regarding the design and effects of acknowledging patient preference, some of the debate relates to the introduction of new treatments. In complex trials involving behaviour the situation is more complicated. In practice, clinical judgement would be used and



care individualised. While not advocating that we only practise according to clinical judgement, there are some interventions that are so dependent on patient motivation and cooperation to undertake behaviour change that it seems unwise to take no account of preference and to allocate such patients to a group they would not choose to join. For example, when designing a study to investigate 'The effect of an exercise support group on risk factors for renal disease in a diabetic population', it was agreed that patients meeting inclusion criteria could self-select to either the exercise group or the control group (usual care) as it seemed ridiculous for patients who were willing to participate in the study, but who did not wish to take exercise, to be allocated into the exercise group. However, when the project was submitted for ethical approval, the Research Ethical Committee responded that, for the purposes of validating and concluding on the study results, randomisation into study groups would be preferential and further explanation and greater detail were requested regarding the reasons for not randomising study subjects. Although self-selection seemed justifiable to the clinicians and the research team in a situation where patient motivation would play a large part, it was challenged during the ethical process.

6. The control group

In drugs trials a placebo drug is used for those people not in the intervention group; as the placebo can be made to look like the trial drug, neither the patient nor the staff know who is in the intervention/control groups. However, the situation is quite different in behavioural research as it is unlikely that the intervention can be disguised or that the researchers can remain unaware of the group to which the participants belong.

Waber *et al.*¹⁸ illustrated how individuals can be subjectively influenced during the course of a trial when they tested the pain relief from a placebo pill. All the pills looked the same – one group of participants were told the price of the pills was \$2.5 each, while those in the other group were told they cost \$0.10. Electrical shocks

were given to the wrists of the participants, calibrated to each individuals' pain tolerance and their perception of pain rated on a scale. It was found that pain reduction was greater in the group who thought they were receiving the more expensive pill. (Figure 1 provides an illustration of the trial.)

In a clinical trial the control group is often reported to receive 'usual' care. The notion of providing usual care in the context of a trial is deceptively simple. What exactly is usual care? Is it standardised? In a multi-centred trial does usual care differ and, if so, does it matter? Is it open to bias by the clinical team who do not want their care to look deficient? The importance of the design of this comparative group is illustrated when study results reveal that significant changes have taken place in both groups. If the intervention allows patients to have more time with clinicians than those in the control group, is it the extra time rather than the actual intervention that makes the difference?

Dr TC Skinner pondered this point in an editorial entitled: 'What does make the difference?' in relation to RCTs on patient education. He suggested that: '... when enthusiastic and motivated health care professionals invest more time with patients in a structured way to help them manage their diabetes, we see improved biomedical outcomes'.¹⁹

The design of the control group needs to be thought through to

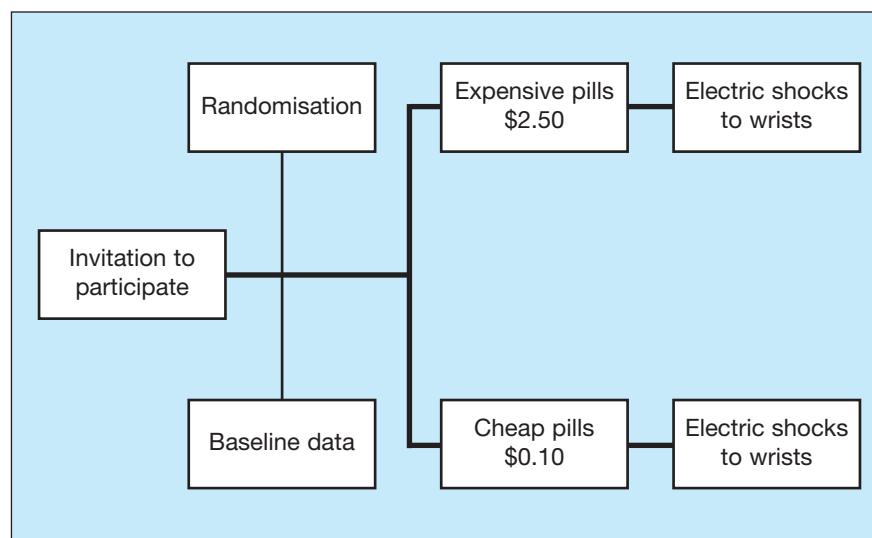
ensure that it does provide an appropriate comparison with the intervention group.

Framework for development of complex interventions

Six areas in which snags are likely to occur, and which need to be taken into account when designing RCTs for complex interventions, have been discussed. This is not a novel issue and in recognition of the difficulties in investigating complex interventions the UK Medical Research Council²⁰ produced a framework for the design and evaluation of complex interventions (Figure 2).

It is acknowledged that it would be too ambitious and inappropriate to conduct a definitive RCT without undertaking a series of preliminary studies leading up to this point. According to the framework, theoretical exploration, modelling and exploratory trial work should precede embarking on a full RCT which would then be followed by long-term evaluations. Following this framework would enable some of the issues raised as 'snags' to be explored. Few research teams have been in a position to work through this framework; however, the research of Mühlhauser and Berger²¹ into patient education is exemplary in this area. Nevertheless, they do report that: 'The time span for gathering the evidence from the theoretical phase to surveillance after implementation was about 20 years.'²¹ This time frame helps emphasise

Figure 1. An illustration of the Waber *et al.* trial¹⁸



that the design issues, when investigating complex interventions, need a long-term commitment and that the problems to be overcome should not be underestimated if results are to be robust.

External validity

The final point of my presentation was to consider the design of RCTs and how it can have an effect upon the external validity of the study. A study has high external validity if the results can be generalised to other people. Berk says: ‘It cannot be overemphasised that unless an experiment can be generalised at least a bit, time and resources have been wasted. One does not really care about the results of a study unless its conclusions can be used to guide future decisions.’²²

In order to achieve the necessary levels of control in the experiment, RCTs tend to have high internal validity at the expense of external validity. There are several threats to external validity, but sampling and intervention support can serve as examples.

Firstly, when sampling, inclusion and exclusion criteria are useful as they eliminate other confounding variables. However, if too many variables are excluded, they may generate a final sample that does not represent the population from which it is drawn and the intervention, if found to be effective, may not be of benefit to the wider population.

Secondly, the nature of the intervention and the amount of support it requires are also very important. If the intervention requires extra resources, it may be possible to apply these in a research setting while staff or additional support have been funded, but such interventions are less likely to be translated to general health care. If the experiment is conducted in a somewhat artificial environment, the RCT may not mirror everyday reality and this may be particularly true if they are only conducted in ‘centres of excellence’.

The appropriateness of RCTs for complex, multi-level interventions – such as diabetes management – needs to be considered in the light of

the challenges of designing a robust study that can overcome, or at least allow for, snags such as have been outlined above.

Conclusion

- When using evidence created by RCTs for complex interventions be aware of design issues that may mean the ‘gold standard’ has not been achieved.
- When using published evidence to plan the care for a patient ask: ‘is the research generalisable to the population from which the research sample was drawn?’ and ‘are the results applicable to this patient?’
- If considering undertaking an RCT – be prepared to be challenged!
- For complex interventions, we need to keep an open mind about other robust means of evaluation.

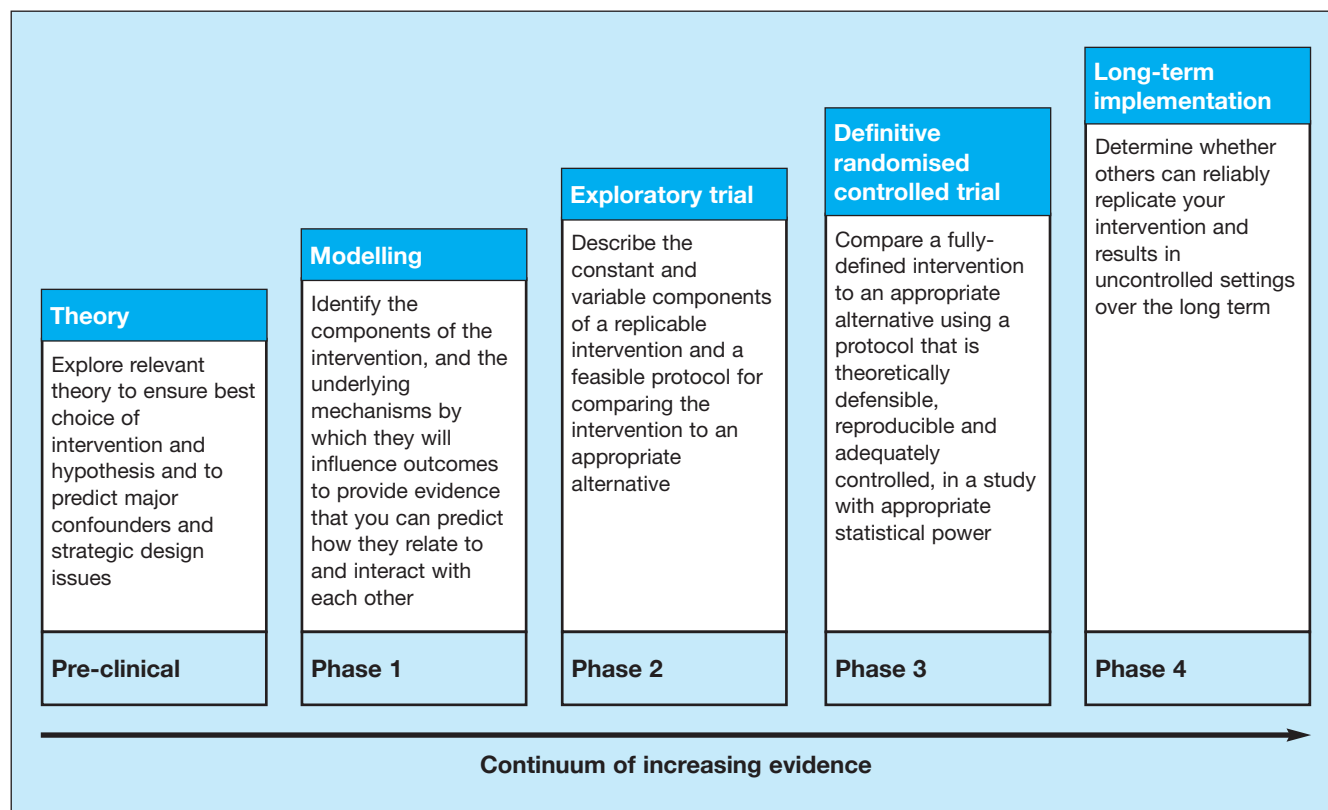
Conflict of interest statement

There are no conflicts of interest.

References

References are available at www.practicaldiabetesinternational.com.

Figure 2. Diagram showing the phases described in the Medical Research Council’s discussion document: ‘A Framework for development and evaluation of RCTs for Complex Interventions to Improve Health’. (Reproduced with permission from the Medical Research Council)²⁰





References

1. Da Costa S. To lead or not to lead: that is the question. The 2007 Janet Kinson Lecture. *Pract Diabetes Int* 2007; **24**(5): 239–242.
2. Torgerson DJ, Torgerson CJ. *Designing randomized trials in health, education and social sciences*. Basingstoke: Palgrave Macmillan, 2008; 1.
3. Campbell D, Stanley J. *Experimental and quasi-experimental evaluations in social research*. Chicago: Rand McNally, 1963.
4. Pawson R, Tilley N. *Realistic Evaluation*. London: Sage, 2008.
5. Pocock SJ. *Clinical trials: A practical approach*. Chichester: John Wiley & Sons, 1983.
6. The Cochrane Library. 2009. www3.interscience.wiley.com/cgi-bin/mrw/home/106568753/HOME?CRETRY=1&SRETRY=0 [accessed 6 Oct 2009].
7. Cochrane AL. *Effectiveness and efficiency: random reflections on health services*. Abingdon: The Nuffield Provincial Hospitals Trust, 1972; 2.
8. Lindsay B. Randomized controlled trials of socially complex nursing interventions: creating bias and unreliability? *J Adv Nurs* 2004; **45**(1): 84–94.
9. Glasgow RE. Outcomes of and for diabetes education research. *Diabetes Educ* 1999; **25**(6 Suppl): 74–88.
10. O’Kane M, Bunting B, Copeland M, *et al.*, for the ESMON Study Group. Efficacy of self-monitoring of blood glucose in patients with newly diagnosed type 2 diabetes (ESMON study): randomised controlled trial. *BMJ* 2008; **336**: 1174–1177.
11. Blackwood B. Methodological issues in evaluating complex healthcare interventions. *J Adv Nurs* 2006; **54**(5): 612–622.
12. Paterson BL, Russell C, Thorne S. Critical analysis of everyday self-care decision making in chronic illness. *J Adv Nurs* 2001; **35**(3): 335–341.
13. Paterson B, Hopwood M. The relevance of self-management programs for people with chronic disease at risk for disease related complications. In *Translating chronic illness research into practice*. Kralik D, Paterson B, Coates VE (eds). Oxford: Wiley Blackwell (in press for 2010).
14. Altman D, Schulz KF, Moher D, *et al.*, for the CONSORT Group. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001; **143**: 663–694.
15. Snowdon C, Garcia J, Elbourne D. Making sense of randomisation; responses of parents of critically ill babies to random allocation of treatment in a clinical trial. *Soc Sci Med* 1997; **45**: 1337–1355.
16. Featherstone K, Donovan JL. ‘Why don’t they just tell me straight, why allocate it?’ The struggle to make sense of participating in a randomised controlled trial. *Soc Sci Med* 2002; **55**: 709–719.
17. Schulz KF. Unbiased research and the human spirit: the challenges of randomised controlled trials. *Can Med Assoc J* 1995; **153**(6): 783–786.
18. Waber RL, Shiv B, Carmon Z, *et al.* Commercial features of placebo and therapeutic efficacy. *JAMA* 2008; **299**: 1016–1017.
19. Skinner TC. What does make the difference? *Diabet Med* 2006; **23**(9): 933–934.
20. Medical Research Council. *A Framework for development and evaluation of RCTs for Complex Interventions to Improve Health*. Discussion document, April 2000. www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC003372.
21. Mühlhauser L, Berger M. Patient education – evaluation of a complex intervention. *Diabetologia* 2002; **45**: 1723–1733.
22. Berk RA. Randomised experiments as the bronze standard. *J Exp Criminol* 2005; **1**: 417–433.

ONLINE ONLY