

Randomisation and chance-based designs in social care research

Bob Woods and Ian Russell

Improving the evidence base for
adult social care practice

Methods Review 17



The School for Social Care Research

The School for Social Care Research was set up by the National Institute for Health Research (NIHR) to develop and improve the evidence base for adult social care practice in England in 2009. It conducts and commissions high-quality research.

About the authors

Bob Woods is Professor of Clinical Psychology of Older People at Bangor University, and Associate Director (Clinical) of NWRTH, the Bangor Clinical Trials Unit. Professor Woods has been a practising clinical psychologist working with older people for over 35 years combined with academic appointments at the Institute of Psychiatry, London and University College London. Current research primarily concerns the development and evaluation of therapeutic approaches for people with dementia and their care-givers.



Ian Russell is Professor of Clinical Trials at Swansea University. He is a career health services researcher who specialises in the design, conduct, analysis and reporting of pragmatic randomised trials to evaluate complex interventions in health care. He was an early protagonist of patient-reported outcome measures in trials and systematic reviews to aggregate the findings of trials. He founded, or laid foundations for, clinical trials units in Newcastle, Aberdeen, York, Bangor and Swansea.



NIHR School for Social Care Research
London School of Economics and Political Science
Houghton Street
London
WC2A 2AE
Email: sscr@lse.ac.uk
Tel: +44 (0)20 7955 6238
Website: www.sscr.nihr.ac.uk

© School for Social Care Research, 2014

This report presents an independent review commissioned by the NIHR School for Social Care Research. The views expressed in this publication are those of the authors and not necessarily those of the NIHR School for Social Care Research, the Department of Health, NIHR or NHS.

Randomisation and chance-based designs in social care research

ABSTRACT

When there is little evidence about the relative effectiveness of two alternative approaches in social care, simple comparisons of the participants who happen to receive one approach with those who happen to receive the other are prone to bias arising from whatever social mechanism was responsible for these allocations. Chance-based designs (CBDs) address this intrinsic problem by allocating participants between approaches at random, thus maximising internal validity. This has led many commentators to describe CBDs (or ‘randomised trials’ as they are called in medicine) as the ‘gold standard’ of health care research and made them the dominant research design in interventions research and in developing ‘evidence-based medicine’ and practice.

However, CBDs are less widely used in social care research in the UK. Indeed, they are generally viewed with suspicion, especially by critics who equate them with drug trials. More specifically, Illsley (1982), a social scientist who chaired the Scottish Health Services Research Committee, argued that CBDs were difficult outside drug trials and that, even when they were feasible, the constraints of the experiment limited their generalisability. Over the past 25 years, health care research has addressed and overcome both of Illsley’s objections. First, there are now more than 750,000 CBDs listed in the Cochrane Library. Secondly, researchers have distinguished between two distinct types of CBD. ‘Fastidious trials’ test whether new approaches, typically new drugs, are more efficacious than the status quo under laboratory conditions. That drug trials are typically ‘double blind’, with neither participant nor doctor knowing whether their drug is active or inactive (or ‘placebo’), was a key feature that led Illsley and others to question their external validity.

In contrast, ‘pragmatic designs’ estimate in normal practice whether the benefits of new approaches, typically more complex interventions, outweigh their costs. More realistically, only those assessing outcomes do not know which intervention each participant is receiving. These pragmatic designs assess the differential effects of the intervention on the lives of participants and thus contribute to future decisions about the allocation of resources and where they are most effective and cost-effective. Such pragmatic designs require clarity about desired outcomes and the development of valid and reliable measures of these outcomes. Furthermore, though CBDs are well suited to evaluating the average effects on participants of a standard intervention, they increasingly assess interventions tailored to individual needs and use participant-centred outcome measures that focus on individual goals and concerns.

In working at the interface between health and social care, and focusing on the care of people with dementia, we have used pragmatic designs extensively. We conclude that such designs have the flexibility and naturalism to contribute effectively to social care research, especially if they exploit and extend the wide range of methods used in health care research. For example, one can assign either individuals or ‘clusters’ of individuals by chance when it would be difficult for some to receive the intervention and others not, for example residents of a care home. However, cluster-randomised designs generally require

Randomisation and chance-based designs in social care research

larger samples and even more care to avoid bias than individually randomised designs.

Critics of CBDs argue that it is unethical to deny 'control' participants access to the new intervention. When there is good evidence that the new intervention is superior, there is no case for a CBD. When practitioners believe that the new intervention is superior, they should not take part in a CBD. When neither of these constraints applies, the authors argue it is unethical not to use the unique but naturalistic rigour of a pragmatic CBD to inform the best use of scarce resources. That said, it is important to ensure that all participants receive at least current 'best practice'. Furthermore, there is a strong case for including service users in research design and management to limit the burden of research and ensure that CBDs are fair to control participants. Sometimes, waiting-list or interrupted time series designs may be appropriate so that all participants receive the experimental intervention.

In short, the increasing pragmatism and rigour of CBD methodology over the past 20 years have enhanced research on complex interventions in health care, notably at the interface with social care. The authors argue that it is time for social care research to exploit this progress.

RECOMMENDATIONS

- Many of the valid concerns raised about use of CBDs in social care have been addressed over several decades, and more consideration should be given to use of CBD for social care.
- CBDs are challenging and require a team approach to set up and establish appropriately, including involvement of service users in research design and management. Careful consideration needs to be given as to how best to use CBD appropriately. However, properly conducted and reported, they may well be cost-effective in terms of the value of knowledge and information gained.
- Open and transparent reporting of CBDs is essential to maintaining and raising their quality, and ensuring that the results and recommendations do justice to the contribution of all those taking part.
- There is a need for further development and testing of outcome measures in many cases before proceeding with a full CBD evaluation.
- It is recommended that work goes on in parallel evaluating the processes of the intervention and its effects within the study, and wherever possible, an evaluation of cost-effectiveness should be carried out as part of the CBD.

Randomisation and chance-based designs in social care research

KEYWORDS

Social care, complex interventions, evaluation of care, methodology, randomised designs, pragmatic trials, cluster randomisation, stepped wedge designs

ACKNOWLEDGEMENTS

The authors are grateful to Kat Algar for carrying out the preliminary literature searches for this review.

CONTENTS

1	Introduction	1
2	Chance-based designs in social care research	1
3	Chance-based designs – more than drug trials?	3
3.1	The classic drug trial	4
3.2	Fastidious and pragmatic CBDs	5
4	Chance-based designs – are they ethical?	7
4.1	Winners and losers	8
4.2	One size does not fit all	12
4.3	Outcomes that are meaningful	12
5	Chance-based designs – practically impossible?	14
6	Chance-based designs – the gold standard?	19
6.1	Development	21
6.2	Feasibility/piloting	21
6.3	Evaluation	22
6.4	Implementation	22
7	Chance-base designs – only useful for effectiveness?	23
8	Conclusions and recommendations	24
	References	27

1 INTRODUCTION

Social care is complex. We want to ensure those requiring, or in receipt of, social care have the best lives possible, but it can be difficult for the individual practitioner or service to be sure how much difference their involvement has made. If the situation improves, was it just luck or some other change independent of the care package? If there is a change for the worse, does that mean the care offered was inappropriate or ineffective, or again, was it the result of factors and circumstances beyond the control of the practitioner or the service? From the service-user's perspective, how can assurance be gained that the most beneficial care package is being offered?

We know that budgets for social care – as for all public services - are limited, and making the best use of the resources we have will always be a priority. We need to invest in those services that, for a given level of resource, make the most difference and achieve the outcomes that are most valued by the community. Conversely, identifying those services and interventions that are ineffective allows the potential of re-directing resource to where it can have greater impact.

It is, therefore, in the interests of practitioners, service-users, services and society as a whole to be able to ascertain the connections between services provided and outcomes achieved. Is it possible to design research in the social care context that can help to evaluate the effectiveness of services and interventions and can plausibly establish the likelihood that any changes seen are indeed related to the service rather than to extraneous factors? This review examines study designs involving the allocation by chance of participants or participating services to receive, or not receive, the service or intervention of interest. This is a methodology that has the potential to fulfil this demanding brief, but which has generally not found favour in the domain of social care research. The approach certainly has its limitations, and careful consideration needs to be given as to how best to use it appropriately, but this review will argue that it merits a place in the repertoire of social care research methods. It is worth noting that these designs have been the subject of controversy in other areas of social science, such as education (Cook 2001), while they have more recently proliferated in areas such as economics (Card *et al.* 2011) and politics (Arceneaux and Nickerson 2009).

In the health care context, chance-based designs (CBDs) are typically referred to as 'randomised controlled trials' (RCTs) and have been widely accepted as the cornerstone of the evidence-base for therapeutic interventions in the NHS, informing decisions regarding commissioning of services and investment in treatments, as well as guiding professional practice. In this review, we will examine some of the reasons for the relative neglect of these approaches in social care research. In part, this appears to have arisen from a mythology that has developed about the nature of RCTs, based perhaps on the situation some years ago, and not taking into account the evolution of these approaches in the health-care context. Accordingly, in this review we have preferred the term 'chance-based designs' to encourage a fresh look at the full-picture of what these approaches can and cannot offer.

2 CHANCE-BASED DESIGNS IN SOCIAL CARE RESEARCH

How widely used and accepted are CBDs in social care research and practice in the UK? In preparing this methods review, we examined the social care evidence and knowledge base as assembled by the Social Care Institute for Excellence (SCIE). A systematic review of two domains of this evidence base was undertaken encompassing the 43 SCIE Research Briefings¹ and 23 Knowledge Reviews² on the SCIE website at 1st January 2014. In each case, the relevant document was searched for the terms 'randomised', 'randomized' and 'RCT'. Reference lists were also searched to ensure that no relevant items were missed. In total, 84 papers were identified meeting the search criteria across the 66 documents.

In 58% of the Research Briefings (25/43), no studies were identified that used a CBD. One briefing (28: Assistive technology and older people) specifically commented that no studies were found that used RCTs or, indeed, experimental or quasi-experimental designs without randomisation (p.5); the authors of this briefing highlight the different perspectives of health and social research, suggesting that, traditionally, the former would not see the 'case studies and other research based on observable evidence' as providing the 'solid' evidence that RCTs could offer. Research Briefing 36 (Reablement: a cost-effective route to better outcomes, p.12) suggested that research could be strengthened by the use of randomised designs, and another (39: Preventing loneliness and social isolation: interventions and outcomes) identified an 'urgent need for more longitudinal randomised controlled trials' (p.1). However, a further briefing (26: Mental health and social work) suggests that a medical model 'uncritically assuming the superiority of ways of producing knowledge through randomised control groups and quantitative methods is not always appropriate' and argues that other methods are required to 'include the needs and concerns of people who use services, their carers and the wider mental health community' (p.10). In a similar vein, Briefing 42 (Returning children home from public care) argues that the 'wide range of variables involved in this area of practice' makes it unsuitable for 'random allocation to tightly defined interventions and control groups' (p.3).

The median number of RCTs mentioned in the remaining 18 briefings was 3. Two briefings included 10 or more relevant papers: the topics of these were 'Preventing falls in care homes' (Briefing Paper 1) and 'Therapies and approaches for helping children and adolescents who deliberately self-harm (DSH)' (Briefing Paper 16). Although these two domains might be said to be at the interface of health and social care, a number of topics with a similar degree of overlap had no mention of RCTs, e.g. 'Access to primary care services for people with learning disabilities' (Briefing Paper 2). It may be the focus on specific therapies and interventions that influences whether RCTs are considered.

Almost three-quarters (17/23) of the Knowledge Reviews made no mention of RCTs. Two reviews mentioned a number of RCTs and suggested that more such were needed. These focused on 'the adoption of looked after children' (Briefing Paper 2) and 'recovery

1. www.scie.org.uk/publications/briefings/index.asp

2. www.scie.org.uk/publications/knowledgereviews/index.asp

Randomisation and chance-based designs in social care research

approaches in community-based adult mental health services' (Briefing Paper 21) respectively. Two reviews made passing reference, while the two final reviews had discussion of the pros and cons of RCTs; Briefing Paper 3 'Types and quality of knowledge in social care' stated that their use was mainly confined to therapeutic/ psychiatric interventions, while Briefing Paper 5 'An exploration of the research literature in foster care' critiques RCTs, which do not feature at all, it appears, in the literature reviewed. The suggestion is made that evaluating social work is more complicated than evaluating therapeutic drugs!

Together, these two Knowledge Reviews contribute to the following critique of CBDs (drawn from Knowledge Reviews 3 and 5 and Research Briefings 26 and 28). Given advances in CBD methodology over the past 30 years, however, we see this critique as mythological.

1. CBDs are only really suited to drug trials. Here it is important to control for all other factors and ensure that any change arises from the pharmaceutical agent itself and that changes would not have come simply with the passage of time. In social care, it is more complex to define both the context of the treatment and the treatment itself. Pills can be manufactured to contain a known dose of the chemical involved, which can be given in virtually the same manner in different centres. Social care cannot be so readily standardised. Drug trials involve patients with clearly defined 'problems', typically with a specific diagnosis. Recipients of social care cannot be classified so precisely. Drug trials should be 'double-blind' so that both the person receiving the drug, and the doctor and nurse dispensing it, are 'blind' as to whether the person is receiving the actual drug or a placebo pill. In social care, it would usually be impossible for the practitioner delivering the service or care package not to be aware of whether the person is receiving the active intervention or not.
2. CBDs are associated with ethical problems, e.g. it would be inconceivable to randomly allocate children to adoption or foster care (or indeed older people to enter a care home as opposed to receiving a care package at home). More generally, deciding on the treatment plan for an individual client by the toss of a coin (effectively) is often impossible or unacceptable in practice.
3. CBDs require large sample sizes to have any chance of showing an effect, and there are likely to be practical problems in achieving these in practice.
4. CBDs are seen in the medical model as the gold-standard of research, the pinnacle of research design, and as superior to all other approaches to the generation of knowledge. In social care, weight is given to other sources of evidence, such as case-studies and qualitative methods.
5. CBDs are well-suited to showing whether an effect occurs but are less able to indicate the mechanisms and processes underlying the effect. This makes it difficult to establish what circumstances are needed for the effect to be repeated, and CBDs frequently produce discrepant results when repeated.

3 CHANCE-BASED DESIGNS – MORE THAN DRUG TRIALS?

3.1 The classic drug trial

The critique of CBDs as only really being suited to drug trials echoes that of Illsley (1982), a social scientist who chaired the Scottish Health Services Research Committee, who argued that CBDs were difficult outside drug trials and that, even when they were feasible, the constraints that they required meant that it would be difficult to apply their findings in the real-life context. In the intervening period, literally hundreds of thousands of CBDs of non-pharmacological interventions have been published in the domain of health care. Difficult they may have been, but CBDs are increasingly recognised as achievable by funders of research and by peer-reviewed journals.

What are the constraints imposed by a CBD in the classic stereotypical drug trial?

First, the **participants** are likely to be a carefully selected group. They will be patients who are clearly identified as having the condition for which the drug has been developed. If the diagnosis is uncertain, or the person has other conditions alongside the target condition, the patient will typically be excluded. In the past, it was common to exclude people over a certain age, as they were more likely to have other conditions that would complicate the treatment of the target disease. Participants will need to be prepared to undergo a number of assessments to evaluate the effects of the drug.

Secondly, the **setting** is likely to be carefully selected. The clinic or hospital is likely to be affiliated to a University with good diagnostic and treatment facilities. It may not reflect a typical NHS centre. The patient is likely to be seen more often and more regularly than in usual clinical practice with close attention to any side-effects.

The **treatment** itself will be carefully prepared to ensure that it provides a standard dose of the drug, identical from tablet to tablet and free of impurity. The patient's compliance with taking the medication may be monitored through, for example, blood tests. The patient will be encouraged not to start on any other medication that may interfere with the effects of the drug in question. The dose of the medication being taken will be carefully supervised.

The participants in the drug trial must be prepared to receive a placebo rather than the active treatment. This will be determined randomly, with the randomisation being carried out independently, so that the person metaphorically 'tossing the coin' is not in a position to bias the outcome of the randomisation. A primary aim of this type of trial is to account for and remove the placebo effect where patients improve to some extent simply from taking a pill, even if it contains no active ingredient. Patients' expectations are seen as having an influence on outcomes but are dealt with as a source of interference.

Similarly, clinicians' expectations are viewed as creating a potential bias, hence the need for these trials to be *double-blind* with neither the clinician nor the patient being aware of what the pills being taken contain. The assessment of any changes in the patient's

Randomisation and chance-based designs in social care research

condition and his/her response to treatment is thus carried out without knowledge of whether or not the patient was receiving active treatment.

3.2 Fastidious and pragmatic CBDs

With these constraints, it is only to be expected that the relevance of this kind of drug trial to 'real life' has been questioned. These studies might be described as 'fastidious trials', testing the efficacy of new approaches under laboratory conditions. However, it is mistaken to equate all CBDs (or indeed all drug trials) with these fastidious trials. 'Pragmatic designs' have been developed and are widely used to establish whether new approaches are effective in normal practice (Schwartz and Lellouch 1967, Hotopf 2002, Torgerson and Torgerson 2008). They have often been used to evaluate more complex interventions and will typically have the potential to indicate whether the benefits of the approach outweigh the costs. Pragmatic CBDs seek to include the participants who are intended to be the eventual beneficiaries of the approach (with any exclusions carefully justified) in the typical services where these approaches will be implemented. Although efforts will be made to ensure participants receive the treatment as planned, once the randomisation has been made, the participant's data will be analysed in relation to the treatment they were offered irrespective of whether they participated fully or not ('analysis by treatment allocated' approach). The evaluation accordingly seeks to ascertain the effects on a typical population, in standard services, with the range of compliance and commitment that would ordinarily be encountered.

While the 'dose' of the 'treatment' will vary according to the participant's engagement, efforts will be made to offer the intervention in a standardised, uniform manner as far as is possible in a complex intervention. However, those delivering the intervention will not usually be the experts in the approach – rather, they will be people who are trained and supported to be competent with a similar background to those who will ultimately deliver the intervention if it is found to be effective.

In pragmatic CBDs, the comparison condition is often 'usual care' (see Box 1) with less emphasis on the requirement for a placebo condition. In fastidious trials, it is seen as important to identify the active ingredients in the intervention offered, and so a placebo comparison or comparison with an existing established treatment is essential. In pragmatic CBDs, the interest is in whether the package as a whole – including the expectations of participants and those delivering the intervention as well as other factors not specifically linked to the intervention – has benefits for those receiving it. Pragmatic CBDs often give rise to the query as to whether the benefits are due to the participants in the intervention group receiving more 'attention'. They may, for example, have attended a number of weekly group or individual sessions over and above what those in the 'usual care' comparison condition have received. The pragmatic CBD is able to inform us as to the effects of *that form of attention*; it does not aim to tease out and unpick the relative contribution of the range of intended and unintended aspects of any complex intervention.

Randomisation and chance-based designs in social care research

Box 1: Specialist behaviour therapy team for challenging behaviour in service for adults with learning disabilities (Hassiotis *et al.* 2009)

Background	Services for adults with learning disabilities, often provided jointly by health and social care services, identify challenging behaviour as a significant issue, often placing the person or others at risk, or leading to breakdown of the person’s living arrangements or placement. Anti-psychotic medication is often used, despite the risk of side-effects, in part because there is a lack of robust evidence regarding the benefits of alternative approaches. ‘Standard treatment’ in the study area was offered by a full multi-disciplinary community learning disabilities team, but an additional specialist team had also been established, including behaviour specialists with training in applied behaviour analysis.
Research objective	This study aimed to evaluate the effectiveness of a specialist behaviour therapy team in addition to standard treatment in comparison with standard treatment alone.
Design	Chance-based allocation of referrals meeting the team’s inclusion criteria to either ‘standard treatment’ or the specialist team. Participants had a 50% chance of receiving the specialist service. Participants were 63 male and female service users with mild to severe intellectual disability who presented with challenging behaviour.
Ethical issues	As the specialist team was already in operation, those not allocated to receive this input might be seen as disadvantaged. Accordingly, a 6 month limit was set on maintaining the chance-based allocation, as this was consistent with typical waiting times to receive the specialist team’s input in practice.
Outcomes	The main indicator of change was the Aberrant Behaviour Checklist, completed with the participant and caregiver by an assessor unaware of which treatment the person had been allocated. Service use and associated costs were also evaluated.
Strengths and weaknesses	<p>Pragmatic trial rooted in usual practice. Health economic analysis adds to the findings.</p> <p>A check was made on whether blindness of assessor was maintained (66% accurate in ‘guessing’ which arm participant was in).</p> <p>The six-month limit on maintaining randomisation restricted the ability of the study to estimate longer-term effects.</p> <p>Sample size calculated to be able to show the anticipated effect, but too small to allow analysis of effects of severity of learning disability, for example.</p>
Generalizability	This design could be used to evaluate the effectiveness of any type of additional service provided by a specialist team in social care.
Comment	Indicates how a basic additional intervention vs. usual treatment pragmatic chance-based design can be carried out in a ‘real-life’ social care context.

Randomisation and chance-based designs in social care research

The pragmatic CBD does not require the participant and those delivering the intervention to be blind to what is being received – for many complex interventions, this is indeed impossible. More realistically, only those assessing outcomes do not know which intervention each participant is receiving. In this sense, they may be described as ‘single-blind’. This is desirable to avoid the assessors’ expectations inadvertently biasing the results obtained. In some studies, there are challenges to the assessor remaining blind, in that the participant may spontaneously talk about the intervention, despite exhortations not to let slip such details. One approach to addressing this potential source of bias is to ask assessors to rate their confidence as to whether the participant had or had not received the intervention after each assessment (for example, see Box 1), so that the extent of bias can be estimated and controlled for in the statistical analyses of the study.

4 CHANCE-BASED DESIGNS – ARE THEY ETHICAL?

4.1 Winners and losers

CBDs are sometimes perceived as being intrinsically unfair. For example, if the study involves comparison of a new treatment approach in comparison with usual care, some practitioners feel uncomfortable, interpreting this design as **withholding** something of potential value from perhaps half the participants. This would also be the case in a study with a placebo comparison group where enrolment in the study necessarily involves the possibility that an inactive substance will be received. To determine the intervention approach for a person who is in need of social care on the toss of a coin, a random approach to decision-making, appears difficult to countenance.

An alternative view is that pragmatic CBDs, comparing a new treatment with usual care, are offering something **additional** to a proportion of the participants – in a sense they are the lottery winners, but there is no penalty, no loss for those without the winning ticket. In the study described in Box 1, it is interesting to note that those randomised to ‘standard treatment’, in this case input from a full multi-disciplinary team, were indeed offered the new treatment approach after 6 months. This resulted from practitioner concerns, but had the consequence that the study was not able to ascertain the longer-term effects of the new service beyond the 6 month follow-up period. In many studies, an attempt is made to ensure that all participants receive **something**, whether it is an opportunity to receive the intervention at a later date, or a package of information, seen as representing the best of what might be offered as usual care. These strategies can be useful for pragmatic CBDs, as it has been suggested that participants in the control group may experience ‘resentful demoralisation’ if not receiving the new treatment. If the research team ensures that those in the control arm receive what is considered ‘best practice’, this may of course in fact exceed ‘usual care’ for many.

The ‘winning the lottery’ analogy does not give the full picture, however. While many participants may enrol in a study simply to be in with a chance of receiving a new treatment, and which perhaps has limited availability outside the study (this is certainly the case in drug trials in conditions such as cancer), others may have other motivations – interest, curiosity, altruism, for example. In the REMCARE study (Woods *et al.* 2012) those involved in the development of the large reminiscence groups for people with dementia and their carers being evaluated assumed that those not randomised to attend would be dissatisfied. In fact, some couples expressed relief that they would not be expected to attend, and a significant proportion (11%) of those invited to attend did not take up the offer.

Perhaps the key ethical issue here relates to how balanced our knowledge of the alternative approaches is. This is sometimes described as ‘equipoise’ (Freedman 1987). If there exists convincing evidence that the new intervention is superior, then there is little justification for designing a study to evaluate it. If experts in the field are already

Randomisation and chance-based designs in social care research

convinced of the benefits of the new approach, then a readiness is required to consider not continuing with the approach if the results are not positive. Once an approach is embedded in 'good practice', this can be remarkably difficult. This can be a challenging area, as inevitably new approaches are developed by enthusiasts, but to undertake an evaluation opens up the potential for negative as well as positive findings.

A welcome trend in health care research has been the involvement of patients and public in research studies from conception to completion (for example, see Box 2). The input of service users in the design of the research can help to ensure proper attention is given to ensuring the study is fair to control participants. Sometimes, the design may go beyond the offer of treatment to all those in the control group wishing to receive it (as was the case in Box 1) to a design where the effects of the intervention on the control group after a delay are considered alongside the effects on the initial intervention group. This 'waiting list' control design is taken further in some studies where all participants receive the intervention, but for each participant or group, the intervention commences at a point randomly determined resulting in baseline measurements of varying duration, which can then be compared with the outcomes following the intervention being implemented (a stepped-wedge or interrupted time-series design, described in more detail below).

Box 2: Improving the health of adults with learning disabilities – evaluation of a hand-held health record (Turk *et al.* 2010)

Background	People with learning disabilities are known to have high health needs but have little involvement in their own health care or health promotion activities. There is clear evidence of health inequality. Hand-held health records have been used in a number of other fields, empowering the person to have more knowledge of their own health needs and the services they receive or require. In this study, the hand-held record was a 50-page Personal Health Profile (PHP) that has been developed and piloted widely
Research objective	Evaluation of whether adults with learning disabilities with a hand-held health record attended their GP more often, had greater knowledge of their health and had improved communication and satisfaction with their GP in comparison with usual treatment.
Design	Participants were 201 adults with a learning disability identified from joint social services/health service registers. A cluster design was used, with general practice as the unit allocated to intervention or control condition, because the use of the PHP requires training for the GP and primary care team in its use.
Ethical issues	Some participants were unable to give informed consent, and carers were consulted regarding their involvement. The study pre-dates the Mental Capacity Act (2005), which provides a framework for the involvement of people who are unable to give informed consent in research studies in England and Wales.

Randomisation and chance-based designs in social care research

Box 2 (continued): Improving the health of adults with learning disabilities – evaluation of a hand-held health record (Turk *et al.* 2010)

Outcomes	Number of GP consultations over a 12-month period. Questionnaires for people with learning disabilities and carers regarding health and satisfaction with GP.
Strengths and weaknesses	<p>There was extensive involvement of people with learning disabilities and their carers throughout the research process:</p> <ul style="list-style-type: none"> • Advocacy group representatives were involved throughout, from the submission of the grant application. • Four people with learning disabilities were employed and trained to interview all the research participants with a learning disability. They were supported by four professionals from the Learning Disability team who recorded the interview responses and assisted in coding them. • Four carers were employed and trained as researchers and interviewed the carers participating in the study. <p>The study used mixed methods with qualitative interviews adding greatly to, and serving to clarify, the quantitative results.</p> <p>The GP practices were stratified on several characteristics to increase comparability of the practices allocated to the two conditions.</p> <p>There was a high drop-out rate of people with a learning disability, with nearly a third unavailable at follow-up – one reason was finding the evaluation too burdensome. There was also a high turn-over of carers. This may have reduced the power of the study to detect an effect.</p>
Generalizability	This design could be used where a service innovation is being introduced at a team level in social care. The involvement of users of services and carers is an excellent model.
Comment	Good example of a pragmatic cluster chance based design and the involvement of users of services, their carers and representative organisations throughout the research process.

A more radical (and still controversial) approach is to adopt what has been described in health-care literature as a ‘patient-preference design’ (Torgerson and Torgerson 2008). This approach recognises that while some participants may be unconcerned regarding the outcome of randomisation, others may have a strong preference, such that they may not join the study at all, or may withdraw if not randomised to their preferred arm. This may lead to an unrepresentative sample participating. In a typical participant preference CBD, those without a strong preference are randomised as normal whereas those with a strong preference are allocated to their preferred condition. Box 3 provides an example of such a study comparing two counselling approaches in primary care with a usual care control condition. Here, sufficient people were prepared to be randomised to provide an adequate randomly selected ‘usual care’ control group. Most people who had a

Randomisation and chance-based designs in social care research

Box 3: Non-directive counselling and CBT for people with depression in primary care
(Ward *et al.* 2000)

Background	Evidence for the effectiveness of counselling for people with depression in primary care has been limited, and specialist therapies such as CBT (cognitive behavioural therapy) have been mainly evaluated in specialist settings. ‘Talking therapies’ are popular among people with common mental health difficulties, but recruitment to studies has been difficult, perhaps because potential participants do not want to risk not receiving the intervention. Participants’ pre-existing expectations regarding the therapies may also act as a source of bias.
Research objective	This study aimed to evaluate the effectiveness of non-directive counselling in comparison with CBT and usual care from the General Practitioner (GP).
Design	Participants were 464 people with depression or mixed depression and anxiety referred by their GP. Participants could either agree to be allocated to one of the three conditions (counselling; CBT; usual GP care) by chance, or they could choose to be allocated to their treatment of choice. During the course of the study, it emerged that many participants had a strong preference to receive a talking therapy, but both CBT and counselling were equally acceptable. In these cases, the allocation to either CBT or counselling was made by chance.
Ethical issues	The opportunity for participants to be allocated to their preferred treatment meant that potential allocation resentment was avoided.
Outcomes	Main outcome measure was score on a validated self-report depression questionnaire assessed at 4 months and 12 months.
Strengths and weaknesses	Able to check that those refusing allocation by chance do not differ in important ways or in outcome from those who only agree to take part if allocated to the treatment of their choice. No independent verification of outcomes – reliant on self-report. Checks made to ensure treatment adhered to CBT and counselling approaches respectively.
Generalizability	Applicable to areas of social care where participants may have strong preferences for a particular approach.
Comment	Provides evidence that the potential bias, by including only those agreeing to chance allocation, may not be great in some contexts.

preference wished to receive one of the two active therapeutic approaches, although a number simply had a preference for an active intervention and were prepared to be randomised to either of the two approaches. It might be thought that deviating to this extent from the principle of random allocation would threaten the internal validity of the study (i.e. it would give different results than if all participants were randomly allocated). In fact, a systematic review by King *et al.* (2005) identified 32 studies making use of

Randomisation and chance-based designs in social care research

participant preference designs and concluded that differences in outcome between chance-based groups and preference-based groups are relatively small, especially in larger studies. It is interesting to note that the characteristics of those who agreed and those who refused to be randomised were not as markedly different as anticipated, suggesting that including only those agreeing to randomisation would not necessarily produce an unrepresentative sample.

4.2 One size does not fit all

Given the complexity of social care provision, the notion of being able to standardise a service intervention, as if it were a pill providing a treatment in a standard, defined manner, is one that leads to discomfort in many practitioners. Slavishly following a treatment manual, specifying what to say, how to respond, appears to risk becoming mechanistic and simplistic. Each individual, each situation is different, and so each service intervention must in turn be different to meet the needs of that specific set of circumstances. The challenge then becomes how to describe and capture the essence of the approach taken so that other services, other practitioners, may learn and benefit from the experience and expertise used to avoid the necessity of re-inventing the wheel in every case, and to provide assurance that we are offering the most cost-effective service. This may require analysing and understanding better the common features in the individualised interventions that are offered.

For example, an expert social worker may develop highly effective, highly individualised care packages for service users that lead to good outcomes for many. However, if service commissioners wished to roll out this service in other areas, other parts of the country, employing other social workers, there is no guarantee the results would be the same. Each social worker would have a distinct skill set and differing experience and expertise. It is not enough to say 'social work is the answer' without being able to say more regarding the details of the approach taken and its fit with the situations encountered. Good practice can only be communicated and disseminated if it can first be described.

Developing an intervention approach requires a process that seeks to identify the essence of the approach taken, specifying the key elements, aiming to standardise these, while retaining those aspects which allow tailoring to individual needs.

How can an approach be standardised and individually tailored at the same time? One approach that has proved useful is to have a standardised process of assessment and action planning, but for the action plan to be completely driven by the needs of the individual concerned. This is illustrated by the study described in Box 4 where care home residents with dementia were assessed using a comprehensive needs assessment tool (the Camberwell Assessment of Need for the Elderly – CANE) and an action plan developed based on the individual resident's profile of needs. The intervention then consisted of additional input seeking to find ways of meeting the residents' unmet needs in an attempt to improve quality of life. The potential range of interventions was then great, depending on the creativity of those involved.

Randomisation and chance-based designs in social care research

Meeting the needs of people with dementia living in care homes (Orrell *et al.* 2007)

Background	The majority of older people living in care homes have dementia. Considering the range of needs – physical, social, psychological – they typically have significant unmet needs and reduced quality of life. Systematically reviewing the person’s needs, identifying unmet needs and implementing an individualised action plan based on addressing unmet needs may improve quality of life.
Research objective	The aim of this study was to evaluate whether additional specialist input to care homes related to meeting individual needs of residents with dementia would lead to a reduction in unmet need and an improvement in quality of life.
Design	Participants were 238 people with dementia living in 24 care homes. All were assessed using a comprehensive needs assessment instrument (the Camberwell Assessment of Needs for the Elderly – CANE) and an action plan developed. In half the homes, allocated by chance, an additional one-hour-a-week liaison input was provided to assist in implementing the action plan.
Ethical issues	Family members consulted regarding relative’s participation. Urgent issues identified by assessment fed back immediately to the care home even if allocated to control condition. Action plans for control care home residents fed back after follow-up assessments completed.
Outcomes	Unmet needs assessed using the CANE and Quality of Life (self-report and completed by staff) after a 20 week intervention period.
Strengths and weaknesses	<p>Interventions highly individualised, based on comprehensive individual assessment.</p> <p>Follow-up assessments carried out by assessor blind to allocation of care home.</p> <p>The intervention period may have been too brief – meeting some needs required systemic change (e.g. activities; awaiting referral to hospital consultant).</p> <p>Many participants unable to complete quality of life measure, meaning sample size too small for this measure.</p> <p>Staff turnover would make it difficult for same staff member to complete staff ratings of quality of life at baseline and follow-up.</p> <p>Carrying out the individualised assessment with staff in control homes may have alerted staff to unmet needs, leading to reduction in unmet needs in both control and intervention homes, reducing apparent effect of intervention.</p>
Generalizability	This cluster CBD could be used in a variety of social care contexts where the intervention involves staff training or introduction of a new approach.
Comment	Basing interventions on an individual assessment allows a standardised approach that is still person-centred.

Randomisation and chance-based designs in social care research

4.3 Outcomes that are meaningful

In any evaluation study, there is a need to ensure that we are asking the right questions. It is always tempting to measure constructs that are simple and easy to measure and shy away from more difficult and complex aspects, even though they may be of more direct relevance to people's lives and experiences. This is another area where the input of those with experience of using services is invaluable, advising on the most relevant and appropriate outcome measures to be used. For an overview of outcome measurement for adults using social care services and support see Netten (2011).

For example, in a study with people living with mental health difficulties, social engagement may be a more appropriate outcome to evaluate than the severity of symptoms in terms of the person's quality of life. Increasingly, quality of life and well-being measures are being used that reflect the person's own experience and perception rather than relying on the judgements of others. Evaluating such complex constructs is not easy, of course, and there is a need for further development and testing of measures in many cases before proceeding with a full evaluation. A good example is the development of quality of life measures that can be completed by people with mild-to-moderate dementia. A scale of this kind was the primary outcome measure in the REMCARE study of joint reminiscence groups for people with dementia and their carers mentioned previously (Woods *et al.* 2012). Before such scales were available, outcomes for people with dementia were often evaluated simply in respect of whether they remained at home or were admitted to an institution or in terms of scores on tests of memory and orientation. Neither of these approaches captured the experience of the person continuing to live with dementia. For an overview of quality of life measures in social care research see Bowling (2014).

With the development of interventions tailored to individual needs, there has also been attention paid to developing and using participant-centred outcome measures that focus on individual goals and concerns. For example, in an evaluation of individualised cognitive rehabilitation (typically delivered by an occupational therapist) with people suffering early-stage dementia, each participant is asked to identify goals he/she would like to achieve that are personally important and relevant (Clare *et al.* 2013). The success of the intervention is then evaluated according to the extent to which the participant rates the goal as achieved and their satisfaction with this.

Randomisation and chance-based designs in social care research

5 CHANCE-BASED DESIGNS – PRACTICALLY IMPOSSIBLE?

It is certainly the case that most CBDs are large undertakings with sample sizes typically numbered not in tens, but in hundreds or even thousands (for example, see Box 5). When reviews are undertaken of a number of CBDs in a particular field, small sample sizes are often the subject of critical comment. It is important to note that it is not the absolute sample size that is of concern, but rather whether the sample size utilised is large enough to detect the amount of change in outcomes that is considered worthwhile and relevant. This is described as the ‘effect size’. If an intervention has a large effect, so that most people receiving it show a clear response, an obvious improvement on the outcome measure being used, then the sample size needed for a rigorous CBD may actually be quite small. Unfortunately, most psycho-social interventions have small-to-medium effect sizes, and these necessitate the larger sample sizes referred to. The issue here is that if a study is ‘under-powered’, i.e. has a sample size insufficient to detect the anticipated difference, then there is a risk that an intervention may be wrongly viewed as being ineffective. This could lead to potentially helpful approaches not being pursued.

Box 5: Telecare for people with social care needs (study protocol) (Bower *et al.* 2011)

Background	There is great interest in the potential of technology to assist people with health and social care needs to remain as independent as possible. Telecare is defined as ‘the remote, automatic and passive monitoring of changes in an individual’s condition or lifestyle’ (e.g. movement sensors; fall sensors) while telehealth is ‘the remote exchange of data between a patient and health care professional’ (e.g. blood glucose monitoring). Little is yet known about their actual effects.
Research objective	The aim of this study is to evaluate the effects of telecare and telehealth on two groups of people: adults with social care needs and adults with long-term health conditions compared with usual treatment.
Design	Primary care practices will participate in a cluster CBD. Practices will be allocated by chance to have either access to telehealth (for patients with diabetes, heart failure and Chronic Obstructive Pulmonary Disease) or access to telecare (for adults with social care needs). Adults with social care needs in practices having telehealth will act as ‘usual treatment’ for telecare, and patients with long-term health conditions in practices having telecare will act as telehealth controls. It was planned to include up to 6,000 participants, whose service use would be monitored using routinely collected data and up to 900 participants who would be asked to complete measures of quality of life etc.
Ethical issues	Avoids allocation at the individual level, but individuals retain the right to participate or not (including data sharing aspect). After 12 months, control participants are assessed for the appropriate technology.

Randomisation and chance-based designs in social care research

Box 5 (continued): Telecare for people with social care needs (study protocol)
(Bower *et al.* 2011)

Outcomes	Service utilisation; quality of life.
Strengths and weaknesses	<p>Pragmatic trial in real-life context, but all practices involved are within demonstration centres, where whole system changes are being attempted, so generalisation to other centres will need to be tested.</p> <p>Each GP practice having access to an intervention for some of their patients may assist in retaining their cooperation.</p> <p>Wide variety of technology (and, importantly monitoring and back-up systems) being used – not standardised. If ineffective, it may be argued that more up-to-date technology would give a different result.</p> <p>Use of routine data allows much larger sample, but there may be concerns regarding its accuracy. This can be tested by comparison with the data from the smaller group who will complete service use inventories.</p> <p>Long-term condition and social care needs groups might be expected to overlap, but authors suggest in practice one or other technology is received.</p> <p>An extensive qualitative evaluation is built into the protocol, looking at the experiences of service users, carers and health and social care professionals in relation to the introduction and use of telehealth and telecare.</p>
Generalizability	This design could be applied across a variety of social care contexts where a new service is delivered through a team, day centre or care home, although having two distinct interventions for two distinct client groups may be unusual.
Comment	A good example of an ambitious, mixed-methods evaluation of a service development that has attracted a great deal of public funding with little evidence for effectiveness to date.

Large sample sizes have become accepted as the norm in health care research, and there is the economic argument that it is more cost-effective to expend research funds on one large study that provides a conclusive answer to a research question than on several smaller studies that raise more questions than answers.

Recruitment to studies is often challenging, but knowledge is growing of what factors assist in achieving target sample sizes. Again, involvement of people from the population who form the focus of the research in designing the research protocol can be extremely helpful in reducing barriers to participation and identifying what might make participation more attractive. Involvement of ‘gatekeepers’ of services is also helpful. Professionals who have direct contact with potential participants are best placed to introduce the project and respond to any initial queries and concerns. Such professionals need to be fully aware of the potential benefits and risks of the research so that they do not inadvertently become over-protective and act as a barrier to participation.

Inevitably, only a proportion of those who are eligible to take part agree to do so. It is

Randomisation and chance-based designs in social care research

important to establish that those who do agree to participate are representative of the population in question, but this can be difficult, as data on those who refuse to participate may not be available or cannot be accessed by the research team. It is feasible to think that there may be systematic biases in those who choose to participate in terms of, for example, age, frailty, relationship quality with family, personality type etc – all of which may be relevant to the interventions being evaluated.

Not all chance-based designs can randomise individual participants. Randomisation may have to occur at a higher level, known as a cluster; for example, when the intervention creates a new service, or teaches care professionals skills for participants in the intervention group they cannot 'forget' for control participants (Torgerson and Torgerson 2008). Cluster designs have been widely used in research involving people in care homes where the unit of randomisation is the care home. For example, in the study described in Box 4 (Orrell *et al.* 2007), 24 care homes took part and 12 of these were randomly selected to receive the additional intervention – in this case, a nurse visiting to assist with implementation of individual care plans for residents with dementia. Clusters might also be social work teams covering different geographical areas, or primary care practices (see, for example, Boxes 2 and 5), or day-care centres, for example. In these contexts, it may indeed be practically impossible for some individuals to receive the intervention and others not. Although assessments are often carried out with individuals within the cluster, there may alternatively, or in addition, be service level measures, e.g. number of service users attending a day-care centre remaining at home after 12 months or proportion of care plans judged to meet a quality standard. Typically, the total sample size required for a cluster chance-based design will be larger than for a study not involving clustering; and because the number of clusters will typically be relatively small, care is needed to avoid bias so that there are not systematic differences between clusters allocated to the different conditions. For example, if the size of care home was thought to be a potential influence, the randomisation process might be set up to ensure that in each condition there are large, medium and small homes. In the study described in Box 2 (Turk *et al.* 2010), primary care practices were stratified according to a number of factors, including whether or not they were a single-handed practice and whether there were more or less than 10 adults with a learning disability registered with them. Even when randomisation is carried out on an individual basis, there are circumstances where there may be a clustering effect, which must be allowed-for in analysing the outcomes. This can occur if the intervention is group-based or several participants receive input from the same worker, for example. In these examples, there may be systematic differences between groups or workers which complicate the comparison with the control condition.

A development of cluster chance-based designs that could be of use in social care research is the 'stepped wedge' or 'interrupted time-series' design (Brown and Lilford 2006). Here a number of clusters are involved, each receiving the intervention in turn, with the remainder acting as controls until all have received the intervention. In effect, clusters act as 'waiting-list' controls until they receive the intervention. The order of the clusters

Randomisation and chance-based designs in social care research

Box 6: Challenging behaviour and people with dementia in care homes (study protocol)
(Zwijnsen *et al.* 2011)

Background	Challenging behaviour in dementia is a complex matter, leading to distress for residents, family members and staff. Its causes are many and varied, and careful assessment and analysis is required in order to prevent and respond to distress. Challenging behaviour is associated with reduced quality of life and increased risk of being prescribed anti-psychotic medication, which may have adverse effects in people with dementia.
Research objective	This study aims to evaluate a new multi-disciplinary programme for responding to challenging behaviour, based on Dutch national guidelines, consisting of four steps: detection, analysis, treatment and evaluation in comparison with usual treatment.
Design	A stepped wedge CBD will be used involving 14 special care units, each including around 20 people with dementia in 14 care homes. The participating units are divided into five groups (4 of 3 and 1 of 2 units) by chance. 6 rounds of measurement will occur over a period of 20 months. The first will be a baseline for all participating units. After each round of measurement, except the last one, a new group of units will start the intervention. The order in which the groups commence the intervention is determined by chance.
Ethical issues	The stepped wedge design means that all units will receive the intervention during the project. Residents' legal representatives have the opportunity to withdraw their resident from the study.
Outcomes	Extent of behavioural problems is the primary outcome. Other outcomes include quality of life, prescription rate of antipsychotics, use of physical restraints and workload and job satisfaction of nursing staff. Cost-effectiveness is also evaluated.
Strengths and weaknesses	The sample size has been calculated in relation to behavioural problems. Inevitably, over a 20 month period in this population, a proportion of residents will die or move to other facilities. The plan is for these to be replaced by newly admitted residents – this makes analyses of change over time more complex, especially if there are differences in attrition between groups of units. In view of the severity of dementia, all the assessments are completed by proxies; in this case, staff of the homes that will not be blinded to whether the programme has started in the home, although the research team collecting the data from the staff will be blinded. A key issue will be the extent to which the unit takes up and follows the intervention programme. A process analysis is planned including qualitative interviews on barriers and facilitators to implementation with key care home staff. The steps to implementation (training/monitoring/support) are not detailed in the protocol.

Randomisation and chance-based designs in social care research

Box 6 (continued): Challenging behaviour and people with dementia in care homes (study protocol) (Zwijssen *et al.* 2011)

Generalizability	This design is applicable to a range of social care settings where a new service or intervention is being introduced and shows how a stepped wedge design can be implemented with groups of facilities rather than one unit at a time.
Comment	The actual programme to be implemented here would need adaptation in a UK context where care homes typically do not have such ready access to a multi-disciplinary team.

receiving the intervention is determined randomly, of course, to avoid bias. The advantage of this design is that all the participating clusters receive the intervention and avoids any resentment from those not selected. Also, if the intervention requires resource input (e.g. a specialist trainer), then this input can be spread over time rather than being required in each centre at the same time.

There are some disadvantages of stepped wedge designs (Kotz *et al.* 2012, Dreischulte *et al.* 2013). Typically, assessment measures need to be completed for all clusters at the point where each cluster commences the intervention. This could mean that if there were eight clusters, ten rounds of assessment would be needed (including a baseline before and a follow-up assessment after all clusters have received the intervention). This contrasts with a typical three assessment rounds in other chance-based designs and has implications for the extent and type of assessment measures used. The study described in Box 6 using a stepped wedge design in 14 care homes reduces this effect by grouping the homes into 5 groups for purposes of initiating the intervention.

There is also the likelihood that the intervention will develop with experience over time and that, for example, the eighth cluster may receive an intervention markedly different from the first. In order to be clear as to what has been evaluated, ensuring consistency over time of the intervention – using a manual and checking adherence – would be a requirement. The effects of time itself may be especially important in some contexts, e.g. where participants have progressively worsening social or health conditions where the timing of the intervention may be critical. Changes in the external context may also have an influence – a change of policy or procedures at national or local levels, or even severe weather. In standard chance-based designs, these sources of influence are assumed to be equivalent between intervention and control conditions, but the situation is more complex in the stepped-wedge design. However, this design does provide a good evaluation model for evaluating the extent and nature of effects of a new service model that is being rolled out into practice.

6 CHANCE-BASED DESIGNS – THE GOLD STANDARD?

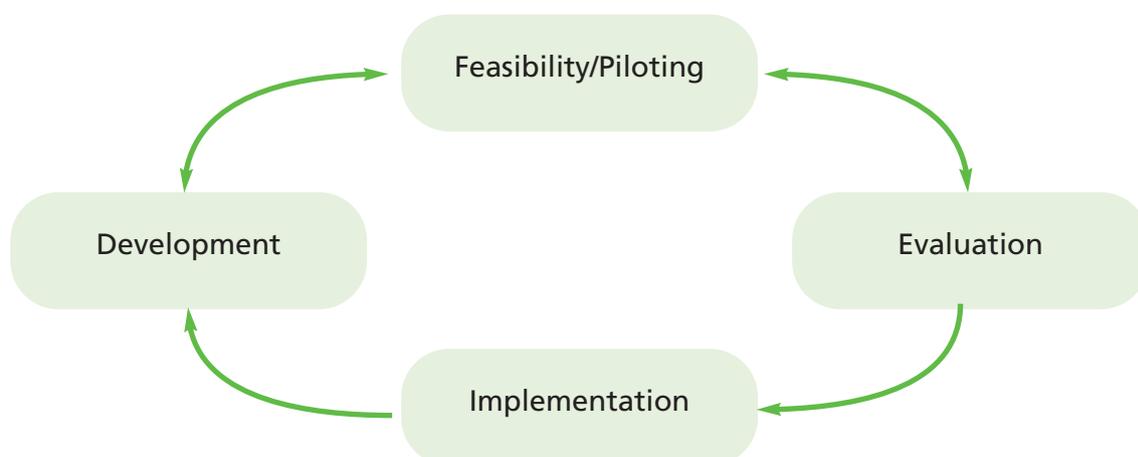
It is probably fair to say that chance-based designs have to a large extent been viewed as the gold standard in medical and health care research. Other methodologies were seen as precursors to establishing the definitive chance-based study that would form the cornerstone of the evidence-base in the specific area under examination. Increasingly, there is a recognition that the ultimate aim of research is the implementation of research into practice and that there can be a massive gap between the evidence-base and practice. This has led to consideration of implementation as an integral part of the process of development of new interventions and not simply something to be considered after the definitive chance-based study.

This changing emphasis is reflected in the widely accepted pathway, developed by the Medical Research Council (2008), for the evaluation of complex interventions. The Framework states that

Developing, piloting, evaluating, reporting and implementing a complex intervention can be a lengthy process. All of the stages are important, and too strong a focus on the main evaluation, to the neglect of adequate development and piloting work, or proper consideration of the practical issues of implementation, will result in weaker interventions, that are harder to evaluate, less likely to be implemented and less likely to be worth implementing. (p.4).

A key point to note is that the pathway from development of a new intervention to implementation is not seen as a linear progression. This is different from the situation for a new drug, which may be seen as proceeding in stages from laboratory testing, through initial trials in humans, to full-scale effectiveness trials to licensed use in medical practice. The framework is presented as a cycle with interactions between stages (see Figure 1), but it is acknowledged that this is an over-simplification.

Figure 1: MRC Framework for the evaluation of complex interventions (adapted from MRC 2008)



Randomisation and chance-based designs in social care research

Within this framework, there is scope for a variety of methodologies to be used, including qualitative methods. Integrating quantitative and qualitative methods can be challenging, but guidance is available on approaches to working effectively with mixed methods (e.g. O’Cathain *et al.* 2010). For a discussion of the framework in the context of end of life care and social care see Evans *et al.* (2013).

The four components identified in the Framework are as follows:

6.1 Development

In considering a new intervention, or adapting previous approaches, a great deal of preparatory work needs to be undertaken. This should include identifying and systematically reviewing the evidence base to establish what is already known.

Understanding and developing the theory underlying the approach is also important: Why would this approach make a difference? What difference exactly would it make? For example, a creative activity in a day-care centre might be enjoyed at the time by the older people attending, but by what process could this enjoyment influence the person’s level of function in his/her own home? Perhaps the enjoyment in a social setting may lead to sustained mood improvements, which in turn leads to improved function at home. If so, this has implications for how the intervention might be evaluated with measures of mood and functional level being appropriate. If the theory is that the enjoyment is ‘in the moment’, then evaluation has to be targeted at the immediate effects during the activity session.

It is important to consider implementation at this early stage. Could this intervention be practically implemented? How does it fit with the policy context? Who would deliver the intervention in practice? Who are the target population? Would it be seen as unaffordable even if it does turn out to be effective? It is possible to conceive of interventions that could well be effective, but might be socially unacceptable, e.g. offering financial payments for compliance with a social work programme.

6.2 Feasibility/piloting

Assessing the feasibility of the intervention and proposed evaluation procedures is essential. Trying out the intervention on a small scale will allow it to be refined so that a clear and full description of it can be drawn up, typically leading to a tried and tested intervention ‘manual’. Feedback from participants, perhaps using focus groups or interviews, analysed qualitatively, should inform this process. In some situations, it is necessary to develop, pilot and validate new evaluation measures, perhaps where consideration of the theoretical basis of the intervention and the expected outcomes has led to the view that existing measures would not be fit for purpose.

Preliminary work in feasibility and pilot studies also allows issues relating to recruitment of study sites and participants to be considered. What are the barriers to both participation in and remaining in a study in this area? How difficult is it to identify

Randomisation and chance-based designs in social care research

potential participants, and what proportion would be likely to agree to participate? As mentioned previously, these are areas where the input of users of services, their carers and representative organisations is especially valuable.

If the intention is to proceed to a full-scale CBD, this preliminary work should include work to estimate the likely effect size on the outcome measures of primary interest. This can be achieved through a pilot study where the aim is simply to gain an indication of how large a difference appears to be made by the intervention in this context. This then allows a calculation to be made of the sample size needed, which can be adjusted to allow for the likelihood that the full-scale CBD will be carried out under less controlled and favourable conditions than the pilot study.

6.3 Evaluation

Here, a CBD of sufficient size to detect important effects of the intervention(s) under study is recommended, but as we have seen, there are a variety of approaches to consider, including cluster and stepped wedge designs and studies where participants who have a strong preference may select which condition they receive.

Within the context of a CBD, it is recommended that work goes on in parallel evaluating the processes of the intervention and its effects within the study. This might include examining the extent to which the intervention is actually carried out as planned and the extent to which participants engage with the intervention procedures. Qualitative interviews and focus groups with those carrying out the intervention and the participants may help to clarify how the intervention is perceived, what has been helpful, and what have been unforeseen effects, if any.

Wherever possible, an evaluation of cost-effectiveness should be carried out as part of the CBD. This will allow estimates to be made of what the additional costs may be of the effects obtained, or perhaps of the cost-savings that can be achieved through the intervention. Even if the intervention is not effective, in terms of leading to change on the primary outcome measures, it may be cost-effective if it is associated with lower costs without detriment to those in receipt of it.

6.4 Implementation

Dissemination of research findings has to be much more active than simply publishing an academic paper. Identifying the key policy makers, those who have influence in guiding service commissioning and provision and providing the important messages in readily digestible form, are good starting places. Making intervention manuals and other materials readily available, providing workshops and training may all help. It is often important to study and understand the processes of behaviour change in those who could deliver the intervention. What are the barriers for them in changing practice or in initiating a new model of service delivery? What drivers and levers for change could be incorporated? If implementation has been fully considered when the intervention was

Randomisation and chance-based designs in social care research

under development, then there is more chance it will be a good fit with current frameworks and policies. For example, if the intervention has been designed to be implemented by an existing group of staff, this may be more readily assimilated than one where a whole new type of staff would need to be employed.

When an intervention is being widely implemented, further study is needed to ascertain the extent to which it is actually being implemented in practice and whether it is having the desired effects in the 'real-life' setting in the long-term, as well as whether it is associated with any adverse effects which were not apparent in the evaluation studies. Further development of the intervention may be required where research in the implementation phase suggests there are barriers to embedding it within practice as it stands.

Implementation research requires consideration of the social care systems in which the intervention might be located. Staff training provides a good example here. Often seen as a panacea for many difficulties and short-comings in services, evidence for the effectiveness of training in many domains is limited, in part related to staff not being able to put into practice new knowledge and skills within the system in which they work.

An example, of implementation research is provided by Streater *et al.* (2013) who report a study where staff in social care settings for people with dementia receive training to deliver a group intervention. They examine whether the proportion of those receiving training that actually run groups can be increased through additional support. In addition, they plan to evaluate whether the effects of the group intervention are comparable in the context of this implementation to those obtained in CBDs where there was the full involvement of a research team.

7 CHANCE-BASE DESIGNS – ONLY USEFUL FOR EFFECTIVENESS?

As we have seen from the above discussion, the emphasis in pragmatic CBDs tends to be on the effectiveness of the whole package rather than unpicking individual components. However, within the MRC Framework, there are a number of points where processes and mechanisms may be addressed. These include the development phase where theoretical models of change are examined and explored; the feasibility and pilot work where different components may be explored and compared on a small scale; the evaluation phase where processes and mechanisms may be explored using mixed methods; and the implementation phase where the interaction of the intervention with the ‘real-world’ becomes the focus of attention and some of the systemic influences on the effectiveness of the intervention become apparent.

The MRC Framework suggests that having established effectiveness, work may be needed to increase its efficiency. Identifying factors in the CBD that are related to a positive response to the intervention may be helpful in highlighting potential ways of targeting the intervention more effectively. Sub-group analyses can be informative; for example, a support programme for family care-givers might be highly effective for spouse care-givers but much less so for adult daughters and sons providing care to a parent; or a programme to engage people with a learning disability in day-time activity might be much more effective for those with a moderate degree of impairment but less so for those who have either milder or more severe impairment. Such findings can inform further development of the intervention and the underlying theory. Inter-relationships between outcomes in different domains may also be informative; for example, in a study of a cognitive stimulation activity group for people with dementia in care homes and day centres (Woods *et al.*, 2006), improved quality of life was related to improved cognitive function, suggesting the mechanism of change was not simply arising from the social benefits of being part of a group.

Another approach, having established effectiveness, is to vary a key element of the intervention across arms of a further CBD study. For example, Williams *et al.* (2013) report what they describe as a ‘dismantling’ CBD study. In this study, the primary intervention – Mindfulness-based Cognitive Therapy (MBCT) – was compared with both usual treatment and an active intervention, which removed one component (meditation) from the usual delivery of MBCT. The participants had all experienced recurrent episodes of depression but were well at the time of entry to the study. The results of this study showed no differences between the three participant groups in terms of experiencing a further episode of depression, but analyses of those who responded best to the intervention indicated that those participants with the most adverse childhood experiences benefitted most from MBCT (including the meditation aspect). These results will inform the future development and implementation of this widely-implemented approach to maintaining mental health in adults vulnerable to depression.

8 CONCLUSIONS AND RECOMMENDATIONS

We have suggested that the resistance to making full use of CBDs in social care research arises from a mythology regarding CBDs based on their use in double-blind randomised controlled trials of drugs. Over several decades, researchers in health care and public health, who operate in a context akin to that of social care, have addressed many of the valid concerns that would indeed be raised if the drug trial methodology was simply parachuted into social care research. Pragmatic CBDs have been placed within a framework which recognises the complex nature of the interventions developed in both health and social care, with well-developed methodologies to evaluate services and their effects, rather than simply focusing on interventions at the individual level. Although the worship of CBDs as the 'gold standard' continues to an extent in the medical domain, there is a growing realisation that they are just one piece of a complex inter-locking jigsaw puzzle, and that the complete picture is only achieved when an approach is routinely and widely implemented into practice.

There is no doubt that CBDs can be challenging: they require a team approach if they are to be set up and established appropriately. Input from statisticians, methodologists, health economists, qualitative researchers, outcome measure specialists, data managers, study managers and quality assurance specialists is needed in addition to the domain-specific expertise from social care researchers and from users of services and carers (Torgerson and Torgerson 2008). However, properly conducted and reported, they may well be cost-effective in terms of the value of knowledge and information gained.

CBDs have been repeatedly subject to ethical scrutiny over many years, and – using designs appropriate to the situation – they have been judged to be acceptable by lay people, including representatives of those using services. Indeed, the argument can be made that using a design for an evaluation not involving the allocation or the timing of the intervention being determined by a chance-driven process could in some circumstances be unethical, as it runs the risk of a biased conclusion to the evaluation. Not to use the unique but naturalistic rigour of a pragmatic CBD to inform the best use of scarce resources could itself be seen as unethical.

However, our emphasis is on CBDs being used appropriately and when they are the best and most appropriate design for the particular research question. There are certainly some situations where it would be unwise to use a CBD:

- Where an intervention is not yet fully developed or has not been extensively piloted. Premature evaluations of under-developed interventions confuse rather than build an evidence-base.
- Where the desired outcome is to prevent a relatively rare (though important) adverse event, where the sample size would have to be very large indeed to detect a difference between groups.

Randomisation and chance-based designs in social care research

- Where the intervention is already widely implemented, as may be the case with a policy directive or initiative. A stepped-wedge design may be possible if all sites are not implementing simultaneously, but otherwise a naturalistic observational evaluation may be needed.
- Where the intervention is not in a situation of equipoise, i.e. if the intervention will be implemented regardless of the results of any evaluation (perhaps because opinion leaders are already convinced of its benefits, or simply perhaps because it is the 'right thing' to do), then it is arguable whether a CBD is justified.

Finally, in view of the relative infrequency of use of CBDs in social care research, it may be helpful to highlight some of the key features from Zwarestein *et al.* (2008) for readers to watch out for, as consumers of research, in evaluating research reports involving CBDs in social care.

- Are the outcome measures appropriate? Do they evaluate the key areas that will make a difference to the lives of users of services? Have the measures been validated so that they provide an accurate picture of what they purport to measure?
- Are the results fully reported? When there are numerous outcome measures, statistical significance may be found on one or two measures by chance. It is good practice for the researchers to publish the protocol for the study before the study data collection is complete so readers can check that the study is carried out and analysed as planned. Several of the example studies provided in the Boxes in this paper are study protocols rather than reports of completed studies.
- Have attempts been made to avoid potential sources of bias? This might include having independent blind assessors and ensuring that the intervention is delivered as planned. Has the chance allocation been made independently so it could not be influenced, e.g. by a researcher, to retain a participant who has a clear preference not to be in the control group in the study?
- Has the sample size been properly justified in terms of a credible expectation of the difference the intervention would need to make for it to be seen as useful?
- Are the participants fully described so that it is clear what population have been included? Are the numbers of participants leaving the study, and the reasons for withdrawal, clearly stated? Are the results analysed in terms of the study allocation to different conditions regardless of whether the participant took up the intervention or not?

Open and transparent reporting of CBDs is essential to maintaining and raising their quality, and ensuring that the results and recommendations do justice to the contribution of all those taking part.

9 REFERENCES

- Arceneaux K, Nickerson DW (2009) Who is mobilized to vote? A re-analysis of 11 field experiments, *American Journal of Political Science*, 53, 1–16.
- Bower P, Cartwright M, Hirani SP, Barlow J, Hendy J, Knapp M, Henderson C, Rogers A, Sanders C, Bardsley M, Steventon A, Fitzpatrick R, Doll H, Newman S (2011) A comprehensive evaluation of the impact of telemonitoring in patients with long-term conditions and social care needs: protocol for the whole systems demonstrator cluster randomised trial, *BMC Health Services Research*, 11, 184.
- Bowling A (2014) *Quality of Life: Measures and Meanings in Social Care Research*, SSCR Methods Review 16, NIHR School for Social Care Research, London.
- Brown CA, Lilford RJ (2006) The stepped wedge trial design: a systematic review, *BMC Medical Research Methodology*, 6, 54.
- Card D, Della Vigna S, Malmendier U (2011) The role of theory in field experiments, *Journal of Economic Perspectives*, 25, 39–62.
- Clare L, Bayer A, Burns A, Corbett A, Jones R, Knapp M, Kopelman M, Kudlicka A, Leroi I, Oyeboode J, Pool J, Woods B, Whitaker R (2013) Goal-oriented cognitive rehabilitation in early-stage dementia: study protocol for a multi-centre single-blind randomised controlled trial (GREAT), *Trials*, 14, 152.
- Cook TD (2001) Science phobia: why education researchers reject randomised experiments, *Education Next*, Fall, 63–68.
- Dreischulte T, Grant A, Donna, P, Guthrie B (2013) Pro's and con's of the stepped wedge design in cluster randomised trials of quality improvement interventions: two current examples, *Trials*, 14, Supplement 1, O87.
- Evans CJ, Stone KA, Manthorpe J, Higginson IJ (2013) *MRC Guidance on Developing and Evaluating Complex Interventions: Application to Research on Palliative and End of Life Care*, SSCR Methods Review 15, NIHR School for Social Care Research, London.
- Freedman B (1987) Equipoise and the ethics of clinical research, *New England Journal of Medicine*, 317, 141–145.
- Hassiotis A, Robotham D, Canagasabay A, Romeo R, Langridge D, Blizzard R, Murad S, King M (2009) Randomized, single-blind, controlled trial of a specialist behavior therapy team for challenging behavior in adults with intellectual disabilities, *American Journal of Psychiatry*, 166, 1278–1285.
- Hotopf M (2002) The pragmatic randomised controlled trial, *Advances in Psychiatric Treatment*, 8, 326–333.
- Illsley R (1982) Research and the NHS, *Health Bulletin (Edinburgh)*, 40, 1, 54–57.

Randomisation and chance-based designs in social care research

King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, Sibbald B, Lai R (2005) Impact of participant and physician intervention preferences on randomized trials: a systematic review, *Journal of American Medical Association*, 293, 9, 1089–1099.

Kotz D, Spigt M, Arts ICW, Crutzen R, Viechtbauer W (2012) Use of the stepped wedge design cannot be recommended: a critical appraisal and comparison with the classic cluster randomized controlled trial design, *Journal of Clinical Epidemiology*, 65, 1249–1252.

Medical Research Council (2008) *Developing and Evaluating Complex Interventions: New Guidance*, Medical Research Council, London.

Netten A (2011) *Overview of Outcome Measurement for Adults Using Social Care Services and Support*, *SSCR Methods Review 6*, NIHR School for Social Care Research, London.

O’Cathain A, Murphy E, Nicholl J (2010) Three techniques for integrating data in mixed methods studies, *British Medical Journal*, 341, c4587.

Orrell M, Hancock GA, Hoe J, Woods B, Livingston G, Challis D (2007) A cluster randomised trial to reduce the unmet needs of people with dementia living in residential care, *International Journal of Geriatric Psychiatry*, 22, 11, 1127–1134.

Schwartz D, Lellouch J (1967) Explanatory and pragmatic attitudes in clinical trials, *Journal of Chronic Diseases*, 20, 637–644.

Streater A, Spector A, Aguirre E, Hoe J, Hoare Z, Woods R, Russell I, Orrell M (2012) Maintenance cognitive stimulation therapy (CST) in practice: study protocol for a randomized controlled trial, *Trials*, 13, 91.

Torgerson DJ, Torgerson CJ (2008) *Designing Randomised Trials in Health, Education and the Social Sciences: An Introduction*, Palgrave, Basingstoke.

Turk V, Burchell S, Burrha S, Corney R, Elliott S, Kerry S, Molloy C, Painter K (2010) An evaluation of the implementation of hand held health records with adults with learning disabilities: a cluster randomized controlled trial, *Journal of Applied Research in Intellectual Disabilities*, 23, 100–111.

Ward E, King M, Lloyd M, Bower P, Sibbald B, Farrelly S, Gabbay M, TARRIER N, Addington-Hall J (2000) Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression I: clinical effectiveness, *British Medical Journal*, 321, 1383–1388.

Williams JMG, Crane C, Barnhofer T, Brennan K, Duggan DS, Fennell MJV, Hackmann A, Krusche A, Muse K, Von Rohr IR, Shah D, Crane RS, Eames C, Jones M, Radford S, Silverton S, Sun Y, Weatherley-Jones E, Whitaker CJ, Russell D, Russell IT (2013) Mindfulness-based cognitive therapy for preventing relapse in recurrent depression: a randomized dismantling trial, *Journal of Consulting & Clinical Psychology*, 82, 2, 275–86.

Randomisation and chance-based designs in social care research

Woods RT, Bruce E, Edwards RT, Elvish R, Hoare Z, Hounsome B, Keady J, Moniz-Cook E, Orgeta V, Orrell M, Rees J, Russell IT (2012) REMCARE: reminiscence groups for people with dementia and their family caregivers – effectiveness and cost-effectiveness pragmatic multicentre randomised trial, *Health Technology Assessment*, 16, 48, 1–116.

Woods B, Thorgrimsen L, Spector A, Royan L, Orrell M (2006) Improved quality of life and cognitive stimulation therapy in dementia, *Aging & Mental Health*, 10, 219–226.

Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D for the CONSORT and Pragmatic Trials in Healthcare (Practihc) groups (2008) Improving the reporting of pragmatic trials: an extension of the CONSORT statement, *British Medical Journal*, 337, a2390.

Zwijssen SA, Smalbrugge M, Zuidema SU, Koopmans R, Bosmans JE, van Tulder MW, Eefsting JA, Gerritsen DL, Pot AM (2011) Grip on challenging behaviour: a multidisciplinary care programme for managing behavioural problems in nursing home residents with dementia. study protocol, *BMC Health Services Research*, 11, 41.

NIHR School for Social Care Research
London School of Economics and Political Science
Houghton Street
London
WC2A 2AE

020 7955 6238
www.sscr.nihr.ac.uk
sscr@lse.ac.uk



Phase I (2009-2014) of the NIHR School for Social Care Research (SSCR) was a partnership between the London School of Economics and Political Science, King's College London and the Universities of Kent, Manchester and York. Phase II (2014-2019) of SSCR is a partnership between the London School of Economics and Political Science and the Universities of Bristol, Kent, Manchester and York, and is part of the National Institute for Health Research (NIHR) www.nihr.ac.uk/.