WSD Action Network

Evaluating telecare and telehealth interventions

WSDAN briefing paper

Anna Davies and Stanton Newman
University College London

The King's Fund

This paper, the second in a series of WSDAN briefing papers, provides guidance about how to carry out good-quality evaluations of telecare and telehealth interventions. We identify different types of evaluations, and briefly review the evidence from systematic reviews of telecare and telehealth to illustrate key issues to be addressed in future evaluations.

We describe the main elements of an evaluation protocol. These include drawing up the questions, selecting an appropriate evaluation design, selecting participants, defining outcome measures, and deciding on analytic methods. We discuss some of the key practical considerations for developing an evaluation, including piloting, staffing, ethics, funding, project and data management, and progress monitoring. The 'References' and 'Resources' sections at the end of the paper provide further information on how to carry out a telecare or telehealth evaluation.

The WSDAN Steering Group members are:

Mike Clark	WSDAN Co-project Lead
Tim Ellis	WSD Programme Director, Department of Health
Richard Giordano	Senior Lecturer, Management and Leadership in Health and Social Care, University of Southampton
Nick Goodwin	WSDAN Co-project Lead
Stephen Johnson (Chair)	Head of Long Term Conditions, Department of Health
Stanton Newman	Dean, School of Community and Health Sciences, City University
Jeremy Porteus	National Programme Lead - Networks (Housing and Telecare), Department of Health
Susan Royer	WSDAN Co-ordinator

If you have any questions about this briefing paper, please email us at: wsdnetwork@kingsfund.org.uk.

Alternatively, you can write to us at:
WSD Action Network
c/o The King's Fund
11-13 Cavendish Square
London
W1G 0AN

Contents

1 Introduction

We begin by defining the terms 'telecare', 'telehealth' and 'evaluation', and go on to describe the types of evaluations that can be conducted in relation to their ability to provide good-quality evidence for telecare and telehealth. We briefly discuss the nature of the evidence for the effectiveness and cost-effectiveness of telecare and telehealth, with the purpose of demonstrating the quality of the evidence that is currently available, and issues that should be addressed by future evaluations.

We then offer guidance about how to plan and implement an evaluation locally, focusing on specific practical issues, such as obtaining ethical approval and data management. However, we do not review the evidence for the effectiveness of telecare and telehealth. Those with an interest in the evidence base for telecare and telehealth should refer to the Whole Systems Demonstrator Action Network (WSDAN) evidence database.

This briefing paper focuses on how to plan and carry out a good-quality evaluation to assess the effectiveness of telecare and telehealth interventions. While the evaluation methods described lend themselves to a cost-effectiveness evaluation, we do not cover the specific methods and analytic techniques needed for this kind of evaluation. Those with an interest in cost-effectiveness evaluations should refer to Professor Martin Knapp's presentation at a recent WSDAN Roadshow, on 'Telecare and telehealth: economic issues', for further information (Knapp 2010) .

Defining telecare and telehealth

Within this briefing, we use the following definitions of telecare and telehealth:

Telecare has been defined by the Department of Health as a service that uses '*a combination of alarms, sensors and other equipment to help people live independently. This is done by monitoring activity changes over time and will raise a call for help in emergency situations, such as a fall, fire or a flood* ' (Department of Health 2009).

Telecare therefore combines monitoring equipment with a monitoring service, and is most frequently used in the home. A telecare user may activate their own alarm if they use a pendant. For those users with passive monitoring equipment, their behaviour patterns are monitored, and changes outside of their normal behavioural parameters are flagged for action (e.g. not getting out of bed at the usual time, exiting the house at night). This monitoring is intended to support people and enable them to continue living in their own home, independently or with the assistance of carers, for as long as possible.

Telehealth has been defined by the Department of Health as a service that '*uses equipment to monitor people's health in their own home… [monitoring] vital signs such as blood pressure, blood oxygen levels or weight* ' (Department of Health 2009).

People use the equipment at home, and in some cases outside the home, to measure the vital signs that would normally be measured by a health care professional, helping to reduce frequent visits to the GP surgery. It is also anticipated that telehealth may reduce the number of unplanned hospital admissions by helping to identify changes

in people's health status before any problems become serious enough to warrant emergency intervention. The data is transmitted automatically via broadband or a dial-up telephone line to a monitoring centre or health care professional. Readings that indicate changes outside the normal parameters, which may indicate deterioration in health, are then flagged for action.

2 Evaluations of telecare and telehealth

What is an evaluation?

An evaluation is the systematic collection of data or information, which is then analysed to inform future decision-making. By carrying out evaluations of interventions and services, we can develop an evidence base to inform decisions about best practice. An evaluation of telecare or telehealth can provide evidence about the effectiveness and cost-effectiveness of the various interventions in managing or improving the well-being of people with care and support needs and/or long-term health conditions.

People with care and support needs are those who, '*by reason of age, illness, disability or any other circumstances are in need of care and attention which is not otherwise available to them' (National Assistance Act* 1948). Long-term conditions have been defined by the Department of Health as '*those conditions that cannot, at present, be cured, but can be controlled by medication and other therapies. They include diabetes, asthma, and chronic obstructive pulmonary disease [COPD]*' (Department of Health 2007). They include non-communicable diseases (e.g. cardiovascular disease, diabetes, COPD) and communicable diseases (e.g. HIV and AIDS), impairments in structure (e.g. joint disorders, blindness), and certain mental health disorders.

Good-quality evaluations provide the best available scientific evidence to inform policy decisions about treatments or changes in the way that care is delivered for people with care and support needs and/or long-term conditions. In turn, this can improve the quality of care provided, increase the efficiency of care delivery, and improve the allocation of resources.

What is a good-quality evaluation in telecare and telehealth?

A good-quality evaluation is one that answers the evaluation question(s) with a high degree of certainty that what is found is an accurate reflection of the effect of the telecare or telehealth intervention. In other words, it should attempt to measure the 'true' effect of telecare or telehealth.

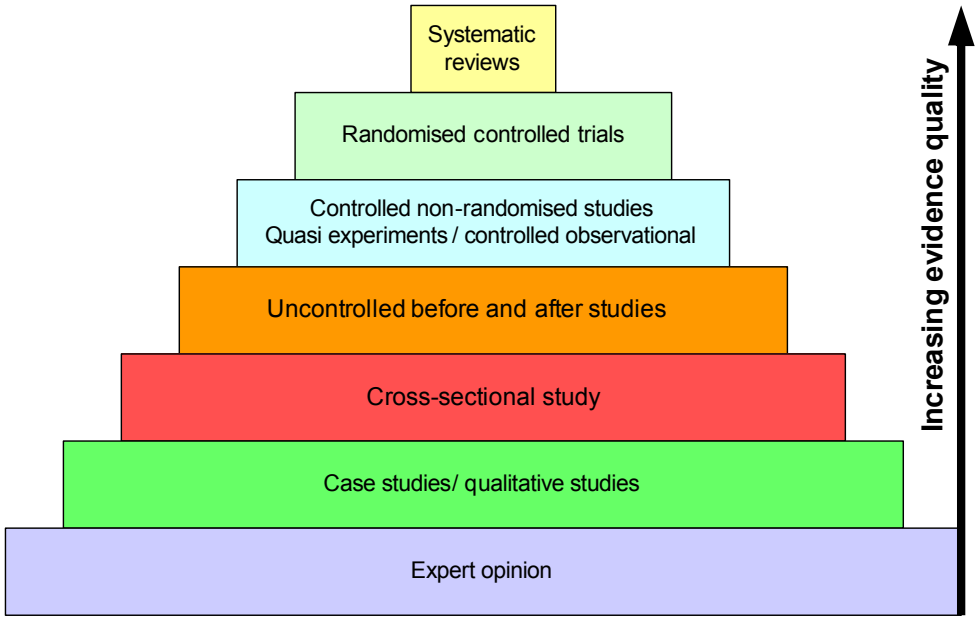
Evaluation designs vary in the extent to which they offer good-quality evidence to inform decision-making about telecare and telehealth. A hierarchy of evidence, presented as a pyramid (*see* Figure 1), has been developed to indicate the frequency with which different evaluation designs are used, and the extent to which they offer good-quality evidence to support evidence-based decision-making in trials of effectiveness. Different hierarchies are available for different research questions (e.g. Evans 2003; see also table 2)). At the bottom of the pyramid is 'expert opinion', frequently used for evidence-based decision-making, but offering the poorest quality evidence. At the top are systematic reviews, less frequently carried out because they are time-consuming and labour-intensive, but offering the best quality evidence. Systematic reviews offer the best quality evidence only if robust methods have been used, and if the evidence selected for inclusion in the review is also of good quality. We discuss the resources needed to carry out each type of evaluation in 'Funding' (*see* page 27).

Different types of evaluation

Systematic reviews

Systematic reviews are generally regarded as the highest quality evidence to inform evidence-based practice. This is because they attempt to draw conclusions about the phenomenon under investigation through a systematic inspection of the accumulated literature about it. Systematic reviews usually answer a specific research question or questions.

Figure 1 Hierarchy of evidence (adapted from Newman and Davies 2009)



To answer evaluation questions, systematic, replicable methods are used to:

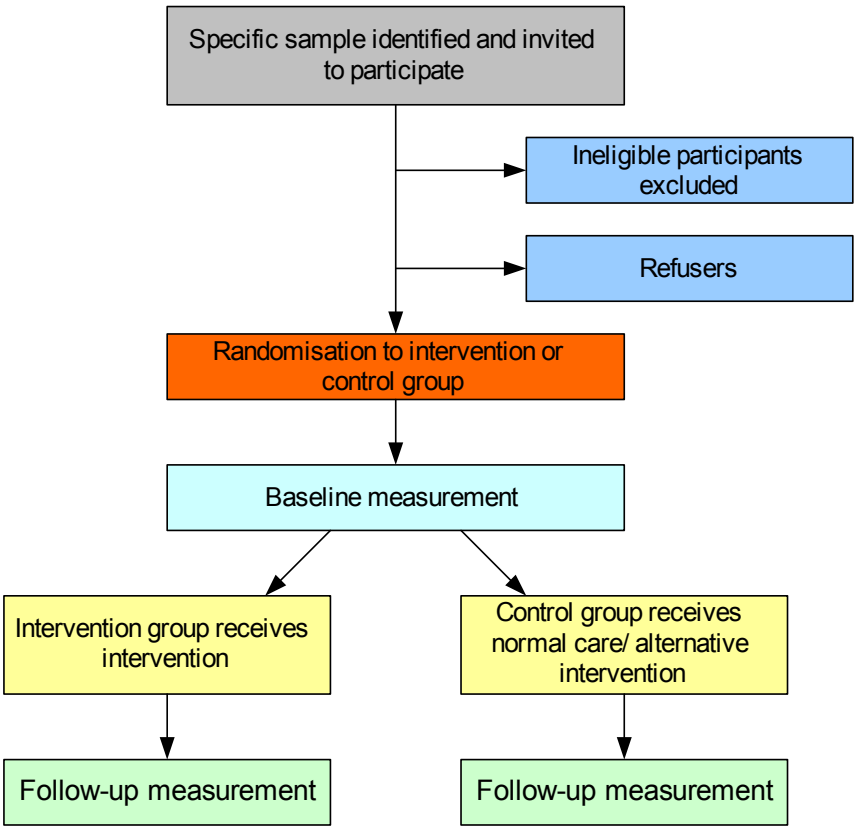
- **identify publications** relating to the topic – an exhaustive search is carried out to identify all relevant publications. In many cases, 'grey literature' is also identified. Grey literature comprises reports that are not published in peer-reviewed journals.
- **select the studies** for inclusion in the review. Specific criteria are set to decide which publications are included. This can be based on factors such as the type of intervention used, the evaluation design, or the quality of the available papers.
- **synthesise and interpret the findings** using a replicable method so that the findings across different evaluations can be drawn together and interpreted in an unbiased manner. Synthesis can be narrative, or can use meta-analytic procedures, to combine statistical data.

Systematic reviews can be located through electronic searching of databases of reviews (e.g. The Cochrane Collaboration), databases of journal articles (e.g. PubMed), or handsearching of topic-relevant journals. WSDAN has also developed a database of peer-reviewed and non-peer reviewed literature that is specific to the telehealth and telecare field (WSDAN evidence database). Where there are no systematic reviews available, it may be necessary to conduct one's own systematic literature review to inform evidence-based policy and practice.

Randomised controlled trials

A randomised controlled trial is the best method of determining whether a cause and effect relationship exists between telecare/telehealth and the outcomes under examination (Roland and Torgerson 1998; *see* Figure 2).

Figure 2 Diagram representing basic randomised controlled trial



RCTs rule out the possibility that the effect of the tested intervention can be explained by other factors associated with both the evaluation group and outcome (Sibbald 1998). The key features of an RCT are as follows.

- **Random allocation to intervention/control groups:** this ensures that there are no systematic differences (both known or unknown) between people in the intervention and control groups.
- **All groups receive exactly the same treatment, except the intervention under investigation** (in this case, the telecare or telehealth intervention): outcomes can then be attributed to the effect of the intervention.
- **Some randomised controlled trials are blind or double blind:** concealing the group allocation for the participant and the investigator is one way of minimising bias. However, this may not be possible in a pragmatic trial that examines a choice between methods of care such as in the case of telecare and telehealth (Roland and Torgerson 1998).

The size of the difference in outcomes between intervention and control group participants is examined.

Evaluations of telecare and telehealth lend themselves to a pragmatic randomised controlled trial design. In contrast with explanatory trials, in which the efficacy of interventions is tested in tightly controlled conditions, pragmatic RCTs assess how effective an intervention is in the context of routine practice (Roland and Torgerson 1998). This gives the trial findings external validity; they are more likely to indicate the likely intervention effects when applied in real life. Some differences between explanatory and pragmatic RCTs are presented in Table 1.

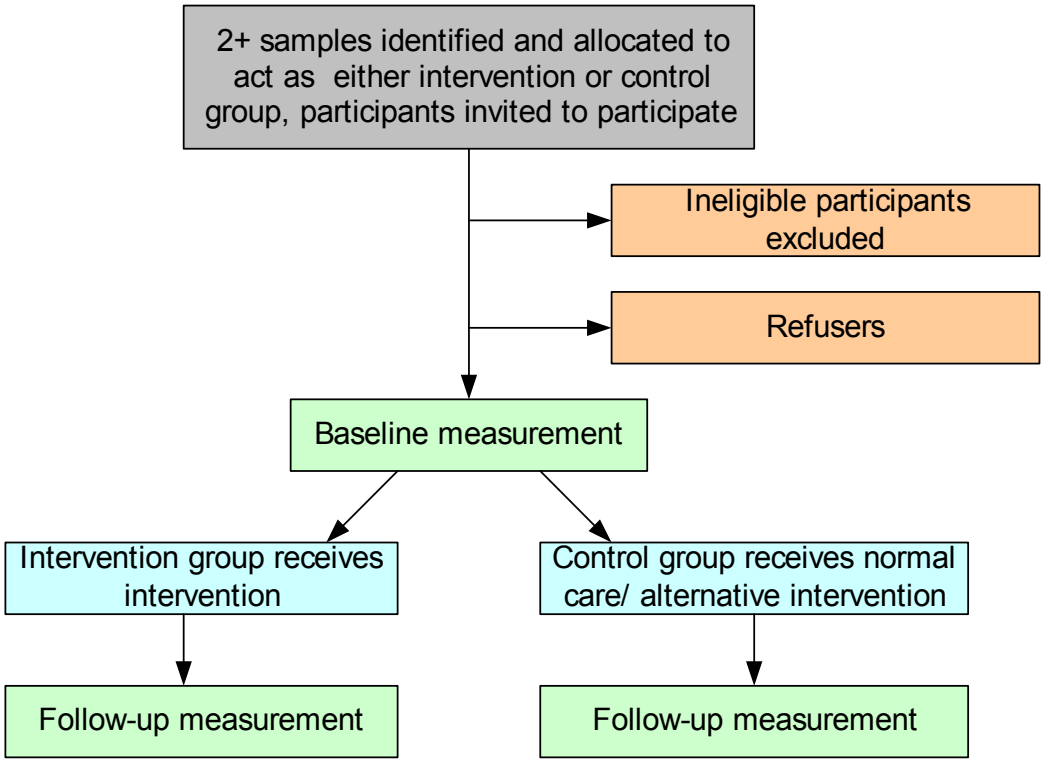
Table 1 Differences between explanatory and pragmatic RCTs (adapted from Zwarenstein *et al* 2008)

	Explanatory RCTs	**Pragmatic RCTs**
Question	Tests efficacy – does the intervention work?	Tests effectiveness – does the intervention work when used in normal practice?
Setting	Well-resourced, ideal setting	Normal practice setting
	Highly selected, participants may be part of a pool who are willing to participate in research. Participants may be excluded if they have other co-morbidities than the condition of interest	Little selection beyond condition of interest
	Clearly defined, strictly enforced and closely monitored	Intervention may vary between individuals, e.g. in relation to telecare/telehealth equipment selected on the basis of need
	Often short-term outcomes, surrogate indicators, or process measures	Outcomes relevant to participants, funders and practitioners
	Limited relevance – design may not match decision- making needs of patients/ practitioners for whom intervention will be implemented	High relevance – trial is designed to meet practitioner/ patient decision-making needs

Controlled non-randomised studies (quasi-experiments/ controlled observational)

Quasi-experiments are natural experiments. Participants are not randomly allocated to the intervention they receive, but are regarded as an intervention or control group participant by virtue of membership to a naturally occurring group, such as living in a particular geographic area, or being treated by a particular health care practitioner. For this reason, quasi-experiments are mostly carried out when randomisation is unethical or impractical. Quasi-experiments cannot give the same assurance as RCTs about cause and effect, as greater effectiveness of an intervention may be attributable to systematic differences between the groups, rather than as a result of the intervention being tested (Shadish *et al* 2002).

Figure 3 Quasi-experimental design



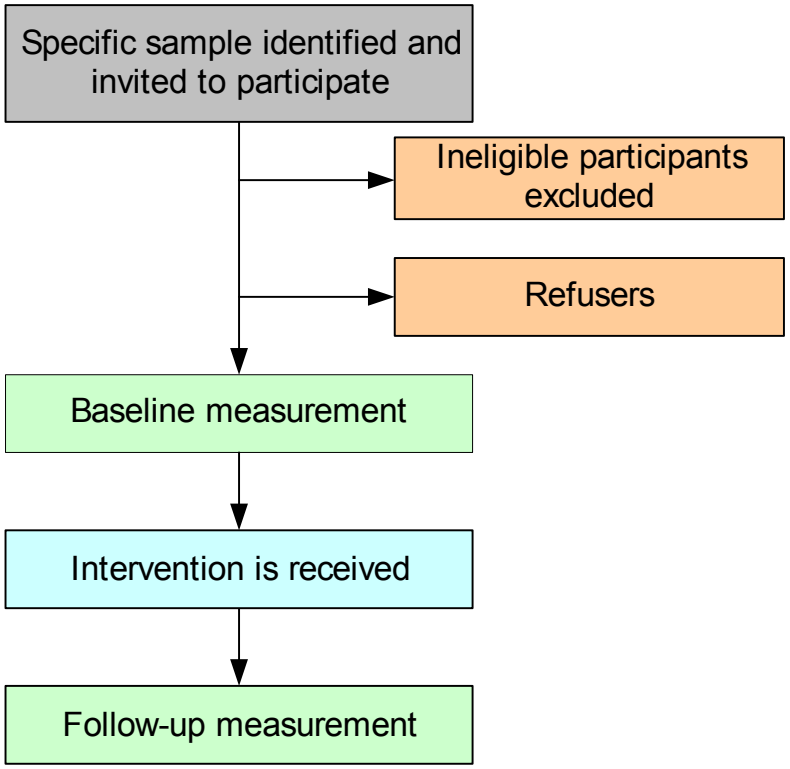
Uncontrolled 'before and after' evaluations

This design is frequently used to evaluate telecare and telehealth interventions. It was used by the US Veterans Health Administration (VHA), for example, in its recent evaluation of the effects of its Care Coordination/Home Telehealth programme on health care resource utilisation (Darkins *et al* 2008). In a 'before and after' evaluation, the outcomes of interest are measured before the telecare or telehealth technology is introduced, and afterwards, to determine whether the intervention results in a change in the outcomes.

The external validity of before and after studies can be limited. This is because known or unknown variables other than receipt of the intervention may be

responsible for any change in outcomes. In addition, they can only demonstrate change resulting from the intervention, and cannot tell us whether the intervention exerts a larger effect or greater benefit than current normal care.

Figure 4 'Before and after' design



Cross-sectional studies

A cross-sectional study is one of the simplest evaluation designs. Data is gathered from participants at a single point in time. A cross-sectional study cannot answer questions of effectiveness, and is more likely to be used to describe participants' characteristics and perceptions of telecare and telehealth.

Qualitative evaluations/ case studies

Qualitative evaluations incorporate various designs and specific techniques, including interviews, observations and focus groups. For each of these methods there is a range of approaches that can be used, including phenomenology, action research and case studies (Bowling 1997). Qualitative evaluations typically result in data from a relatively small number of participants in comparison with quantitative evaluations. Participants are selected using various methods, including convenience, purposive, snowballing and theoretical sampling (*see* Appendix 1 for a glossary of sampling terms).

Qualitative evaluations of telecare or telehealth might be used to answer questions about participants' experiences of receiving the intervention, and staff experiences of providing it. Qualitative evaluations that include case studies are not appropriate for answering questions about the effectiveness of the intervention. They typically provide insights into people's beliefs and experiences relating to the intervention,

rather than attempting to objectively measure the outcomes.

The specialist skills required of a qualitative evaluator are equal to those required of a quantitative evaluator. Therefore, staff who are not experienced in conducting qualitative work will require rigorous training to be able to carry out a robust evaluation.

3 Improving the evidence base

A number of systematic reviews suggest that telecare and telehealth can be effective in improving the quality of life and health outcomes of the people who use the devices or services (e.g. Barlow *et al* 2007), and may also result in cost savings (e.g. Polisena *et al* 2009). However, there is a consensus amongst reviews that further, more robust evaluations are needed to address the limitations of the current evidence for the effectiveness of telecare and telehealth interventions. Previous systematic reviews have suggested a number of ways to develop a more robust evidence base, and these are detailed below.

The need for better quality evidence

Several systematic reviews indicate that there have been few evaluations of sufficient quality to determine effectiveness (e.g. Polisena *et al* 2009; Polisena *et al* 2010; Glueckauf and Ketterson 2004). For example, in a systematic review of telecare for people with physical disability, learning disability or cognitive impairment, there were no studies that fulfilled the quality assessment criteria set by the reviewers (Martin *et al* 2008). Quality assessment involves determining whether an evaluation fulfils a number of criteria. These relate to use of an appropriate design, participant selection, methods of intervention, sample size calculation, and appropriate analytic methods to accurately answer the evaluation questions.

As stated earlier, randomised controlled trials are generally considered the best method of assessing cost-effectiveness and effectiveness, and previous systematic reviews have frequently used an RCT design as a selection criterion for inclusion. However, numerous reviews have found that a small proportion of RCTs have been conducted to answer questions of effectiveness and cost-effectiveness (e.g. Maric *et al* 2009).

The need for more evidence on key aspects of telecare and telehealth

Economic outcomes

Reviews have identified that only a small proportion of evaluations of telecare and telehealth interventions address the question of economic outcomes, and that the quality of evaluations needs to be improved (Davalos *et al* 2009). For example, a recent evaluation of telehealth for COPD found that only 2 out of 23 identified studies examined the costs associated with the intervention (Jaana *et al* 2009). Another review identified 130 evaluations of telehealth for heart failure, but only 22 provided sufficient data to allow an appraisal of economic outcomes (Bensink *et al* 2006).

Another issue is that many economic analyses evaluate cost only, and do not examine the non-cost related outcomes following the introduction of telecare or telehealth (*see* Bergmo 2009; Davies and Newman 2010). Without examining the benefits beyond cost-savings – for example, improved quality of life – it is not possible to determine whether telecare and telehealth offer the same or better outcomes compared with the current level (and cost) of care. Better quality economic evaluations are needed (Polisena *et al* 2009), and specific indicators of economic outcomes have been suggested (Rojas and Gagnon 2008). It is also important that more economic analyses are carried out in the UK; extrapolation

of findings from American or Western European evaluations to the UK health care system may be inadvisable, given the differences in how the health and social care systems are structured, funded and managed.

Which devices or services are most effective?

There is a need to evaluate the effectiveness of specific devices. While some evidence reviews report findings for specific devices, using clear criteria to select studies on the basis of the devices used, many reviews report findings about an array of different devices or services (e.g. Klersy *et al* 2009; Maric *et al* 2009). For example, a recent review of telehealth interventions in heart failure considered the evidence for videoconferencing, telephone consultation and doctor–patient communication via the internet (Neubeck *et al* 2009). Making generalisations across these different modalities makes it difficult to identify which interventions are the most effective or cost-effective. Since telehealth interventions often comprise a number of components – for example, a user may have a blood pressure monitor, scales and a pulse oxometer – reviews have called for more work to be done to identify which is the most effective component (Dang *et al* 2009; Schmidt *et al* 2010).

Which groups of people are most likely to benefit?

More evaluations are needed to attempt to identify which groups of people are most likely to benefit from telecare or telehealth interventions. Groups may include people with specific health conditions, or health status, or in relation to socio-demographic characteristics. There is also some evidence to suggest that telecare and telehealth are more effective in managing some health conditions than others. For example, a review of a small number of studies found positive effects of telehealth monitoring for participants with heart failure or psychiatric conditions, but not for diabetes (DelliFrane and Dansky 2008). These findings have been supported by another review (Paré *et al* 2007).

What are the long-term effects?

The Whole System Demonstrators Established Users Evaluation aims to evaluate the effectiveness of telecare and telehealth for people in Newham and Kent who have been using equipment for more than a year. The results of this evaluation will be available in 2011. However, most evaluations have not considered the effectiveness of the intervention for a period longer than 12 months after installation (*see* Klersy 2009; Wainwright and Wootton 2003). A recent review of telehealth for people with heart failure has identified that most interventions are evaluated at either 6 or 12 months (Klersy 2009). Many telecare or telehealth users are likely to continue using the equipment for periods longer than 12 months, and patterns of telecare and telehealth equipment use may change over time – both of which will affect health and social care outcomes. Therefore it is essential to develop a better understanding of whether an intervention's effectiveness and cost-effectiveness persists beyond the first year.

4 Planning a good-quality evaluation

Rigorous planning is essential for a good-quality evaluation, so adequate time should be given to the planning and preparatory stages. You should allow at least six months to develop a comprehensive protocol, engage relevant stakeholders, and complete the necessary NHS ethics and research and development procedures (*see* 'Obtaining ethical approval for an evaluation' on page 24).

Developing an evaluation protocol

Setting clear questions

The key to designing a high-quality evaluation of telecare and telehealth is to clearly set out the question(s) to be answered. The research questions guide every aspect of the evaluation, including the design, inclusion criteria for selecting participants, methods, outcome measures and analytic techniques. Questions should arise from issues identified from previous studies, or from local practice. The questions should have a clear purpose – for example, to evaluate the impact of different technologies for people with the same long-term conditions, evaluate current practice, compare different methods of practice, or examine effects of practice in different populations.

An evaluation may address one or more questions. The WSD trial aims to answer a number of questions relating to the effectiveness and cost-effectiveness of telecare and telehealth, the processes of change, and the experiences of participants and providers (Department of Health 2010).

Questions might include the following:

- Do telecare and telehealth improve quality of life for people with long-term conditions and care and support needs?
- Does telehealth encourage people to change their behaviour, and engage in more self-care behaviours?
- Are telecare and telehealth more effective than current normal care in improving or maintaining people's quality of life?
- Do telecare and telehealth reduce the frequency of hospitalisation or length of stay for people with care and support needs or those with long-term conditions?
- Do telecare and telehealth delay or prevent admission to care homes?
- Are there specific client/ patient groups for whom telecare and telehealth are more or less beneficial?
- What are people's experiences of using telecare and telehealth and related services?
- Are telecare and telehealth more cost-effective than the care that is currently provided?
- What changes need to be made to provide an effective and cost-effective telecare and telehealth service?

Selecting an evaluation design

You should select an evaluation design that will directly address the research question(s). While randomised controlled trials (RCT) are generally regarded as the gold standard, certain research questions may be better addressed using other quantitative or qualitative designs. As Table 2 shows, an randomised controlled trial design is the best method for addressing questions of effectiveness and cost-effectiveness. Questions about acceptability and changes to the way services are provided can be addressed using both qualitative and quantitative methods.

Defining the intervention and control conditions

The evaluation protocol should clearly specify the intervention to be used. The uniformity of the intervention received by participants will depend on the evaluation questions and the nature of the design. Telecare and telehealth interventions are likely to be complex (*see* Craig *et al* 2008). This is because participants in an evaluation are likely to be using more than one device, and the devices will aim to improve not only the individual's well-being, but also health and service-use related outcomes (e.g. quality of life, blood glucose, hospitalisations). For this reason, they lend themselves to evaluation using a pragmatic RCT design. Furthermore, telecare and telehealth can be considered to be complex interventions because they will be required to operate in a complex social and health service system (Shiell *et al* 2008).

Control group

Where the intervention is compared with the effects of receiving current normal care, it is also unlikely that control group participants will receive completely uniform care. This is not a problem, as it will determine intervention effectiveness – its effect when applied in real life (Roland and Torgerson 1998). However, it is extremely important that the nature of normal care is carefully documented to allow generalisations to be made about the effects of the intervention in comparison with it.

Table 2 Evaluation designs and their ability to address evaluation questions

Example question	RCT	Controlled non-		Cross-
Effectiveness: Is telecare/telehealth more effective than current care?	✓✓	✓	✓	x
Cost-effectiveness: How cost-effective is telecare/telehealth?	✓✓	x	✓	x

Satisfaction: Are people satisfied with the service?	✓	x	x	✓✓
Acceptability: Will people want to use telecare/telehealth?	✓✓	✓	x	✓
Safety: Does telecare/telehealth do more good than harm?	✓✓	✓	✓	x
Process of service delivery: How well does it work?	x	x	✓	□□

✓✓ Best method/s ✓Good method/s, x- not appropriate to answer question

Selecting participants

Selecting the population and sample

When selecting participants for an evaluation, you should clearly identify the population of interest, and the method used to sample from it.

A **population** is all the people in the group of interest. A population can be as heterogeneous or as homogeneous as required. For example, it may be very narrowly defined to a specific sub-group (e.g. all people with type I diabetes who have an HbA1c higher than 6.5), or it may incorporate a wide range of conditions (e.g. all people using telecare, or all people in primary care who are using telehealth).

A **sample** is a group of individuals in the population who are used to represent the population.

Population selection should be determined by the aims, objectives and questions identified during the planning of the evaluation. The extent to which the population sampled is homogeneous will determine your ability to generalise outside of it. **Generalisability** is the extent to which evaluation findings, when tested in a population sample, can be applied to the wider population.

Homogeneous population with narrow eligibility criteria: Narrow eligibility criteria result in a sample from a population with very specific characteristics. Examples of narrow eligibility criteria might include people with type I diabetes with an HbA1c higher than 6.5, or men aged over 70 with COPD, or people with social care needs who have fallen at home at least twice in the past year. Using a homogeneous population with less variation in participant characteristics will allow you to identify reliable findings that are specific to the population under evaluation. However, evaluation findings cannot then be used to make generalisations about the intervention's likely effect for people with a different condition, or people who

have not fulfilled your eligibility criteria.

Heterogeneous population with broader eligibility criteria: Broad eligibility criteria result in a sample from a more heterogeneous population. An example might be inviting all patients registered with a GP in a primary care trust (PCT) area who have had a recent unplanned hospital admission relating to diagnosis of a long-term condition (such as diabetes, COPD or heart failure) to participate in an evaluation. In relation to telecare, it might be inviting people who are currently in receipt of night-sitting services, or people who have an informal carer.

The benefit of using broader eligibility criteria for telecare or telehealth evaluations is that it can provide information about the likely effectiveness for people who are frequent users of primary care services or social services. However, variation between the participants may mask an effect of the intervention on outcomes. To overcome this, you may need to use a sample stratification procedure to ensure that you recruit sufficient numbers of participants from specific sub-groups of interest (e.g. those with specific health conditions). This can allow further analysis to determine the effects in each group.

Identifying participants for your sample

Sampling methods

The most likely method for recruiting participants is a convenience sample (also known as purposive; *see* Appendix 1 for sampling terms and definitions). With convenience sampling, the evaluator recruits participants from a known, easily accessed population. For example, in a telecare or telehealth evaluation, participants may be recruited from those who have contacted the telecare service and have been referred to the care management or social work team for a needs assessment.

It may be possible to recruit potential participants for telecare or telehealth via their GPs. Participants can then be sampled from the population who have been identified as having a need that may be appropriately addressed with a telecare/telehealth service. A number of random sampling methods can be used including simple random sampling, systematic random sampling, stratified random sampling, or cluster sampling. Non-random sampling methods might include quota sampling.

Setting clear eligibility criteria

You should develop a list of criteria that potential participants must fulfil to be eligible to participate (inclusion criteria). You should also draw up a list of criteria that exclude someone from taking part (exclusion criteria). These lists should be used by staff referring potential participants to the evaluation (e.g. GPs, social care staff), and by the evaluation team members responsible for recruiting participants into the trial.

For example, to be recruited to the social care evaluation in the WSD trial, participants must be aged 18 or over, and currently be receiving seven hours or more home care per week, or 3.5 hours or more home care per week plus a meals service.

To be recruited to the heart failure telehealth evaluation, individuals must be aged 18 or over, have had a heart failure diagnosis confirmed by echocardiogram or by specialist assessment, and have had at least one unplanned health service use event in the past 12 months (such as an unplanned hospital admission, ambulance use or A&E visit). In addition, for participants recruited to the diabetes or COPD evaluation groups, clinical measures relating to their conditions are required to have been taken within a specific period of time prior to their entry into the evaluation. For example, a measure of HbA1c should be taken within six months of inclusion in the trial to act as a baseline clinical measure, against which changes over time can be assessed. Evidence of participants' fulfilment of these criteria should be carefully recorded, both by those referring the participant to the study, and by the evaluation team.

With regard to exclusion criteria, individuals who lack mental capacity to consent to participation were not included. Those who live in sheltered or warden-assisted accommodation, and those who do not have a telephone line, were also excluded from taking part.

Allocating participants to evaluation groups: randomisation in RCTs

The method of randomisation is likely to be chosen on a pragmatic basis, to reflect the routes through which participants are recruited to the trial, and the staff resources available for recruitment and participant management.

Cluster randomisation

Participants may be randomised to their group by cluster. Randomisation of individuals to clusters can be done for reasons of feasibility or to reduce administrative burden. Clusters may be defined in terms of geographic areas, social or health care providers, hospital clinics or GP practices.

Cluster randomisation has some drawbacks: statistical power is reduced because of the possibility that participants in a cluster may report similar outcomes due to the care they receive. For example, it is possible that receiving care from a specific GP may influence telehealth participant outcomes. Therefore, cluster-randomised evaluations require participation of a large enough number of clusters to accurately detect intervention effects. There is also increased likelihood that the effects of an intervention might be explained by systematic differences between those participants allocated to intervention and those allocated to control clusters. It should be noted that cluster-randomised evaluations require the use of specific analytic procedures.

Individual randomisation

Individual participants are randomised to the intervention or control group. Where successful, this carries the benefit of reducing the likelihood of confounding by extraneous variables. For example, it prevents the person who is referring a participant to the evaluation from allocating him or her to receive the telecare or telehealth intervention (rather than allocating them to the control group) because they think they are most likely to benefit from it. Therefore, individual randomisation may prevent systematic differences between those randomised to the intervention and control groups, other than the care that they receive. Thus, if randomisation has

been successful, the likelihood that a third variable may explain intervention effects is limited.

However, individual randomisation in a large evaluation is likely to be a labour-intensive administrative task, particularly if participants are recruited across a number of sites. It may be easier to randomise on an individual basis in trials where all participants are recruited through the same care provider, or where the number of participants in the trial is small.

Ensuring sufficient participants: sample size and power

Many evaluations of telecare and telehealth suffer from having a small number of participants. With too few participants, there is increased risk of not detecting an effect of the intervention when there is one to be found (called type II error). Caution should therefore be exercised when extrapolating findings to the larger population from studies with a small number of participants.

On the other hand, with large samples, there is an increased probability of finding an effect, but the effects found may be smaller and less likely to be clinically important. Therefore, a key task when planning your evaluation is to calculate the number of participants required to enable detection of effects of the intervention on the outcomes of interest when carrying out statistical analyses. This is referred to as a **power calculation**. A statistician can calculate the number of people required to maximise the likelihood of finding a clinically significant effect.

Power can be calculated using various online tools (e.g. Faul *et al* 2007). However, it is advisable to engage a statistician skilled in clinical trial evaluation early on in the planning process in order to identify the number of participants required. To determine the sample size needed, the evaluation design and methods will need to be carefully planned. A statistician will require information about:

- the planned statistical analyses
- **the number of measurement points:** how many times you are going to ask participants to complete measures, or how many times data will be gathered about participants from databases
- **the size of the expected effect of the intervention on the primary outcome:** for example, if 5 per cent fewer A&E visits are expected in the intervention group compared with the control group, a larger sample will be required than a situation where only a 10 per cent difference in A&E visits is expected between the intervention and control groups
- **other statistical parameters** (e.g. □ [also known as a p-value] and □): the probability of detecting an effect of the intervention.

Allowing for non-response: Your sample size should be calculated to allow for a proportion of those eligible to participate in the trial refusing to do so. Previous evaluation studies can be used to gain an idea of likely response rates for participation in such an evaluation.

Allowing for drop-outs: Your sample size should also be calculated to allow for attrition (drop-outs), which happens in all evaluations, no matter how well they are conducted. Attrition can result from various factors, including change in health status, lack of time, moving away from the area, death of the participant, or simply because

some participants may wish to withdraw from the evaluation (*see* Ellis and Price 2009). Likely attrition rates can be identified from pilot work, or from a review of relevant studies.

Selecting measures to answer the evaluation question(s)

Use **established measures of the outcomes** that you are interested in, whose properties have been evaluated to determine whether they are valid and reliable. This is particularly important when measuring non-observable outcomes (referred to as latent variables), including physiological measures and psychological measures such as quality of life, stress, coping, depression and anxiety (Bowling 1997).

Measures are **valid** if they accurately measure what they intend to measure. For example, a measure of depression is considered to be valid if it captures the extent to which a person is depressed (construct validity), and can predict depression-associated outcomes such as likelihood of hospitalisation (predictive validity). Other indicators of validity can also be used to assess whether a measure captures what it intends to measure (e.g. concurrent validity, face validity).

Measures are **reliable** if they consistently measure what they intend to. This means that they have been assessed to find out whether they will result in the same measurement from one day to the next, unless change has taken place. For example, a set of weighing scales can be considered to be reliable if they assess a person's weight as being the same two days in a row, where a person has not gained or lost weight. The same rule is applied to psychological and quality of life outcome variables.

It is not advisable to develop new measures for the purposes of an evaluation unless you have conducted a pilot (*see* page 22) to establish whether the proposed measure is valid and reliable. Developing and testing a new measure is a time-consuming process (*see* Bowling 1997).

Measures used in the WSD trial were presented at the Manchester WSDAN Roadshow presentation (*see* Newman and Davies 2009). These measures can be used to compare data across different evaluations. It should be noted that some evaluation measures are licensed and cannot be used or distributed for the purposes of an evaluation without permission.

Measurement points

Measurement points – that is, the time intervals at which measurements for the evaluation should be taken – should be selected on both a theoretical and pragmatic basis. The number of measurement points may be determined by anticipated response burden for participants, expected patterns of behaviour in relation to the telecare and telehealth equipment, and budget considerations, including availability of staff to administer measures.

Where measures are taken at baseline (that is, before the introduction of the telecare or telehealth intervention), data should be accurate. As a minimum, data should be gathered and checked in relation to participant eligibility for the trial, including their diagnoses or needs where appropriate, and use of NHS and social services. This is an important, but potentially complex and time-consuming process. Depending on resources, it may also be possible to assess patient well-being through questionnaires. For more on the issues to consider in relation to

security of data, *see* 'Data collection agreements and data sharing' on page 30.

When choosing an endpoint for the evaluation, consider the timeframe in which effects of the telecare or telehealth intervention are likely to be seen. Furthermore, as described earlier (*see* 'What are the long-term effects?' on page 13), there continues to be a dearth of evidence about the likely effectiveness of telecare or telehealth in the longer term (beyond 12 months). In any evaluation of telecare or telehealth, outcomes should be measured in the period between the introduction of the device(s) and services and the endpoint. Measurements in this intervening period can be used to determine whether, for example, there were any changes in the patterns of telehealth equipment use over time. The novelty of having the equipment installed may result in more frequent use immediately after installation. But this may decline once users become accustomed to it, providing a possible explanation for patterns in evaluation outcomes.

Planning an analytic method

The analytic method(s) selected should be appropriate to the evaluation questions, and will be driven by the evaluation design and methods. You will need to recruit appropriately skilled staff in the planning stage to ensure that planned analyses are achievable and appropriate. For example, a clinical trial statistician should be recruited to inform an evaluation of effectiveness, and a health economist should be recruited to inform an evaluation of cost-effectiveness. Depending on staff skills, they may also be required to conduct the analyses.

Minimum standards for reporting

Guidelines for the reporting of evaluations are an excellent resource to inform intervention planning. They describe the basic information that should be reported about health care evaluations, highlighting the key steps that should be addressed during development of the intervention. These are also likely to be applicable to evaluations of social care interventions. Guidelines are available to address the range of evaluation designs (*see* Table 3).

Table 3 Frameworks to guide intervention development (see 'Resources' section on page 42 for web links)

Trial design	Example reporting guidance
Resource for all evaluation designs	See EQUATOR
Randomised controlled trials	CONSORT, CONSORT plus
Observational studies	STROBE, MOOSE
Qualitative research	COREQ
Economic analyses	ISPOR RCT-CEA
Quality improvement	SQUIRE

5 Practical considerations

Piloting prior to a full-scale evaluation

We strongly recommend that you pilot your planned evaluation. Pilot studies are a way of assessing the feasibility of a planned evaluation. They are generally small-scale versions of the trial (Polit *et al* 2001), and are used to fulfil a number of aims. These include:

- testing the measures to be used
- establishing whether the evaluation protocol is feasible
- establishing whether the recruitment methods work, and understanding potential levels of attrition
- estimating variability in outcome measures to inform a sample size calculation for a main trial
- identifying the staff and other resources needed for the evaluation.

Pilots provide the opportunity to address any problems before embarking on a large-scale evaluation, thereby maximising the likelihood of a successful evaluation process. Equally, qualitative evaluations should also be piloted to ensure that the planned schedule of questions elicits useful and rich data for analysis.

Independent evaluation

Evaluations of telecare and telehealth interventions are often carried out by those who are providing the service. However, evaluations are considered to be more rigorous if carried out by independent evaluators. By clearly stating that the evaluation is independent, participants will be reassured that their responses will be treated with confidentiality. They are also more likely to feel reassured that their care will not be negatively affected by their responses. It may also prevent the participants from giving responses that they think will please the evaluator, which is a risk if they have an existing relationship with them. The outcomes will also be less subject to biases if the evaluation is carried out by people with no vested interests. Independent evaluators can be contracted from independent research organisations or universities.

The ethics of using randomised controlled trial design

Ethics of withholding a potentially beneficial technology

Service providers and stakeholders frequently raise concerns about the ethics of withholding telecare or telehealth from evaluation participants who are randomised to the control group within a trial. However, it should be noted that at the current time, telecare and telehealth are considered to be unproven technologies; as we have stated, further evidence is needed about the effectiveness of specific devices or service delivery methods, particularly in relation to specific client or patient groups. Therefore, further, robust evaluations of the effectiveness and cost-effectiveness of these technologies are needed to establish whether they can be ethically applied in normal care. A randomised design is the best method of doing this where telecare

or telehealth are not currently available to service users. A randomised controlled trial cannot be carried out where a service is currently available to users, because, for example, participants in a telecare control group would be able to request the service using their personal budget. In such instances, other methods of evaluating telecare or telehealth should be considered.

Those aiming to evaluate telecare and telehealth interventions may not currently provide them as part of their normal service. Alternatively, they may not provide a particular type of telecare or telehealth service, or may not provide it to a particular patient or client group. These present ideal situations to conduct an RCT evaluation to provide the best quality evidence about the devices and their effects on your outcomes of interest.

One method of overcoming stakeholder concerns about withholding a potentially beneficial intervention is to allow control group participants the option to have telecare or telehealth installed at the end of the evaluation, if the evaluation results in robust evidence of its effectiveness and cost-effectiveness. However, it is important that control group participants are not given telecare or telehealth kit and services where the evaluation provides evidence that it is unlikely to benefit certain participants or groups of participants. Depending on the length of the trial length, waiting times for installation of equipment for the control group may, in fact, be representative of local service provision limitations.

Blinding participants and evaluators to group allocation

Double-blind trials, in which the participant, health/social care professional and evaluator are blinded to the participant's allocated group, are the most robust method of evaluating effectiveness and cost-effectiveness, particularly in drug trials. However, as telecare and telehealth involve visible devices, and health and social care professionals are required to act upon alerts generated by the monitoring equipment, it is not possible to blind participants or their health/social care providers to study group allocation in trials of telecare and telehealth. While this is a source of confounding in an evaluation, it is one that cannot be overcome.

Obtaining ethical approval for an evaluation

Types of evaluation requiring review by a research ethics committee

All evaluation activities have the potential to have ethical implications, particularly in relation to participants (for a report on associated ethical issues, *see* Social Care Institute for Excellence (SCIE) 2010). Evaluations may involve research and development, clinical audit or service evaluation. These activities are differentiated by the amount of potential risk involved for the participants. Only those activities that constitute research and development work will need to be reviewed by a research ethics committee. The three types of evaluation are defined below.

Research is '*the attempt to derive generalisable new knowledge by assessing clearly defined methods using systematic and rigorous methods'* (Department of Health 2004).

Clinical audit is '*the quality improvement process that seeks to improve patient care and outcomes through systematic review of care against explicit criteria and the implementation of change. Aspects of the structure, processes and outcomes of care are selected against specific criteria. Where indicated, changes are implemented at an individual, team, or service level and further monitoring is used to confirm improvement in [service] delivery'* (National Institute for Clinical Excellence (NICE) 2002).

Service evaluation is '*a set of procedures to judge a service's merit by providing a systematic assessment of its aims, objectives, activities, outputs, outcomes and costs'* (NHS Executive 1997).

Submission of an ethics application for review by an ethics committee is required for an evaluation of telecare or telehealth in the following instances:

- where participants are patients and users of the NHS. This includes all potential research participants recruited by virtue of the patient or user's past or present treatment by, or use of, the NHS. It includes NHS patients treated under contracts with private sector institutions
- where participants are individuals identified as potential research participants because of their status as relatives of carers of patients and users of the NHS, as defined above
- where the evaluation requires the use of, or potential access to, NHS premises or facilities
- where participants are NHS staff – recruited as research participants by virtue of their professional role.

The criteria for each type of evaluation are set out in Appendix 2.

Applying for ethical approval for research evaluations

You can get information on the process and paperwork necessary for obtaining ethical approval from the National Patient Safety Agency's National Research Ethics Service (*see* 'Resources'). All research ethics applications are now managed through the Integrated Research Application System (IRAS). This is an electronic form, incorporating all of the required information for an application to a research ethics committee. You may be required to approach a specific research committee if your evaluation is likely to involve participants who lack mental capacity to provide consent to participate

You will be required to submit additional paperwork alongside the IRAS form. This will include a full evaluation protocol and finalised copies of the evaluation invitation/ advert, information sheets and consent forms to be distributed to participants.

Applying for research and development approval

In addition to an IRAS application, you will need to submit an application to the local research and development office at each of the proposed evaluation sites. This is to ensure that the organisation has given permission for the evaluation to take place and that someone within the organisation is responsible for it. It also ensures that there are adequate resources to conduct an evaluation, and that the evaluators adhere to any legislation in relation to the research.

Timeline for ethics applications

It can take several weeks or months to complete the necessary paperwork, depending on local resources. This is because the application has to include a full protocol and all materials to be used in the evaluation, which cannot be significantly altered once approval has been obtained from the committee. If you intend to alter the evaluation after approval, you should contact the committee again to determine whether you need to submit a substantial amendment application.

An independent scientific review of the proposed research should be conducted before submitting your application, so you need to factor this into the evaluation timetable. The specified time from submission of an application is up to 60 working days in the first instance. Where the committee asks for additional clarification or amendments, it may take a further 60 days to receive approval once the required revisions are submitted.

Recruitment and retention of participants

Recruitment

Inadequate recruitment of participants can undermine the extent to which the evaluation is able to establish the effects of the intervention on outcomes. Eligible participants who do not agree to participate can also limit the generalisability of your findings if there are systematic differences between those who participate and those who do not (Glasgow *et al* 1999). When planning an evaluation, you need to determine whether there are likely to be enough people who will fulfil your eligibility criteria, and whether they can be accessed and recruited.

To recruit an adequate sample, use evidence-based strategies for maximising recruitment referrals from front-line staff, and maximisation of referred participants. These might include:

- **maximising recruitment referrals:** engaging front-line staff in development of the evaluation protocol, providing clear and good communication with them, providing suitable training for referrers, and providing incentives for referral of eligible participants (Campbell *et al* 2007; Dormandy *et al* 2008; Foy *et al* 2003).
- **maximising recruitment of referred participants:** face-to-face recruitment may be more effective than invitations by post. A recent review has identified a number of evidence-based methods for maximising recruitment (Treweek *et al* 2010)

However, recruitment of participants should not be maximised at the expense of ensuring that the participants have a thorough understanding of what is involved in taking part (e.g. randomisation, repeated measurement and time commitment). Participants should be given the opportunity to make an informed choice as to whether to participate; invitations to participate and information sheets should provide sufficient information to enable them to do so. They should also advise participants that there is a 'cooling-off' period, during which they can withdraw if they decide not to take part, or not to continue taking part, at any time during the evaluation. You will need to submit all the information to be given to participants about what is required of them to the research ethics committee for prior assessment.

Managing participant attrition

Participant attrition (drop-outs) can affect the external validity of your evaluation findings. This is because it can limit their generalisability, particularly where there are systematic differences between those who drop out and those who continue. Data on recruitment and attrition for the WSD evaluation has been reported in previous WSDAN Roadshow presentations (*see* Ellis and Price 2009; Harburn and Carter 2009).

Participants may withdraw or drop out for a number of reasons, including:

- death
- change in eligibility (their medical condition may change, they may move to residential or nursing home care, they may move out of the area, or the carer's status may change)
- personal preference:
- they may not like the group they have been allocated to, particularly those in the control group who may be awaiting telecare or telehealth kit
- they may not like the telecare or telehealth equipment they have been allocated
- they may find completing the questionnaire or other evaluation commitments too onerous.

Some participants may not formally withdraw, but fail to complete follow-up questionnaires, either because they have moved address and contact has been lost with them, or they may be unwilling to complete the questionnaires, and may not respond to contact. In the case of participants moving home, questionnaires are likely to be returned by the postal service.

Retention strategies

There are a number of strategies you can use to minimise attrition among those participants who have not explicitly asked to withdraw (Coday *et al* 2005). These include:

- **Providing comprehensive, easy to read information sheets**

The information sheet participants receive before agreeing to take part should give clear information on the following:

- random allocation to intervention or control group (where appropriate)
- frequency and nature of contact from the evaluation team
- what is required of participants (how many questionnaires they need to complete, and how long each is likely to take)
- the likelihood that they will receive telecare or telehealth at the end of the evaluation (if they are allocated to the control group)
- how to contact the evaluation team.

- **Providing reminder contacts (pre-notification for follow-up phases)**

There is some evidence to suggest that reminder contacts, in the form of repeat questionnaire mailings, written reminders and telephone calls, can be effective in increasing questionnaire completion rates. A systematic review suggests numerous evidence-based strategies for maximising responses to postal questionnaires, some of which have been used in the WSD trial (Nakash *et al* 2006). One method that has been found to increase response and retention rates is Dillman's 'tailored design' (Dillman 1999; Dillman *et al* 2009). This method incorporates the use of multiple reminders using different modes of delivery (e.g. letters, telephone calls, reminder postcards).

Funding

You need to ensure you have adequate funds to:

- purchase telecare or telehealth equipment and services where these are not funded through current arrangements
- employ staff to carry out evaluation-related tasks
- purchase office and other equipment, including computers, licensed software and evaluation measures
- analyse, write up and disseminate findings.

The amount of funding available will dictate the type of evaluation you can carry out.

- **If you have only limited funding available** staff may be required to carry out evaluation-related tasks in addition to their current roles. You may only be able to conduct a small RCT or observational evaluation, using just a few measurements.
- **If you have some funding available** a 'medium' sized evaluation can be carried out, limited by the number of staff that can be appointed to carry it out. An RCT or observational evaluation can be conducted with a moderate number of participants and measurements.
- **If you have substantial funding available** a large-scale RCT with numerous measures can be carried out. You are likely to be able to recruit sufficient staff to carry out evaluation-related tasks.

Staffing

The budget for an evaluation will determine whether evaluation-related tasks are carried out by current service staff. A larger trial will require adequate funding to set up a dedicated evaluation team. A principal investigator should be appointed to take responsibility for the design, conduct and quality of the evaluation, as well as reporting the findings, and resolving any conflicts. You will need to appoint other staff to carry out the following tasks:

- **Manage the evaluation:** overseeing and co-ordinating all evaluation activities
- **Identify and refer potential participants:** identifying people who fulfil criteria for inclusion in the trial, and referring participants to the recruiter

- **Recruit and manage participants:** inviting participants, overseeing randomisation and allocation to intervention, distributing measurement materials and collecting data at each time point (e.g. questionnaires), and employing procedures to minimise attrition
- **Manage data:** developing an evaluation database, inputting data, and organising its editing and storage
- **Analyse data (analyst/statistician):** producing evaluation reports about recruitment and adherence, planning analysis of the design, carrying out analysis, and reporting statistical findings
- **Manage finances:** preparing budgets and managing expenditure.

The nature of the evaluation design and the number of participants required to conduct a powered analysis on outcomes data will determine the number of staff needed to carry out these tasks. One individual may carry out several of these tasks, or a staff member or group of staff may be required to take on each evaluation-related task (e.g. multiple trial administrators). Furthermore, it will be necessary to engage some clinical/ front-line staff to carry out tasks such as referrals of eligible participants.

Changes to staff roles and responsibilities

Staff may be required to alter their job roles when an evaluation is carried out. For example, GPs' working patterns may be altered where they are identifying and referring potentially eligible participants to the evaluation, or for some form of remote care. The working patterns of social and health care staff may be altered by the need to respond to triggers or alerts generated by the telecare and telehealth equipment.

Monitoring progress

Your evaluation planning should incorporate procedures for regular monitoring and reporting of:

- recruitment of referrers
- recruitment of participants and fulfilment of eligibility criteria
- refusals to participate in the evaluation and reasons why (this can be used to describe and compare non-respondents to allow statements about the generalisability of findings)
- attrition rate, and the reasons why participants drop out
- progress with data collection.

This will enable you to identify any problems with recruitment or retention of participants, and make timely changes. A full set of key performance indicators (KPIs) should be agreed with all stakeholders. These should be monitored regularly and reported back to the stakeholders and the evaluation team.

Monitoring workflow in this way also allows better planning of future administrative workload to keep the evaluation running smoothly. A presentation about monitoring KPIs was provided by KPMG at the Bristol WSDAN Roadshow (Harburn and Carter 2009). An audit trail should be used to identify when and why any changes to databases and records have taken place.

Workflow processes

If the evaluation requires participants to be followed through a process which passes between (or to and from) the responsibility of different agencies (e.g. social care, health care, evaluators), it is beneficial to plan the workflow stages. A diagram /flow chart can be created that starts from the identification of a potential participant to their exit from the programme. For example, intermediate steps may include obtaining informed consent, collecting medical/clinical data, completing questionnaires, and installing and maintaining technology – all of which may occur on more than one occasion.

Each step should include the tasks to be carried out, the person or people responsible, the data to be recorded, and the mechanisms by which the individual and their data are passed to the next stage. It is also important that processes for participant exit (withdrawal, deaths, and suspensions) and sharing this information with the other parties involved are incorporated into the workflow.

The process can be managed using workflow software (which has the ability to impose controls and validations on each step) or a centralised records system, with appropriate controls and validation procedures. When each step is completed, it should be date stamped to provide an audit trail for each participant.

Managing data

Databases

To manage participant data, you will need to develop a database. You may need separate databases to manage participant contact information and data collected from questionnaires or from health/ social care sources.

Data management programmes, such as Microsoft Excel or Access, are readily available. Alternatively, specialist database developers can design bespoke web-based programmes. Even when using widely available databases, you should ensure that you use someone with the relevant specialist skills, who is experienced in setting up and using the type of database selected. In addition, all staff who use the database should be skilled in using, or trained to use, the database that has been developed.

You should draw up clear instructions about how the database is to be used. The database and instructions should be thoroughly piloted by all who are expected to use it before the evaluation begins. This will allow you to identify any problems with the instructions or database, and to draw up and refine the validation and controls required to ensure that data is stored accurately.

Local data protection agreements should be adhered to. This will most likely require that the database(s) are password protected and encrypted (*see* 'Data collection agreements and data sharing' below). Standard operating procedures (SOPs)

should be developed by the data manager. These should be shared with all staff who use the database in order to ensure that data is always securely stored.

Data collection agreements and data sharing

In evaluations of telecare or telehealth interventions, it is likely that data will be gathered from numerous sources, such as questionnaires, local routinely collected data used for audit or service evaluation, and from other widely used sources (e.g. NHS or social services data collection systems).

Before the evaluation begins, a binding data collection agreement (DCA) should be drawn up between the evaluators and all sites involved in the evaluation. This should outline the data that is required, and who is responsible for collecting/ gathering that data. Minimum data to allow matching of participants across the databases used by different sites should be specified. This will most likely include an NHS number or social services identifier, date of birth, full name and postcode.

The DCA should clearly set out the format in which the data should be organised to ensure uniformity when files from different sites are merged. For example, where dates are entered into a database, the specific format for the date (eg, DD-MM-YYYY) should be agreed in advance. It may be necessary for individual sites to receive a site-specific DCA; however, data-sharing policies should be as uniform as possible to ensure that the process of merging data from different sites is as smooth as possible.

Data is subject to the Data Protection Act (*see* below). Data sharing between sites involved in an evaluation should only be done with files that are encrypted, and have strong password protection (e.g. using a long password with a mix of letters, symbols and numbers). Where sensitive participant data is to be moved between sites, only encrypted mail services should be used (e.g. NHS mail), or encrypted data storage devices.

Version control

It is likely that you will need to make changes to the format of databases during the course of the evaluation, or that data-sharing agreements may need to be altered to reflect circumstantial changes. A team member should be responsible for managing all such alterations. A version control process is required, whereby one person manages changes, highlights alterations and circulates them to all those who use the databases on a regular basis.

The version control process should also ensure that the database into which participant data is entered is regularly backed up. All staff who use the database should be aware of the relevant version control procedure. In the WSD trial, in addition to nightly server back-ups, the databases are routinely saved, with a new version created each day. This means that data can quickly and easily be restored should there be a problem.

Data protection

The Data Protection Act (1998) and local data protection procedures should be adhered to. These will need to be specified in an ethics application. Any data that

is accessible via a server should be saved on a server that is secure, and to which there is limited access. Encrypted, password-protected files should be used.

Pseudonymisation

Data protection agreements frequently require that data identifying a participant and data collected from questionnaires or routinely used data collection systems are stored separately. In order to link the data, identifiers that do not use an individual's name or contact information should be stored on both databases. Each individual participating in the evaluation should be allocated a unique indentifier. This may include a participant number, NHS number or social service identifiers.

Data screening/ checking

It is normal that some degree of human error may occur during data entry, although procedures should be in place to minimise this. To check for errors, validation rules should be established and checks should be run on a regular basis. Validation rules can be used to ensure that out-of-range data is not entered into the database, as this will result in distorted findings. Regular validation checks of eligibility data, and of data collected subsequently, including questionnaire data and device-monitoring data, will ensure that any problems identified are promptly addressed. For example, checks should be routinely carried out to ensure that participants' questionnaire data responses are within the ranges specified. If a response scale ranges from 1 to 10, an algorithm can be used to check that no keying errors have been made, in which values outside of this range are entered into the database.

6 Summary

This briefing has given an overview of the most important issues to consider when developing a good-quality evaluation of telecare or telehealth interventions.

- Careful planning is the key to carrying out a good-quality evaluation. Allow sufficient time for the planning and design stages.
- The evaluation should be adequately funded and resourced, using staff with the appropriate skills to carry out the necessary tasks.
- Set clear questions to be answered from the outset, to guide your planning. These can be identified from a review of the literature so as to address gaps in current knowledge, and by considering important local issues.
- Select a design that is appropriate to, and can therefore accurately answer, the evaluation questions.
- Clearly define the population for which the intervention will be evaluated. Set out clear criteria for inclusion and exclusion and ensure that these are adhered to. Make sure you identify an appropriate number of participants for the planned analyses.
- Use valid and reliable measures of the outcomes of interest, and consider the time points at which they will be administered.
- Factor in time to allow for completion of ethics or research and development applications, where the evaluation constitutes research.
- Use evidence-based recruitment and retention strategies to ensure high levels of participation and to minimise attrition.
- Ensure that data collection agreements between the sites involved are put into place prior to the start of data collection. Data management strategies, including frequent validation checks, should be factored into the planning.

Appendix 1: Glossary of sampling terms

(Adapted from Bowling 1997)

Convenience Sampling (also known as opportunistic sampling)	Sampling is not based on random selection or probability; the researcher selects convenient participants as respondents. Therefore, it is not guaranteed that all those who might be eligible in a population have an equal chance of being sampled.
Purposive sampling	A deliberately non-random method of sampling. A sample of a group of people, or settings, with a particular characteristic is used. This can be used within qualitative research, or may be used in experimental designs for practical reasons. For example, a medical team may invite all their current inpatients to test the effectiveness of a new treatment.

Random sampling	**Simple random sampling**: The members of a population of interest are numbered and a proportion of them are selected using random numbers. Therefore, each member of the population has an equal chance of being selected for the sample. This may sometimes be referred to as: ***systematic* random sampling** if the population is already systematically organised (e.g. alphabetically), or ***stratified* random sampling**, where the population is divided into strata representing groups, and samples are taken from each stratum using random sampling. This method guards against obtaining an unrepresentative sample which under- or over-represents certain groups of the population.
Snowballing (sampling)	This technique is used where no sampling frame exists and one cannot be created. For example, there may be no list of the people that you aim to recruit. Therefore, an initial group of participants is invited and the respondents invite others that they know are in the target group from their social group (e.g. friends, family, colleagues). Anyone identified is contacted and asked during their interview to identify others who may wish to take part.

Theoretical sampling	Conceptual or theoretical categories are generated during the research process. This sampling method aims to develop and challenge emerging hypotheses. A small number of similar cases of interest are selected and interviewed to develop an understanding of the particular phenomenon. Cases are then sampled who might be exceptions, in an attempt to challenge the developed and emerging hypothesis.
Quota sampling	This is favoured by market researchers for its convenience and the speed with which participants are recruited. It is a method of stratified sampling, often in geographical areas. The area may be selected randomly, and then quotas of participants are interviewed on the basis of selection criteria decided by the interviewers, who are aware how many interviews they need to carry out in each stratum (e.g. the number of men, women, people with diabetes, people with COPD). They may then stand in the street, or go from door to door asking people to participate. This method is unlikely to result in a representative sample because of the likelihood of who might be in the street or at home when the interviewer attempts to recruit.

Appendix 2: Criteria for research evaluation, clinical audit and service evaluation

(Adapted from NRES Ethics Consultation E-Group and NHS Research and Development Forum)

Research evaluation	Clinical audit	Service evaluation
Requires research ethics committee review	Does not require research ethics committee review	Does not require research ethics committee review
Conducted to systematically gain new knowledge to identify best practice, and may change practice	Conducted to produce information to inform delivery about best care/ practice and may change practice to improve quality. It may ask whether: - what ought to be happening is happening - current practice meets required standards - current practice follows published guidelines - clinical practice is applying the knowledge that has been gained through research - current evidence is being applied in a given situation	Conducted to define or evaluate current care
Can be quantitative research, to test a hypothesis, or qualitative research, to identify/ explore themes using established methodology	Can be quantitative or qualitative research Establishes whether the service reaches a pre-determined standard	Can be quantitative or qualitative research Designed to answer –'what standard does this service achieve?'
Addresses clearly defined questions, aims and objectives	Measures against a standard	Measures current service without reference to a standard

Quantitative research may involve evaluating or comparing interventions, may test a new practice/ therapy/ drug which may be invasive Qualitative research- involves studying the experience of interventions and how relationships are established	Involves an intervention currently in use only. Choice of intervention is decided by patient/ client or professional, and determined using current guidance or professional standards	Involves an intervention currently in use only. Choice of intervention is decided by patient/ client or professional, and determined using current guidance or professional standards
Strict selection criteria for the participants involved, defined using eligibility criteria	Eligible participants can be all those receiving the service under investigation	Eligible participants can be all those receiving the service under investigation
May involve contact with participants/ patients	May or may not involve patient contact, but does not generally involve change to normal clinical care	May or may not involve patient contact, but does not generally involve change to normal clinical care
May involve (random) allocation of participants to two or more groups	No allocation to intervention groups: the health care professional and patient have chosen the intervention to be received before audit	No allocation to intervention groups: the health care professional and patient have chosen the intervention to be received before service evaluation
Participants may be patients/clients, patients/ clients as volunteers or healthy volunteers	Participants are always patients/clients currently receiving care within a service	Participants are always patients/clients currently receiving care within a service
Protocol driven: the design will not be altered part-way through the evaluation		
May involve collecting data from participants outside of/ in addition to data that is ordinarily collected. May involve collecting data from medical records, using samples from tissues or investigations additional to routine care	Usually involves analysis of existing data but may include administration of a simple interview or questionnaire	Usually involves analysis of existing data but may include administration of a simple interview or questionnaire

References

Barlow JG, Singh D, Bayer S, Curry R (2007). 'A systematic review of the benefits of home telecare for frail elderly people and those with long-term conditions'. *Journal of Telemedicine and Telecare*, vol 13, no 4, pp 172–79. Available at: www.ncbi.nlm.nih.gov/pubmed/17565772 (accessed on 27 October 2010).

Bensink M, Hailey D, Wootton R (2006). 'A systematic review of successes and failures in home telehealth: preliminary results'. *Journal of Telemedicine and Telecare,* vol 12, supplement 3, pp 8–16. Available at: http://jtt.rsmjournals.com/cgi/content/abstract/12/suppl_3/8 (accessed on 27 October 2010).

Bergmo TS (2009). 'Can economic evaluation in telemedicine be trusted? A systematic review of the literature'. *Cost Effectiveness and Resource Allocation,* vol 24, pp 7–18. Available at: www.resource-allocation.com/content/7/1/18 (accessed on 27 October 2010).

Bowling A (1997). *Research Methods in Health: Investigating health and health services*. Buckingham: Open University Press.

Campbell MK, Snowdon C, Francis D, Elbourne D, McDonald AM, Knight R, Entwistle V, Garcia J, Roberts I, Grant A (2007). 'Recruitment to randomised trials: strategies for trial enrolment and participation study. The STEPS study'. *Health Technology Assessment*, vol 11, no 48. Available at: www.hta.ac.uk/execsumm/summ1148.htm (accessed on 27 October 2010).

Coday M, Boutin-Foster C, Goldman Sher T, Tennant J, Greaney ML, Saunders SD, Somes GW (2005). 'Strategies for retaining study participants in behavioral intervention trials: retention experiences of the NIH Behavior Change Consortium'. *Annals of Behavioural Medicine*, vol 29, pp 55–65. Available at:www.ncbi.nlm.nih.gov/pubmed/15921490 (accessed on 27 October 2010).

Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M (2008). 'Developing and evaluating complex interventions: the new Medical Research Council guidance'. *British Medical Journal*, vol 337, a1655. Available at: www.ncbi.nlm.nih.gov/pubmed/18824488 (accessed on 27 October 2010).

Dang S, Dimmick S, Kelkar G (2009). 'Evaluating the evidence base for the use of home telehealth remote monitoring in elderly with heart failure'. *Telehealth Journal and e-Health*, vol 15, no 8, pp 783–96. Available at: www.ncbi.nlm.nih.gov/pubmed/19831704 (accessed on 27 October 2010).

Darkins A, Ryan P, Kobb R, Foster L, Edmonson E, Wakefield B, Lancaster AE (2008). 'Care Coordination/ Home Telehealth: the systematic implementation of health informatics, home telehealth, and disease management to support the care of veteran patients with chronic conditions'. *Telemedicine and e-Health,* vol 14, no 10, pp 1118–26. Available at:www.ncbi.nlm.nih.gov/pubmed/19119835 (accessed on 27 October 2010).

Data Protection Act (1998). www.opsi.gov.uk/acts/acts1998/ukpga_19980029_en_1

Dávalos ME, French MT, Burdick AE, Simmons SC (2009). 'Economic evaluation of telemedicine: review of the literature and research guidelines for benefit-cost analysis'. *Telemedicine Journal and e-Health*, vol 15, no 10, pp 933–48. Available at: www.ncbi.nlm.nih.gov/pubmed/19954346 (accessed on 27 October 2010).

Davies A, Newman S (2010). 'Telecare and telehealth: critically examining the evidence'. WSDAN Roadshow presentation, Stansted, February. Available at :www.wsdactionnetwork.org.uk/past_events/wsdan_roadshow_11.html (accessed on 27 October 2010).

DelliFraine JL, Dansky KH (2008). 'Home-based telehealth: a review and meta-analysis'. *Journal of Telemedicine and Telecare,* vol 14, pp 62–66. http://jtt.rsmjournals.com/cgi/content/abstract/14/2/62 (accessed on 21 October 2010).

Department of Health (2010). *Whole System Demonstrators.* Available at: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Healthcare/Longtermconditions/wholesystemdemonstrators/DH_084255 (accessed on 21 October 2010).

Department of Health (2009). *Whole System Demonstrators: An overview of telecare and telehealth*. Available at:www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_100947.pdf (accessed on 21 October 2010).

Department of Health (2007). *Long-Term Conditions*. Available at:http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Healthcare/Longtermconditions/DH_4128521 (accessed on 21 October 2010).

Department of Health (2004). *The Research Governance Framework for Health and Social Care: Implementation plan for social care*. Available at: www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4082172.pdf (accessed on 21 October 2010).

Dillman DA (1999). *Mail and Internet Surveys: The tailored design method.* New Jersey: John Wiley & Sons, Inc.

Dillman DA, Smyth JD, Christian LM (2009). *Internet, Mail, and Mixed-Mode Surveys: The tailored design method*, 3rd ed. New Jersey: John Wiley & Sons, Inc.

Dormandy E, Kavalier F, Logan J, Harris H, Ishmael N, Marteau T (2008). 'Maximising recruitment and retention of general practices in clinical trials: a case study'. *British Journal of General Practice,* vol 58, no 566, pp 759–66. Available at:

www.ncbi.nlm.nih.gov/pubmed/19000399 (accessed on 21 October 2010).

Ellis T, Price H (2009). 'Whole System Demonstrators (WSD) Programme: An overview of participants leaving the trial and Kent experiences'. WSDAN Roadshow presentation, Leeds, June. Available at: www.wsdactionnetwork.org.uk/past_events/wsdan_roadshow.html (accessed on 27 October 2010).

Evans D (2003). 'Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions'. *Journal of Clinical Nursing,* vol 12, no 1, pp 77–84. Available at: www.ncbi.nlm.nih.gov/pubmed/12519253 (accessed on 21 October 2010).

Faul F, Erdfelder E, Lang A-G, Buchner A (2007). 'G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences'. *Behavior Research Methods,* vol 39, no 2, pp 175–91. Available at:www.ncbi.nlm.nih.gov/pubmed/17695343 (accessed on 21 October 2010).

Foy R, Parry J, Duggan A, Delaney B, Wilson S, Lewin-van den Broek N, Lassen A, Vickers L, Myres P (2003). 'How evidence based are recruitment strategies to randomized controlled trials in primary care? Experience from seven studies'. Family Practice, vol 20, no 1, pp 83–92. Available at:www.ncbi.nlm.nih.gov/pubmed/12509377 (accessed on 21 October 2010).

Glasgow RE, Vogt TM, Boles TM (1999). 'Evaluating the public health impact of health promotion interventions: the RE-AIM framework'. *American Journal of Public Health,* vol 89, no 9, pp 1323–27. Available at:www.ncbi.nlm.nih.gov/pubmed/10474547 (accessed on 21 October 2010).

Glueckauf RL, Ketterson TU (2004). 'Telehealth interventions for individuals with chronic illness: research review and implications for practice'. *Professional Psychology: Research and Practice,* vol 35, no 6, pp 615–27. Available at: http://cat.inist.fr/?aModele=afficheN&cpsidt=16321018 (accessed on 21 October 2010)

Harburn F, Carter R (2009). 'Whole System Demonstrators (WSD) Programme, WSD Programme & Project Management: An overview'. WSDAN Roadshow presentation, Bristol, September. Available at: www.wsdactionnetwork.org.uk/past_events/wsdan_roadshow_10.html (accessed on 21 October 2010).

Holloway I, Wheeler S (1995). 'Ethical issues in qualitative nursing research'. *Nursing Ethics,* vol 2, no 3, pp 223–32. Available at:www.ncbi.nlm.nih.gov/pubmed/7583428 (accessed on 27 October 2010).

Jaana M, Paré G, Sicotte C (2009). 'Home telemonitoring for respiratory conditions: a systematic review'. *American Journal of Managed Care*, vol 15, no 5, pp 313–20. Available at:www.ncbi.nlm.nih.gov/pubmed/19435399 (accessed on 27 October 2010).

Klersy C, De Silvestri A, Gabutti G, Regoli F, Auricchio A (2009). 'A meta-analysis of remote monitoring of heart failure patients'. *Journal of the American College of Cardiology,* vol 54, no 18, pp 1683–94. Available at:www.ncbi.nlm.nih.gov/pubmed/19850208 (accessed on 27 October 2010).

Knapp M (2010). 'Telecare and telehealth: economic issues'. WSDAN Roadshow presentation, Gateshead, June. Available at: www.wsdactionnetwork.org.uk/past_events/wsdan_roadshow_30.html (accessed on 27 October 2010).

Maric B, Kaan A, Ignaszewski A, Lear SA (2009). 'A systematic review of telemonitoring technologies in heart failure'. European Journal of Heart Failure, vol 11, no 5, pp 506–17. Available at: www.ncbi.nlm.nih.gov/pubmed/19332417 (accessed on 27 October 2010).

Martin S, Kelly G, Kernohan WG, McCreight B, Nugent C (2008). 'Smart home technologies for health and social care support (Cochrane Review)'. *Cochrane Database of Systematic Reviews*. Available at:http://www2.cochrane.org/reviews/en/ab006412.html (accessed on 27 October 2010).

Nakash RA, Hutton JL, Jørstad-Stein EC, Gates S, Lamb SE (2006). 'Maximising response to postal questionnaires: a systematic review of randomised trials in health research'. *BMC Medical Research Methodology*, vol 6, no 5. Available at: www.ncbi.nlm.nih.gov/pubmed/16504090 (accessed on 27 October 2010).

Neubeck L, Redfern J, Fernandez R, Briffa T, Bauman A, Freedman SB (2009). 'Telehealth interventions for the secondary prevention of coronary heart disease: a systematic review'. *European Journal of Cardiovascular Prevention and Rehabilitation,* vol 16, no 3, pp 281–89. Available at: www.ncbi.nlm.nih.gov/pubmed/19407659 (accessed on 27 October 2010).

Newman S, Davies A (2009). 'Planning an evaluation of telecare and telehealth'. WSDAN Roadshow presentation, Manchester, December. Available at: www.wsdactionnetwork.org.uk/past_events/wsdan_roadshow_10_2.html (accessed on 27 October 2010).

National Assistance Act (1948). www.opsi.gov.uk/revisedstatutes/acts/ukpga/1948/cukpga_19480029_en_1

NHS Executive (1997). *An Introduction to Service Evaluation.* Royal College of Psychiatrists Research Unit. Cited in E Hardman and C Joughin (1998) FOCUS on Clinical Audit in Child and Adolescent Mental Health Services. Royal college of Psychiatrists, London. Available at http://www.rcpsych.ac.uk/pdf/clinauditPrelims.pdf (accessed on 9 December 2010).

National Institute for Clinical Excellence (NICE) (2002). *Principles for Best Practice in Clinical Audit*. Available at: www.nice.org.uk/niceMedia/pdf/BestPracticeClinicalAudit.pdf (accessed on 21 October 2010).

Paré G, Jaana M, Sicotte C (2007). 'Systematic review of home telemonitoring for chronic diseases: the evidence base'. *Journal of the American Medical Informatics Association,* vol 14, no 3, pp 269–77. Available at: www.ncbi.nlm.nih.gov/pubmed/17329725 (accessed on 27 October 2010).

Perry J, Beyer S, Francis J, Holmes P (2010). 'Ethical issues in the use of telecare'. Social Care Institute for Excellence (SCIE). Available at: www.scie.org.uk/publications/reports/report30.pdf (accessed on 27 October 2010).

Polisena J, Coyle D, Coyle K, McGill S (2009). 'Home telehealth for chronic disease management: a systematic review and an analysis of economic evaluations'. *International Journal of Technology Assessment in Health Care,* vol 25, no 3, pp 339–49. Available at: www.ncbi.nlm.nih.gov/pubmed/19619353 (accessed on 21 October 2010).

Polisena J, Tran K, Cimon K, Hutton B, McGill S, Palmer K, Scott RE (2010). 'Home telemonitoring for congestive heart failure: a systematic review and meta-analysis'. Journal of Telemedicine and Telecare, vol 16, no 2, pp 68–76. Available at:www.ncbi.nlm.nih.gov/pubmed/20008054 (accessed on 21 October 2010).

Polit DF, Beck CT, Hungler BP (2001). *Essentials of Nursing Research: Methods, appraisal, and utilization*, 5th ed. Philadelphia: Lippincott Williams & Wilkins.

Rojas SV, Gagnon MP (2008). 'A systematic review of the key indicators for assessing telehomecare cost-effectiveness'. *Telemedicine Journal and e-Health,* vol 14, pp 896–904. Available at:www.ncbi.nlm.nih.gov/pubmed/19035798 (accessed on 27 October 2010).

Roland M, Torgerson DJ (1998). 'What are pragmatic trials?' *British Medical Journal*, vol 316, no 7127, p 285. Available at: www.ncbi.nlm.nih.gov/pubmed/9472515 (accessed on 27 October 2010).

Schmidt S, Schuchert A, Krieg T, Oeff M (2010). 'Home telemonitoring in patients with chronic heart failure: a chance to improve patient care?' *Deutsches Arzteblatt International*, vol 107, no 8, pp 131–38. Available at:www.ncbi.nlm.nih.gov/pubmed/20300221 (accessed on 27 October 2010).

Shiell A, Hawe P, Gold L (2008). 'Complex interventions or complex systems? Implications for health economic evaluation'. British Medical Journal, vol 336, no 7656, pp 1281–3. Available at:www.ncbi.nlm.nih.gov/pubmed/18535071 (accessed on 27 October 2010).

Shadish WR, Cook TD, Campbell DT (2002). Experimental and Quasi-Experimental Designs for Generalized Causal Inference. New York: Houghton Mifflin Company.

Sibbald B, Roland M (1998). 'Understanding controlled trials: why are randomised controlled trials important?' *British Medical Journal,* vol 316, no 7126, p 201. Available at:pubmed/9468688 (accessed on 21 October 2010).

Treweek S, Pitkethly M, Cook J, Kjeldstrøm M, Taskila T, Johansen M, Sullivan F, Wilson S, Jackson C, Jones R, Mitchell E (2010). 'Strategies to improve recruitment to randomised controlled trials (Cochrane Review)'. Cochrane Database of Systematic Reviews, issue 4. Available at: www2.cochrane.org/reviews/en/mr000013.html (accessed on 27 October 2010).

Wainwright C, Wootton R (2003). 'A review of telemedicine and asthma'. *Disease Management and Health Outcomes,* vol 11, no 9, pp 557–63. Available at: http://adisonline.com/diseasemanagement/Abstract/2003/11090/A_Review_of_Telemedicine_and_Asthma.3.aspx (accessed on 27 October 2010).

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, vol 337, a2390. Available at:www.ncbi.nlm.nih.gov/pubmed/19001484 (accessed on 27 October 2010).

Resources

CONSORT: Schulz KF, Altman DG, Moher D, CONSORT Group (2010). 'CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials'. *British Medical Journal,* vol 340, c332. Available at: www.ncbi.nlm.nih.gov/pubmed/20332509 (accessed on 3 November 2010).

CONSORT PLUS: Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D, 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*, vol 337, a2390. Available at: www.ncbi.nlm.nih.gov/pubmed/19001484 (accessed on 3 November 2010).

COREQ: Tong A, Sainsbury P, Craig J (2007). 'Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups'. *International Journal for Quality in Health Care*, vol 19, no 6, pp 349–57. Available at: www.ncbi.nlm.nih.gov/pubmed/17872937 (accessed on 3 November 2010).

The Cochrane Collaboration www.cochrane.org/

EQUATOR: Simera I, Moher D, Hoey J, Schulz KF, Altman D (2010). 'A catalogue of reporting guidelines for health research'. *European Journal of Clinical Investigation*, vol 40, no 1, pp 35–53. Available at: www.ncbi.nlm.nih.gov/pubmed/20055895 (accessed on 27 October 2010). See also: www.equator-network.org/

G Power www.psycho.uni-duesseldorf.de/aap/projects/gpower/

IRAS www.myresearchproject.org.uk

ISPOR RCT-CEA (International Society for Pharmacoeconomics and Outcomes Research Randomized Clinical Trials - Cost-Effectiveness Analysis): Ramsey, S., Willke, R., Briggs, A., Brown, R., Buxton, M., Chawla, A., Cook, J., Glick, H., Liljas, B., Petitti, D. & Reed S (2005). 'Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA task force report'. *Value in Health*, vol 8, pp 521–33. Available at: www.ncbi.nlm.nih.gov/pubmed/16176491 (accessed on 27 October 2010).

MOOSE: Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB, Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group (2000). 'Meta-analysis of observational studies in epidemiology: a proposal for reporting'. *Journal of the American Medical Association,* vol 283, no 15, pp 2008–2012. Available at: www.ncbi.nlm.nih.gov/pubmed/10789670 (accessed on 27 October 2010).

National Patient Safety Agency National Research Ethics Service www.nres.npsa.nhs.uk

NHS Research and Development guidance www.rdforum.nhs.uk/011.asp

Patients lacking mental capacity – ethical guidance: National Patient Safety Agency National Research Ethics Service (2007). 'Research involving adults unable to consent for themselves'. Available at: www.nres.npsa.nhs.uk/EasySiteWeb/GatewayLink.aspx?alId=321 (accessed on 3 November 2010).

PubMed www.ncbi.nlm.nih.gov/pubmed

SQUIRE: Davidoff F, Batalden P, Stevens D, Ogrinc G, Mooney S (2008). 'Publication guidelines for improvement studies in health care: evolution of the SQUIRE Project'. *Annals of Internal Medicine*, vol 149, no 9, pp 670–76. Available at: www.ncbi.nlm.nih.gov/pubmed/18981488 (accessed on 3 November 2010).

STROBE: von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP (2007). 'The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies'. *Annals of Internal Medicine*, vol 147, pp 573–77. Available at:www.ncbi.nlm.nih.gov/pubmed/17938396 (accessed on 3 November 2010).

WSDAN evidence database www.wsdactionnetwork.org.uk/resources/evidence_database.html