An epistemology of patient safety research: a framework for study design and interpretation. Part 1. Conceptualising and developing interventions

C Brown,[1] T Hofer,[2] A Johal,[1] R Thomson,[3,4] J Nicholl,[5] B D Franklin,[6] R J Lilford[1]

See Editorial, p 154

[1] Department of Public Health and Epidemiology, University of Birmingham, Birmingham, UK; [2] University of Michigan Medical School, Ann Arbor, Michigan, USA; [3] National Patient Safety Agency, London, UK; [4] Newcastle upon Tyne Medical School, Newcastle upon Tyne, UK; [5] University of Sheffield, Sheffield, UK; [6] London School of Pharmacy, London, UK

Correspondence to:
Dr C Brown, Research Methodology Programme, Department of Public Health and Epidemiology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK;
c.a.brown@bham.ac.uk

Accepted 18 January 2008

ABSTRACT

This is the first of a four-part series of articles examining the epistemology of patient safety research. Parts 2 and 3 will describe different study designs and methods of measuring outcomes in the evaluation of patient safety interventions, before Part 4 suggests that "one size does not fit all". Part 1 sets the scene by defining patient safety research as a challenging form of service delivery and organisational research that has to deal (although not exclusively) with some very rare events. It then considers two inter-related ideas: a causal chain that can be used to identify where in an organisation's structure and/or processes an intervention may impact; and the need for preimplementation evaluation of proposed interventions. Finally, the paper outlines the authors' pragmatist ontological stance to patient safety research, which sets the philosophical basis for the remaining three articles.

We have documented a massive rise in patient safety research over the past decade.[1] Much of this consists of basic research in cognate disciplines such as psychology, sociology, organisational studies, ergonomics and education. Improving patient safety requires that the knowledge gleaned from basic science and clinical research should be taken up in the design of interventions to improve patient care. This is the first of four articles dealing with the development and evaluation of such interventions. In these articles we summarise the results of a report to the Medical Research Council (MRC) "The epistemology of patient safety research: a framework for study design and interpretation".[2] In this first article we set the scene. We first describe the causal chain through which interventions designed to reduce the number of patient safety incidents impact on complex systems such as healthcare organisations. Second, we discuss the process through which interventions should be selected and refined even before they are rolled out in practice: preimplementation evaluation. Last, we describe and briefly defend the philosophical (ontological and epistemological) premises to which we subscribe and on which the arguments in subsequent articles will build.

Part 2 of this series will consider the different types of study design that could be used when interventions are evaluated in healthcare organisations and will show that the different methods have strengths and weaknesses that vary according to the type of intervention being evaluated: one size does not fit all. In Part 3 we will discuss how quality and safety may be measured, with particular reference to potential biases. In Part 4 we will bring the various themes together and show how many different sources of knowledge, including the results of preimplementation evaluations and data arising from different points in the causal chain, can be integrated in a bayesian statistical framework.

OUR POINT OF DEPARTURE: THE NATURE OF PATIENT SAFETY INTERVENTIONS

Safer care can sometimes be achieved by replacing unsafe treatments and technologies (such as medical devices or surgical techniques) with safer alternatives and the study of these alternatives is often referred to as health technology assessment (HTA). Healthcare can also be made safer by more appropriate use of existing treatments—that is, improvements in the system in which patient care is embedded. The study of methods to strengthen the system is often considered to fall within service delivery and organisational (SDO) research or health services research. Since the patient safety movement is usually conceptualised in terms of seeking to improve the systems within which staff work,[3] most patient safety research would seem to fit under the broad heading of SDO research.[4] For all that it might be part of SDO research, safety research does have one idiosyncrasy with strong methodological implications: the extreme rarity of many of the events that safety interventions aim to reduce.

THE SAFETY/QUALITY CONTINUUM AND THE RARITY OF SAFETY EVENTS

The safety incidents that make the news usually involve dramatic events such as wrong site surgery, lethal dose miscalculations, deaths following inadvertent intravenous potassium chloride concentrate and inadvertent intrathecal administration of vincristine. As well as leading to severe harm, the link between error and the adverse outcome in these high-profile cases is:

- immediate (or rapid);
- certain (or highly likely).

For these reasons, such errors are sometimes given the sobriquet "egregious": the fact that an error has produced the poor outcome is indisputable. These errors also have a third feature in common:

- they are all very rare—in some cases a country the size of the UK may experience fewer than one case per year.

Not all safety incidents make the news: those that do not tend to be those with lower immediacy and causality, but which occur more frequently.

Table 1 Examples of clinical events/error with differing degrees of causality and immediacy

Error	Causality	Immediacy	Point on fig 1
Intrathecal administration of vincristine/potassium chloride, resulting in death	High	High	A
Failure to vaccinate, resulting in the individual developing the disease that the vaccine was intended to prevent	High	Low	B
Failure to use thrombolytics in myocardial infarction, resulting in death due to further myocardial infarction	Low	High	C
Failure to use β-blockers after myocardial infarction, resulting in subsequent myocardial infarction	Low	Low	D

Table 1 gives examples of different types of error on the dimensions of immediacy and causality. As a general rule, very rare errors with high immediacy and causality generate concerns over *safety*. More frequent events with low immediacy and causality, such as failure to follow evidence-based guidelines on vaccination, are often conceptualised as *quality* rather than safety issues. Indeed, the performance of healthcare providers is assessed against targets such as vaccination rates. However, we do not believe that safety can be distinguished from quality purely by the egregiousness of the link between error and outcome. Like Hofer and colleagues,[5] we identify a safety/quality continuum based on a vector of egregiousness (fig 1), and like them define no clear point on this vector at which safety topples over into quality. Likewise, errors that do not lie on (or close to) this vector (eg, points B and C in fig 1) do not fall into exclusive categories of quality or safety. Hence there is no clear divide between safety research and SDO research more generally. Furthermore, at a population level, high frequency but lower harm, immediacy or causality incidents may contribute more harm overall, for example failure to detect or act on the deteriorating patient[6 7] or the problem of falls in hospitals.[8]

By causality we mean the confidence with which a bad outcome, if it occurs, can be attributed to the error. So if someone who should have been vaccinated against influenza contracts the prevalent strain of the disease, it is quite possible

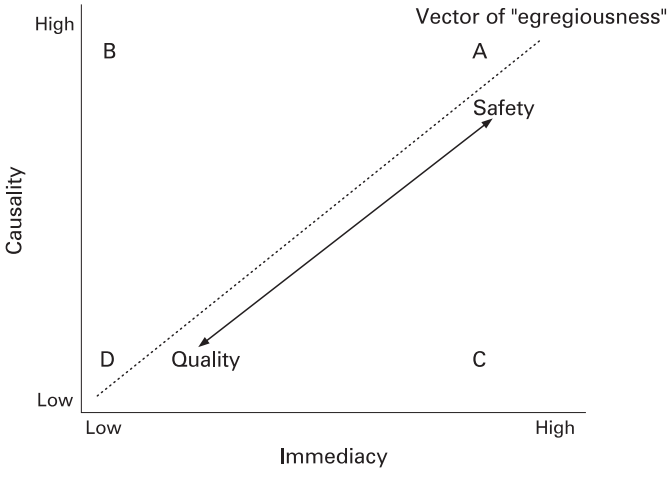

Figure 1 The quality/safety continuum. Note: the letters A–D refer to examples of clinical errors with different degrees of causality and immediacy provided in table 1.

that this could have been prevented. On the other hand, the reoccurrence of myocardial infarction in an individual patient might not have been prevented by β-blockers, even if at a population level the benefits are clear.

In this article we lay out a framework for SDO research with special emphasis on rare events at the safety end of the continuum. The three recurring messages in this four-part series are listed below.

- Interventions should be developed in the light of a clear understanding of the *causal chain* through which they may impact (positively and negatively) on an organisation (and hence on patients).
- Safety interventions should also be examined at all levels in this causal chain they may influence. This examination should allow both the positive and negative effects of the intervention to be identified.
- It is auspicious that, whenever possible, evaluations of safety interventions should be planned concurrently with the intervention itself so that the opportunity to collect baseline data is not missed.

With these ideas in mind we now present a conceptual framework for this causal chain at the system level in healthcare organisations.

A CAUSAL CHAIN LINKING INTERVENTIONS TO OUTCOME

The concept of a causal chain draws heavily on Donabedian,[9] who distinguished between structure, process and outcome, and Reason,[3] who wrote of latent and active errors. Behind both of these concepts lies the generic idea of a service (frontline healthcare) embedded in a system. We have built on these ideas to create a conceptual model of the system within which a healthcare organisation operates as shown in fig 2. Like Donabedian, we start the causal chain with the "structure" within which a service is delivered. By structure, we mean the exogenous factors or "givens" that cannot be completely determined by managers within a particular healthcare organisation. Depending on the national context, these may include national directives, licensing procedures and the resource intensive building blocks of care, such as the provision of buildings, staff and equipment and the budgets that constrain staff–patient ratios.

Next in the chain comes the endogenous processes that are under local control. We distinguish between two types of process: management/organisational processes (eg, human resource policy; training of new staff; management of the supply chain) and clinical or front-end processes (eg, adoption of particular safety/evidence-based practices; the quality of clinician–patient communication). This distinction accords with Reason's distinction between latent errors at an organisational level and active errors which involve direct human interaction.[3] Interventions focused on management processes, such as human resource policies, eg, staff appraisal and management "walkabouts" (the presence of senior management on the wards—see Frankel *et al*[10]) will generally affect patient safety outcomes through their effect on intervening variables and staff behaviours/attitudes, such as morale, culture or sickness absence. Alternatively, interventions may be designed to impact directly on a clinical domain, such as use of a "forced function" engineering solution to prevent misconnecting anaesthetic tubes. Last in the chain are clinical outcome and throughput (eg, number of patients treated). In Parts 3 and 4 of this series we will argue that safety/quality interventions should be studied at all levels along this chain. A systems-level approach in which the causal chain is considered as a whole is also useful

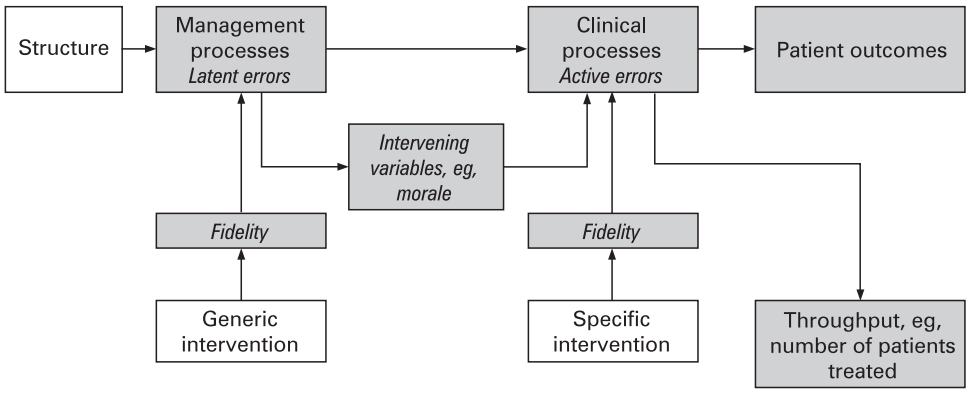

Figure 2 General and specific interventions across the system and evaluation end points. The shaded boxes represent the end points that could be measured in an evaluation of a patient safety intervention. Surrogate end points are shown in italics.

at the development phase for new interventions. We now turn to this topic of developing new interventions and selecting those that are suitable to be rolled out into practice.

DEVELOPING INTERVENTIONS AND PREIMPLEMENTATION EVALUATION

Interventions to improve patient safety do not just appear: they have to be conceived, designed and selected. This preimplementation phase is important in selecting the most propitious interventions. The MRC in the UK[11] has specified a framework for the evaluation of complex interventions which starts before the intervention is introduced in practice—a preimplementation evaluation (PIE). Campbell and colleagues[12] have recently published a practical guide to the implementation of the MRC framework using examples from primary care research. We conceptualise PIE in four broad stages:

Stage 1

PIE begins with recognition of the need for an intervention to improve patient safety. Such evidence could be generated from epidemiological data, internal or external performance management/audit data, local or national error reporting data, the medical literature and the experiences of clinical and non-clinical staff. Patient safety problems may be identified retrospectively (in response to actual errors/adverse events) or prospectively (to mitigate any anticipated errors/adverse events).

Stage 2

The intervention should build on a thorough understanding and description of existing practice that should be studied systematically at all levels in the causal chain shown in fig 2. Using Reason's model[3] this is analogous to identifying where in the system ("Swiss cheese") the largest holes are. Root cause analysis is a method of tracking back from a patient safety incident to potential holes in the system.[13] A system can also be studied systematically irrespective of any particular incident using processes such as prospective hazard analysis[14 15] or human factors engineering, whereby experienced people use their knowledge and imagination to work out how human characteristics and interactions between humans and the tools used in the healthcare system may generate risks to patient safety.[16]

Stage 3

In the next phase an intervention is designed and described.[17] The design should be based on theory generated from basic science in subjects such as psychology, sociology and ergonomics. An intervention may be multifaceted, having more than one component. Anaesthesia, for example, has become safe as a result of a combination of changes, each of which has had a small effect on safety.[18] One approach to the problem of measurement of small effects (discussed in Part 3 of this series) is to evaluate a package of interventions introduced simultaneously—that is, as a "complex intervention". A good example is the package of measures introduced to reduce central line infections in Michigan intensive care units.[19] Such an approach is particularly appropriate if theory suggests that the individual components of the intervention may act synergistically, with no individual component generating particularly high risks or costs.

At this development stage it is important to describe the intervention in detail. Interventions can first be classified according to where, on the causal chain in fig 2, they are targeted in the first instance. Within each domain, the tasks should be described in the order they need to take place.[20] The aim should be to comply with the time-honoured scientific and culinary principle of providing sufficient information to allow others to replicate the process. Some complex interventions comprise standardised processes (eg, development of educational materials) but in local forms (eg, materials developed for the specific education level of the local population).[21] Here, a description of the standardised processes is essential, but this needs to be augmented with detail of local adaptations. This is because, as with other "non-complex" interventions that evolve over time following implementation or that are implemented with varying fidelity across sites, such detail may help to explain the success or otherwise of the intervention (as we describe in Part 3).

Stage 4

The consequences of intervening in a certain way should be modelled in as explicit a way as possible through proactive risk assessment using methods such as failure modes and effect analysis.[14 15] This modelling can involve group discussion (thinking it through, with special reference to possible unintended consequences), formal modelling (with or without probability estimates, value weightings and formal mathematical calculations), simulations (mock-ups of the real world and role play) or any combination of these. One of the outcomes of modelling may be identifying barriers to change, and methods to address these barriers can then be included within the (re-) design of the intervention.[12] Results of this fourth modelling stage can be fed back into the design phase on an iterative basis until an intervention is judged fit for roll out into practice. These iterations may continue once the intervention has been

rolled out in practice, as we discuss in the ensuing articles in this series.

It is often the case that the opportunities for tackling different threats to patient safety exceed the resources available to enable all putative interventions to be implemented. In these circumstances, it will be necessary to prioritise potential interventions and PIE will be an important step in identifying the interventions likely to be most cost-effective (or satisfying an alternative prespecified criterion). Health economic methods, typically used on the demand side to inform decisions of which technologies to deploy, are increasingly used on the supply side to decide what technologies to develop.[22]

EXAMPLE OF A PREIMPLEMENTATION EVALUATION

The development of a method to enhance teamwork can be used to illustrate the above stages. This would start with identifying actual or potential patient safety problems, say on the labour ward. Further work might include indepth studies of current practice: for example ethnographic studies of how different health professionals work in teams and qualitative interviews or focus groups to uncover psychological obstacles to team work. This might confirm problems with teamwork on the labour ward and uncover possible causes, such as confusion of roles. Such studies may show how different professional groups relate to and communicate with each other while undertaking various tasks. A training intervention could be designed in the light of this study. Such an intervention would build on educational and psychological theory—for example, social cognitive theory. The resulting ''solution'' would be carefully described. Modelling would take the full causal chain into account. It could begin by asking stakeholders (managers, clinicians, patients) to comment on the proposed intervention in prototype form. Possible adverse effects would be considered, for example, would the training timetable remove staff from important clinical duties? If the proposed interventions seemed worthwhile, simulations could be created. These simulations could then be analysed by both qualitative and quantitative methods to refine the intervention.

ONTOLOGY AND EPISTEMOLOGY

Before launching into our discussion of study design and measurement in the next two articles in this series we need to do some philosophical ground clearing. This is because our aim of critiquing the methods of study design and measurement in the existing patient safety literature to inform our framework for future work requires us to identify our ontological principles and why we subscribe to these principles. The need for such ground clearing is not exclusive to patient safety, as similar issues would arise in other areas.

Ontology is somewhat loosely described as ''what is considered as truth''. In our opinion the most fundamental distinction within ontology lies between those who believe there is no such thing as an objective truth (eg, relativists) and those who subscribe to the alternative position (eg, positivists). The relativist argument has been made for science as a whole by Feyerabend[23] and in a social science context by Guba and Lincoln.[24] We do not adhere to this philosophy for reasons we and others have articulated elsewhere.[25 26] Moving away from the relativist tradition, there are several alternatives, including positivism and pragmatism.[24] We take a broadly pragmatist position whereby strength of belief accumulates in line with salient evidence: an idea on which we elaborate in Part 4 of the series.

The next source of potential confusion we tackle lies in the distinction between objectivity and subjectivity in science. The idea has grown that objectivity and subjectivity are two distinct epistemologies (ways of getting at the truth). Friedman and Wyatt,[27] for example, write of objectivist and subjectivist approaches to evaluation, as though one or the other must be selected. We reject this dichotomy on the grounds that all scientific interpretation (the derivation of all scientific meanings) is subjective. When Copernicus interpreted planetary observations as evidence for a heliocentric solar system he was attributing meaning to data. When Rutherford concluded that matter was simply made up of space between atoms he was interpreting his famous objective finding that most γ rays pass through a very thin sheet of gold. We therefore adhere to the premise that empirical information (however obtained) inevitably requires subjective interpretation.

It is also useful to make clear the distinction between objective and subjective entities. The entity under study may or may not have an existence independent of human feeling, experience or thought: in the words of John Searle[25] it may be ontologically objective or subjective. But even if the entity is ontologically subjective it can be quantified. Therefore money (which is socially constructed) and pain (which is experienced in a personal, subjective way) can nevertheless be measured.

Lastly we deal with the issue of what is sometimes referred to as quantitative and qualitative research. While primary data may be quantitative or qualitative, this does not allow the research itself to be classified unproblematically into either quantitative or qualitative categories. We have noted that all scientific observations require subjective (and hence qualitative) interpretation in order to acquire scientific meaning. Moreover, an inductive step is always required in deciding whether or to what extent quantitative observations would apply in another place and time, or to what extent biases in the data collection process have muddied the results. Such a step is subjective and usually expressed in qualitative terms. However, qualitative data can be transformed into quantitative data (for example calculating the percentage of respondents in interviews giving a particular response) and, as we will see below, can be used to estimate a parameter.[26 28 29]

Rather than classify research as objectivist/subjectivist or quantitative/qualitative, we therefore prefer to acknowledge that subjective interpretations are always required and to consider the following questions.

- Is the construct being examined ontologically and epistemologically objective (eg, death), ontologically subjective but epistemologically objective (eg, costs) or ontologically and epistemologically subjective (eg, pain)?
- Are the primary data collected in numerical (quantitative) or ''open'' (qualitative) format?
- Are the data analysed quantitatively or quantitatively (or both)?

These considerations will be used to frame our arguments relating to study design in the subsequent articles in this series: for example, in identifying the potential biases arising from quantitative and qualitative methods of data collection. In Part 4 in this series, we will argue that a bayesian approach, which provides a statistical model for updating a belief on the effects of an intervention as more data (both quantitative and qualitative) come to light,[26] may be an appropriate approach in triangulation of findings based on different types of data and/or different studies altogether. Whether or not a bayesian approach is used,

collecting both quantitative and qualitative data is advocated throughout this series of articles.

CONCLUSION

Safety interventions are notoriously prone to back fire (for example alcohol-based handrubs placed on wards have been stolen and consumed by alcoholic patients). For this reason, proposed interventions to improve safety should all be screened through a systematic process of PIE. The idea is to reduce, but of course not to eliminate, the risks that interventions will not work well or will introduce important new hazards. Modelling how an intervention may impact on safety/quality requires an understanding of the causal chain through which an effect may be produced. In Part 3 of this series we will show that this chain is important for modelling and as a guide to study end points. In Part 4 we bring the various strands together to show how measurements of end points along the causal chain can contribute to a holistic evaluation through a process of ''triangulation''. This process should illuminate not just whether an intervention works but why it ''works for whom under what circumstances'', as recommended by Pawson and Tilley in their work on realistic evaluation.[30] We will go further and describe how ''triangulation'' can be rendered transparent in a bayesian statistical framework. Lastly, we have drawn a distinction between objectivity and subjectivity. The phenomenon being studied may be objective (life or death) or subjective (pain). In either case data collection may be more or less objective/subjective (ie, independent/dependent of the observer). However the interpretation/extrapolation of the data is necessarily subjective.

Acknowledgements: We would like to acknowledge the support of the National Coordinating Centre for Research Methodology and the Patient Safety Research Programme. The authors would also like to acknowledge the contributions of attendees at the Network meetings and the helpful comments of the peer reviewers of this article.

Competing interests: None.

Authors' contributions: RL conceived the Network and formulated the first draft of the report and the current paper with assistance from AJ. CB contributed to subsequent drafts of the report and this paper. BDF, TH, RT and JN contributed to the Research Network and provided comments on drafts of the report and papers in their areas of expertise.

This work forms part of the output of a Cross-Council Research Network in Patient Safety Research funded by the Medical Research Council (Reference G0300370). More details of the Research Network can be found at: http://www.pcpoh.bham.ac.uk/publichealth/psrp/MRC.htm

REFERENCES

1. **Lilford RJ,** Stirling S, Maillard N. Citation classics in patient safety research: an invitation to contribute to our online bibliography. *Qual Saf Health Care* 2006;**15**:311–13.
2. **Brown CA,** Patil P, Hofer TP, *et al*. An epistemology of patient safety research: A framework for study design and implementation. http://pcpoh.bham.ac.uk/publichealth/psrp/MRC.htm (accessed 1 Apr 2008).
3. **Reason J.** Human error: Models and management. *BMJ* 2000;**320**:768–70.
4. **Shojania KG,** Duncan BW, McDonald LM, *et al*. Making health care safer: A critical analysis of patient safety practices. AHRQ publication 01-E058. Rockville, MD: Agency for Healthcare Research and Quality, 2001.
5. **Hofer TP,** Kerr EA, Hayward RA. What is an error? *Eff Clin Pract* 2000;**3**:261–9.
6. **McQuillan P,** Pilkington S, Allan A, *et al*. Confidential inquiry into quality of care before admission to intensive care. *BMJ* 1998;**316**:1853–8.
7. **National Patient Safety Agency**. *Safer care for the acutely ill patient: learning from serious incidents*. The fifth report of the Patient Safety Observatory. London: National Patient Safety Agency, 2007.
8. **National Patient Safety Agency**. *Slips, trips and falls in hospital*. The third report of the Patient Safety Observatory. London: National Patient Safety Agency, 2007.
9. **Donabedian A.** Explorations in quality assessment and monitoring. In: Griffith JR, ed. *The definition of quality and approaches to its assessment*. Ann Arbor Michigan: Health Administration Press, 1980:4–163.
10. **Frankel A,** Graydon-Baker E, Neppl C, *et al*. Patient safety leadership walkrounds. *Jt Comm J Qual Saf* 2003;**29**:16–26.
11. **Medical Research Council**. A framework for development and evaluation of RCTs for complex interventions to improve health. London: Medical Research Council, 2000.
12. **Campbell NC,** Murray E, Darbyshire J, *et al*. Designing and evaluating complex interventions to improve health care. *BMJ* 2007;**334**:455–9.
13. **National Patient Safety Agency**. *Seven steps to patient safety*. London: National Patient Safety Agency, 2004.
14. **Marx DA,** Slonim AD. Assessing patient safety risk before the injury occurs: An introduction to sociotechnical probabilistic risk modelling in health care. *Qual Saf Health Care* 2003;**12**:ii33–ii38.
15. **Vincent C,** Taylor-Adams S, Stanhope N. Framework for analysing risk and safety in clinical medicine. *BMJ* 1998;**316**:1154–7.
16. **Gosbee J.** Human factors engineering and patient safety. *Qual Saf Health Care* 2002;**11**:352–4.
17. **Campbell M,** Fitzpatrick R, Haines A, *et al*. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;**321**:694–6.
18. **Leape LL,** Berwick DM, Bates DW. What practices will most improve safety? Evidence-based medicine meets patient safety. *JAMA* 2002;**288**:501–7.
19. **Pronovost P,** Needham D, Berebholtz S, *et al*. An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med* 2006;**355**:2725–32.
20. **Czaja SJ,** Schultz R, Lee CC, *et al*. A methodology for describing and decomposing complex psychosocial and behavioural interventions. *Psychol Aging* 2003;**18**:385–95.
21. **Hawe P,** Shiell A, Riley T. Complex interventions: how ''out of control'' can a randomised controlled trial be? *BMJ* 2004;**328**:1561–3.
22. **Girling AJ,** Freeman G, Gordon JP, *et al*. Modeling payback from research into the efficacy of left-ventricular assist devices as destination therapy. *Int J Technol Assess Health Care* 2007;**23**:267–77.
23. **Feyerabend PK.** *Realism, rationalism and scientific method*. Philosophical papers, volume 1. Cambridge: Cambridge University Press, 1985.
24. **Guba E,** Lincoln Y. Competing paradigms in qualitative research. In: Denzin N, Lincoln Y, eds. *Handbook of qualitative research*. Thousand Oaks, CA: Sage, 1994:105–77.
25. **Searle J.** The construction of social reality. New York: Free Press, 1995.
26. **Lilford RJ,** Braunholtz D. Reconciling the quantitative and qualitative traditions: the bayesian approach. *Public Money and Management* 2003;**23**:203–8.
27. **Friedman CP,** Wyatt JC. *Evaluation methods in medical informatics*. New York: Springer, 1997.
28. **Roberts KA,** Dixon-Woods M, Fitzpatrick R, *et al*. Factors affecting uptake of childhood immunisation: a bayesian synthesis of qualitative and quantitative evidence. *Lancet* 2002;**360**:1596–9.
29. **Lilford RJ,** Braunholtz D. The statistical basis on public policy: a paradigm shift is overdue. *BMJ* 1996;**313**:603–7.
30. **Pawson R,** Tilley N. *Realistic evaluation*. London: Sage, 1997.