Improving the quality of health care

Using research to inform quality programmes

John Øvretveit, David Gustafson

Quality programmes consume more resources than any treatment and have potentially greater consequences for patient safety and other clinical outcomes. So why do we know so little about whether they are effective?

Health resources that could be used for clinical care are increasingly being devoted to large scale programmes to improve the quality of health care. Examples include national quality initiatives, hospital programmes, and quality accreditation, assessment, and review processes. However, little research has been done into their effectiveness or the conditions needed to implement quality programmes successfully. This is partly because the programmes are difficult to evaluate: they change over time, are applied to changing organisations, and need to be assessed from different perspectives. However, research can produce valid and useful knowledge about how to make such programmes work. We describe what research has shown us so far and highlight how better information can be obtained.

Summary points

- Quality programmes are large scale interventions to improve health care
- Little research is available to show if they work or are cost effective
- Such research is difficult because the programmes involve dynamic organisations and change over time
- Research can identify the factors needed for successful implementation

What is a quality programme?

Quality programmes are planned activities carried out by an organisation or health system to prove and improve the quality of health care. The programmes cover a range of interventions that are more complex than a project carried out by a single team (box 1).

Use of quality programmes is increasing worldwide. One recent study noted 11 different types of quality programmes in the NHS over three years.[1] Many countries are embarking on accreditation programmes without any evidence that they are the best use of resources for improving quality and no evidence about the effectiveness of different systems and ways to implement them.[2] Nevertheless, research into some types of programme has produced useful information for decision makers.

Research into quality improvement programmes

Total quality management in hospitals

Most research has been done into hospital quality programmes, particularly total quality management programmes in the United States (now called continuous quality improvement programmes). Unsystematic reviews of the research show that few healthcare organisations have successfully implemented a quality programme.[3–7] However, the evidence provided by the studies is limited. Little is known about long term results or whether the programmes have been sustained. Few studies describe or compare different types of hospital quality programmes. Many studies rely on self reports by quality specialists or senior managers and survey these people once, retrospectively.

Other quality improvement programmes

Few other types of quality improvement programmes have been systematically studied or evaluated. In a study of accreditation, managers reported that organisations that received low scores (probation) on the US Joint Commission for Accreditation of Healthcare Organisations assessment were given high scores three years after but had not made substantive changes.[7] A few studies have described or assessed some of the many quality assessment systems,[8–11] external evaluation processes,[12–17] national and regional quality strategies, or programmes in primary health care.[18] Research is now being done to evaluate quality improvement collaboratives.[19] This research considers the factors critical for success as perceived by different parties.

Clearly, we need more evaluations and other types of studies of quality programmes to answer the questions of decision makers and to build theory about large scale interventions to complex health organisations or health systems. Below, we describe the problems of research and the methods that can be used to provide more information.

Box 1: Types of quality programmes

Quality programmes include programmes for:

- Whole organisations—eg hospital total quality programmes
- Teams from many organisations—collaborative programmes
- External reviews of organisations—eg quality accreditation programmes
- Changing practice in many organisations—eg practice guidelines programmes
- National or regional quality initiatives or strategies (which could include any or all of the above)

This is the first of three articles on research to improve the quality of health care

Medical Management Centre, Karolinska Institute, Stockholm Sweden S-40242
John Øvretveit
director of research

University of Wisconsin, 610 Walnut Street, University of Wisconsin, Madison WI 53705, USA
David Gustafson
Robert Ratner professor of industrial engineering and preventive medicine

Correspondence to:
J Øvretveit
jovret@aol.com

BMJ 2003;326:759–61

Research challenges

Large scale quality programmes are difficult to evaluate using experimental methods. The programmes evolve and include many activities that start and finish at different times. Many programmes are poorly formulated and partially implemented. Most cannot be standardised and need to be tailored to suit the situation in different ways from those used when a treatment is changed to suit a patient. The targets of the interventions are not patients but whole organisations or social groups, which are complex adaptive systems that vary more than the physiology of individual patients.[20]

Another problem is that there are many criteria of success of a programme. Each will usually have short and long term outcomes, and these often need to be studied from the perspectives of different parties. It is also difficult to prove that any change is due to the programme, given their evolving nature, their target, the environment, and the long timescales.[21]

Some people believe that each programme and situation is unique and no generalisations can be made to other programmes. This may be true in some cases, but even a description of the programme and its context allows others to assess the relevance of the programme and the findings to their local situation.

Research designs

The difficulties in evaluating quality programmes do not mean that they cannot or should not be evaluated. The designs described below have been used successfully. Further details are available elsewhere.[21-23]

Descriptive case design

This design simply aims to describe the programme as implemented. There is no attempt to gather data about outcomes, but data are obtained on what knowledgeable stakeholders expect from the programme and their perceptions of the strengths and weaknesses of the programme. The Cochrane Effective Practice and Organisation of Care Group (EPOC) has developed methods for assessing observational studies.

Audit design

The audit design takes a written statement about what people should do, such as a protocol or plan, and compares it with what they actually do. This quick and low cost evaluation is useful when there is evidence that following a programme or protocol will result in certain outcomes. It can be used to describe how far managers and health staff follow prescriptions for quality programmes and why they may diverge from these prescriptions. Audit of quality accreditation or review processes can help managers to develop more cost effective reviews.

Before and after designs

Before and after studies are prospective and may be single case or comparative. The single case design gathers data about the target of the intervention before and after (or during) the intervention. The outcomes are the differences between the before and after data. The immediate target is the organisation and staff, but the ultimate targets are patients.

Comparative before and after designs produce stronger evidence that any changes are due to the programme and not to something else. As with a controlled trial, if the programme is not introduced into the comparison unit, any change seen in the intervention unit is more likely to be due to the programme if the units have similar characteristics and environments.

Box 2: Example of theory building action evaluation comparative design

A comparative study of the quality programmes in six Norwegian hospitals over four years provided evidence about results and factors critical for success.[23] It also provided the first detailed and long term description about what hospitals in a public system actually did and how their programmes changed over time.[5] The common factors critical for success were involvement of management and doctors at all levels, good data systems, the right training, and effective project team management.

Retrospective or concurrent evaluation designs

In these designs, the researcher can use either a quasi-experimental theory testing approach or a theory building approach. An example of a theory testing approach is the prediction testing survey. The researcher studies previous theories or empirical research to identify hypothetical factors that are critical for success (for example, sufficient resources, continuity of management, aspects of culture) and then tests these to find which are associated with successful and unsuccessful programmes.

In a theory building approach, the researcher gathers data about the intervention, context, and possible effects during or after the intervention (box 2). To describe the programme as it was implemented, the researcher asks different informants to describe the activities that were actually undertaken. The validity of these subjective perceptions can be increased by interviewing a cross section of informants, by asking informants for any evidence that would prove or disprove their perceptions, and by comparing data from difference sources to identify patterns in the data.[23]

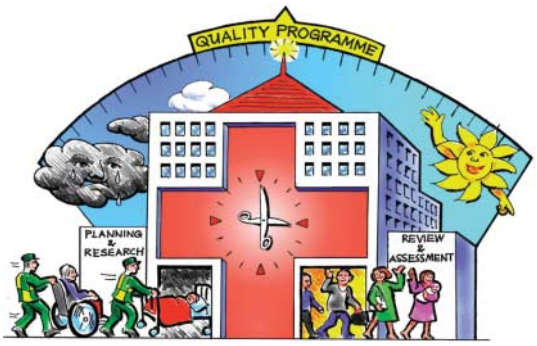
The choice of design depends on the type of quality programme (short or long term, prescribed or flexible, stable or changing?) who the research is for, and the questions to be examined (was it carried out as planned? did it achieve its objectives? what were the outcomes? what explains outcomes or success or failure?).

Improving research into quality programmes

Research into quality programmes could be improved by researchers paying attention to common failures of previous research. These can be categorised as follows.

Implementation assessment failure—The study does not examine the extent to which the programme was actually carried out. Was the intervention implemented fully, in all areas and to the required "depth", and for how long?

Prestudy theory failure—The study does not adequately review previous empirical or theoretical research to make explicit its theoretical framework, questions, or hypotheses.



SUE SHARPLES

Outcome assessment failure—The study does not assess any outcomes or a sufficiently wide range of outcomes such as short and long term impact on the organisation, patients, and resources.

Outcome attribution failure—The study does not establish whether the outcomes can unambiguously be attributed to the intervention

Explanation failure—There is no theory or model that explains how the intervention caused the outcomes and which factors and conditions were critical.

Measurement variability—Different researchers use very different data to describe or measure the quality programme process, structure, and outcome. It is therefore difficult to use the results of one study to question or support another or to build up knowledge systematically.

Conclusions

Although some discrete quality team projects have been shown to be effective, little evidence exists that large scale quality programmes bring important benefits or are worth the cost. However, neither is there conclusive evidence that there are no benefits or that resources are being wasted. Such evidence may never exist: quality programmes are changing multicomponent interventions applied to complex organisations in a changing context with many short and long term outcomes, few of which can unambiguously be attributed to the intervention with the research designs that are possible.

Seeking evidence of effectiveness for evidence based policy is either impossible or premature at this stage. A more realistic and useful research strategy is to describe the programmes and their contexts and discover factors that are critical for successful implementation as judged by different parties. In a relatively short time this will provide useful data for a more research informed management of these programmes.

This is a shorter version of a paper published in *Quality and Safety in Health Care* 2002;11:270-5.

Competing interests: DG has been a speaker on quality improvement at numerous organisations over the past five years and has received speaking fees for those presentations, for example, the Institute for HealthCare Improvements National Forum.

1 West E. Management matters: the link between hospital organisation and quality of patient care. *Qual Health Care* 2001;10:40-8.
2 Shaw C. External assessment of health care. *BMJ* 2001;322:851-4.
3 Bigelow B Arndt, M. Total quality management: field of dreams. *Health Care Manage Rev* 1995;20(4):15-25.
4 Motwani J, Sower V, Brasier L. Implementing TQM in the health care sector. *Health Care Manage Rev* 1996;21(1):73-82.
5 Øvretveit J. The Norwegian approach to integrated quality development. *J Manage Med* 2001;15:125-41.
6 Shortell S, Bennet C, Byck G. Assessing the impact of continuous quality improvement on clinical practice: what will it take to accelerate progress. *Milbank Q* 1998;76:593-624.
7 Blumenthal D, Kilo C. A report card on continuous quality improvement. *Milbank Q* 1998;76:625-48.
8 Thompson R, McElroy H, Kazandjian V. Maryland hospital quality indicator project in the UK. *Qual Health Care* 1997;6:49-55.
9 Cleveland Health Quality Choice Program. Summary report from the Cleveland Health Quality Choice Program. *Qual Manage Health Care* 1995;3(3):78-90.
10 Rosenthal G, Harper D. Cleveland health quality choice. *Joint Commission Journal on Quality Improvement* 1994;2:425-42.
11 Pennsylvania Health Care Cost Containment Council. *Hospital effectiveness report*. Harrisburg: PHCCCC, 1992.
12 National Institute of Standards and Technology. *The Malcum Baldridge national quality award 1990: application guidelines*. Gaithersburg, MD: NIST, 1990.
13 Hertz H, Reimann C Bostwick M. The Malcolm Baldridge national quality award concept: could it help stimulate or accelerate healthcare quality improvement? *Qual Manage Health Care* 1994;2:63-72.
14 European Foundation for Quality Management. *The European quality award 1992*. Brussels: EFQM, 1992.
15 Sweeney J, Heaton C. Interpretations and variations of ISO 9000 in acute health care, *Int J Qual Health Care* 2000;12:203-9.
16 Shaw C. External quality mechanisms for health care: summary of the ExPeRT project on visitatie, accreditation, EFQM and ISO assessment in European Union countries. *Int J Qual Health Care* 2000;12:169-75.
17 Øvretveit J. Quality assessment and comparative indicators in the Nordic countries. *Int J Health Plann Manage* 2001;16:229-41.
18 Wensing M, Grol R. Single and combined strategies for implementing changes in primary care: a literature review. *Int J Qual Health Care* 1994;6:115-32.
19 Øvretveit J. How to run an effective improvement collaborative *Int J Health Care Qual Assurance* 2002;15(5):33-44.
20 Plsek P, Wilson T. Complexity science: complexity, leadership, and management in healthcare organisations. *BMJ* 200;323:746-9.
21 Øvretveit J. Evaluating hospital quality programmes. *Eval* 1997;3:451-68.
22 Cook T, Campbell D. *Quasi-experimentation: design and analysis issues for field settings*. Chicago: Rand McNally, 1979.
23 Øvretveit J. *Action evaluation of health programmes and change: a handbook for a user focused approach*. Oxford: Radcliffe Medical Press, 2002.

(Accepted 6 February 2003)

One hundred years ago

The Nestor of medicine

The oldest member of the medical profession now living is said to be Dr. Jean David of Montpellier, who on February 8th entered on his 103rd year. He was born at Murviel-les-Montpellier on the 19th Pluviôse, Year IX of the Republic one and indivisible, which in the ordinary calendar is February 8th, 1801. He practised his profession for three-quarters of a century at Grabels, a country district near Montpellier, making his daily rounds on horseback. In his youth he is said to have witnessed the march of Wellington and the Peninsular army through Southern France after the battle of Toulouse. His first serious illness—an attack of typhoid fever—occurred when he was 63 years of age. When he was 91 he was attacked by congestion of the lungs, but promptly recovered. Dr. David, who retired from practice a good many years ago, enjoys excellent health, and age has in no way dimmed the brightness of his intelligence. His only infirmity is failing eyesight. On being asked recently by an interviewer to reveal the secret of his length of days, Dr. David replied: "Sobriety in all respects. The human body is a wonderful machine whose organs should never be overtaxed. For my part I continue living much as I have always lived. I am only worried by one thing—the idleness to which failing eyesight has now condemned me." He speaks with enthusiasm of the wonderful progress which he has seen achieved by medicine and surgery in the course of his career.

(*BMJ* 1903;i:511)