Health Services Research – what is it and what has it done for anaesthesia and critical care?*

Gillian Campbell, Andrew Smith
Department of Anaesthesia and Patient Safety Research Unit, Royal Lancaster Infirmary, Lancaster, UK

Tuesday, 14 June 2011 8:30 - 9:15 Room: E108

*This editorial will be published later this year in the European Journal of Anaesthesiology and is based on the Refresher Course lecture with the same title at the European Anaesthesiology Congress, Amsterdam, June 2011. It is reproduced here with the permission of the Editor-in-Chief and the publishers of the European Journal of Anaesthesiology

Health services research is defined as ‘the multidisciplinary field of scientific investigation that underpins improvements in the way that health services are financed, organised, planned and delivered’. Its focus is on patient outcomes (safety, effectiveness and experience) so epidemiology and phase III drug trials are included within the definition whereas basic science studies and phase I and II trials are not. Health services research is therefore an approach that encompasses a wide range of methodologies and techniques. This review aims to provide an overview of these techniques and illustrate their applications with examples from recently published papers in both anaesthesia and critical care.

Evaluation in healthcare

Evaluation has been described as ‘a comparative assessment of the value of an intervention using systematically collected and analysed data, in order to decide how to act’ [1]. This definition may in fact be rather narrow, as it could be interpreted as excluding diagnostic tests, for instance, and some other types of question. However, the fundamental principle of submitting ourselves, our activities and our systems of care to careful scrutiny is essential, for how else can we know that we are acting wisely? Anaesthetists will be most familiar with the comparative evaluation of drugs and techniques using the randomised controlled trial [2] (phase III of the sequence above), and of course this is the best study design to answer questions about the effectiveness of interventions. For this reason, this article will focus on other types of study design and will include observation and interview studies, Delphi type work, evaluation studies and systematic reviews. Epidemiological studies using non-randomised methods (cohort and case-control studies) have potentially greater threats to internal validity and less ability to account for them. The design and conduct of non-randomised and qualitative studies are at least as methodologically challenging as the design and conduct of randomised controlled trials and often less clear-cut. The guiding principle, as the range of questions relevant to the understanding and improvement of health care is wide, is the selection and use a study design and methodology which best allows the specific research question to be answered.

‘Simple’ evaluation

Even those studies that do not use the specific designs mentioned above can shed light on patient care. They also share the problems inherent in all evaluation work, namely the precise description of the model or intervention to be evaluated, the definition of relevant outcomes and the identification and explanation of other factors that might affect observed differences.
These methods are similar to those used in clinical audit, the differences being the use of more rigorous methods, the fact that it is researchers rather than clinical staff who are performing the work, it tends to be more generalisable and maybe larger scale, and evaluation is unlikely to be comparing performance in a practice to an agreed standard. Sometimes, although a randomised approach might be scientifically preferable, an observational study can help to evaluate an intervention in a ‘real world’ setting [3].

These principles are illustrated by two recent publications. ‘Track and trigger’ systems were adopted widely in the United Kingdom in the early 2000s to aid the early identification of deteriorating hospital patients. Scores are assigned to vital signs such as pulse rate, arterial pressure and respiratory rate, the score being higher the greater the deviation from normal physiological values. High scores reach a ‘trigger’ point where a specific corrective action is taken. In a study assessing error rates in scoring [4], the scoring system had just been introduced into the hospital. The researchers performed a retrospective casenote analysis, extracting the scores calculated at the time, recalculating them from the vital signs also recorded, and comparing the two. Scoring errors were found in over 20% of observations, with some patients not reaching the ‘trigger’ score when they should have done. This could have delayed their referral to critical care. However, evaluation of any clinical intervention must take account of the circumstances during which it is used. This study was undertaken during an epidemic of Legionnaires’ Disease and the ‘track and trigger’ system had only recently been introduced. The epidemic caused a substantial increase in workload for staff and so it could be argued that working conditions were not typical. However, other studies in different contexts have found similar error rates in scoring. Further, patients who were tested positive for Legionnaires’ disease tended to be mis-scored more than those who tested negative. This suggests that staff may be prone to subconscious bias, scoring the patient in accordance with their prior knowledge of their disease status.

Another paper from the same team analysed the critical incidents reported to the UK’s National Reporting and Learning Scheme between 2006 and 2008 [5]. In this case, the focus was on developing a methodological approach to retrieving incidents from the database and sharing the lessons for practice more widely than the hospital where they occurred.

Quantitative and qualitative approaches

The studies above involved the analysis of numerical data, namely quantitative research, which is often distinguished from a qualitative approach, which focuses less on measurement and more on exploration, explanation and eliciting meaning. Although qualitative methods are less familiar to anaesthetists, they have been used in a number of studies of relevance to anaesthesia [6]. Principal methods in qualitative research include observation, questionnaires, interviews, focus groups and consensus methods. Analysis proceeds by induction, that is, the themes and findings arise not from some pre-specified hypothesis, but are suggested by the data themselves. In fact, qualitative methods are extremely useful to generate hypotheses for further work, especially in areas where there is little prior knowledge. Clearly, researchers will have some prior knowledge of the subject - there is never a completely ‘blank canvas’ - but should take pains to ensure that this does not overtly colour their view as the work unfolds. One of the strengths of qualitative work is that it builds up a picture of its subject that ‘makes sense’ to those being studied and of course potential bias from the researchers should not be allowed to interfere with this [7]. (Note that the problem of finding what one sets out to find is not unique to any one research approach. Precisely the same caveats apply in quantitative work, although randomisation and blinding can help reduce bias in the results, whether subconscious or deliberate). Qualitative methods provide an insight and rounded view of the various aspects of human behaviour. It is important in reporting qualitative research that the methods are clearly explained and that researchers demonstrate their awareness of possible limitations [8]. ‘Triangulation’ is the term that describes the use of more than one research method to increase the range of information and improve the validity of results [6].

Observation and interview studies

In contrast to the ‘observational study’ referred to above [2] (where data items for collection and outcomes of interest were pre-specified), qualitative research can involve observation and transcription, with detailed analysis of talk and behaviour. This is illustrated by a study examining handovers between anaesthetists and recovery nurses in a post-anaesthesia care unit [9]. The researcher observed 45 such handovers, taking notes by hand for later transcription and analysis. It would have been possible to create a structured tool - for instance, by using a checklist of items to guide the observer as to what to look for - but the research team opted instead for a freer,
open-ended approach. The advantages of this method are that a picture of actual practice can be drawn, rather than relying on interview data, for instance, where respondents may state that they do something different from what they actually do in practice. Disadvantages of observational studies include observer bias, where the observations are inadvertently influenced by the personal views of the observer. Studies may also be subject to the Hawthorne effect (or the ‘effect on the behaviour of the subjects of the watcher watching’) that may again influence the validity of the results [6]. It is therefore essential that in any observational study attempts be made in the methodology to minimise these potential limitations. The cornerstone of a good qualitative study is the process of triangulation where the initial results are further validated using a different stream of data. In this observational study of handovers, structured interviews were also performed, not to ask participants how they handed over, but to explore their perceptions of the process, often drawing on events from observations already made [9]. Interviews also have disadvantages - they are expensive, time-consuming and subject to interviewer bias. Prejudices and prior assumptions can easily influence how the data collection is conducted, and reflexivity - the process whereby the researcher attempts to bring to conscious attention possible biases he or she may have to keep bias to a minimum - is part of the qualitative researcher’s training. Careful phrasing of prompting questions can also help, as can a good rapport between interviewer and respondent. Where patients are being interviewed, it is preferable to have someone who is clearly independent of the hospital staff, as patients may be guarded in their answers if they feel that they may affect the quality of care they receive from staff.

**Questionnaires and surveys**

Questionnaires are structured surveys containing open and/or closed questions to collect data for coding and analysis. A recent study assessing psychological burden in patients with chronic obstructive pulmonary disease and their relatives following admission to intensive care [10] shows good use of different types of questionnaire. The researchers used of the validated French version of the Hospital Anxiety and Depression Scale (HADS) and the Peritraumatic Dissociative Experiences Questionnaire (PDEQ), both of which are examples of quantitative questionnaires with pre-coded response choices that are assigned numerical values. Quantitative patient questionnaires have advantages as they produce numerical data that are easy to count and analyse, they can be administered to large numbers of people and they are relatively economical. There are, however, some disadvantages in that the pre-coded responses may not be sufficiently comprehensive and subtlety may be lost if answers are ascribed to a code which may not fully reflect the initial response. Also, it is assumed the questionnaire is worded and ordered in such a way that patients will understand and give accurate responses. Recall questionnaires are subject to both recall bias and framing where patients’ responses are influenced by the design of the pre-coded response. There may also be a ‘social desirability’ bias, where respondents answer to portray the most positive image of themselves and their behaviour.

Questionnaires are not just used for patients but also to survey clinicians and their clinical practice. A German study examined the use of tracheostomy on intensive care using a postal survey [11]. The authors achieved a good response rate (89%) by enclosing a stamped addressed envelope for the return of the completed questionnaire initially, sending a second questionnaire to units that did not respond to the first, then following up persistent non-responders by telephone. This achieved a high response rate, but generally a poor response to surveys is common, though this varies according to the contact method used. Postal questionnaires have the greatest average response rate (72%), then electronic mail (34%), followed by internet-based surveys (19%) [12]. Non-respondents can be followed up by telephone, thus improving the response rate [12, 13]. However, electronic mail and internet-based surveys are easy, quick and inexpensive and may be useful especially for pilot studies.

There are still limitations to such surveys particularly as the data are reliant upon the self-reporting of physicians. Representativeness is a potential problem in that enquiries about practice may receive more, or less optimistic answers depending on who answers them [14] and there may be something different about responders and non-responders? The length of the questionnaire must also be considered, as it is a key determinant of respondents’ willingness to fill in the form. Finally, design must always be a compromise between ease of answering and usefulness of information obtained. Simple Yes/No questions are easy to answer but will not provide the detailed understanding of longer free text narratives [14].
Consensus methods

Consensus methods aim to determine the extent of agreement amongst experts. These may be most familiar to anaesthetists in ‘consensus statements’ on aspects of clinical care [15].

Consensus can be obtained through face-to-face meetings of a group of experts but this can be expensive as long-distance travel may be needed, and may be limited by lack of common language and shortage of time. The Delphi technique, on the other hand, allows experts’ views to be gathered singly and remotely, by postal questionnaire, electronic mail or telephone. The responses to this first enquiry are then summarised and sent back to the respondents asking them to comment further on the collated material. Commonly, the experts are asked to rank these initial items in order of importance. The results of this second round can be fed back once more for further refinement and the degree of agreement can be assessed formally using statistical techniques.

The Delphi technique is a useful tool in allowing consensus to be obtained without necessarily bringing the experts together to discuss their views. It means that people are less likely to feel inhibited and are more willing to express their views and that discussions are not dominated by one person. Summarising and re-ranking statements improves the accuracy of the original enquiry. A recent study attempting to define excellence in anaesthetic practice made use of a modified Delphi technique [16]. Anaesthetists with an interest in education were invited by electronic mail to list qualities associated with excellence in anaesthesia. However, rather than feeding the responses back to individuals, respondents were invited to take part in focused group interviews - unstructured interviews with small groups of people and a group leader, to rank the items from the first round. This technique is used frequently in business and advertising but also has its place in health services research. Focus groups provide a useful tool to allow freedom to discuss topics in depth, air personal views and exchange ideas with others.

Using both Delphi technique and focus groups provided a structured approach to a difficult question and the combination of the two methods allowed more in depth discussion to contextualise the information and help design modifications to the anaesthetic training curriculum.

Systematic review

Within the body of scientific literature there are often many studies on the same topic. These often differ only slightly in methodology but may give different answers to the same clinical question. A systematic review is a detailed structural analysis, using reproducible methods, of previously conducted primary research. It aims to summarise and critically appraise large bodies of evidence to help bring together and, if necessary, explain differences amongst studies on the same clinical question. A meta-analysis is the mathematical technique that combines the data from two or more methodologically similar studies. [17, 18].

A systematic review differs from the traditional or ‘narrative’ review article [19] in that authors are working to a predetermined protocol in the same way as researchers conducting a randomised clinical study. Setting out how studies will be selected, graded for methodological quality, combined and interpreted should reduce bias in the results of the final product. A recent review of perphenazine for the prevention of postoperative nausea and vomiting illustrates the principles [20]. Strict inclusion criteria were applied, relevant outcomes defined and the search strategy, data extraction and statistical methods all decided in advance of starting work. Eleven trials including over 2 000 patients were included in the meta-analysis, which suggested that perphenazine is effective in both adults and children.

Systematic reviews, with or without meta-analysis, are excellent ways of summarising existing research. However, their results can only be as robust as the quality of the primary trials, and the care with which they are combined, will allow. Reviews can potentially miss relevant trials or be misrepresented as a result of publication bias (the tendency for trials reporting positive findings to be published more readily than those with apparently negative results). When performed well, systematic reviews including all relevant published and unpublished literature can provide statistically significant results with detailed subgroup analysis [18].
Conclusion

Anaesthesia and critical care are specialties that are advancing rapidly and have a number of different facets to them, with the human element integrating closely with science. It follows then that each type of study and methodology is needed for its particular contribution to the scientific advancement of our profession and better care for our patients.

Key learning points

- The randomised controlled trial is the most appropriate methodology for answering questions about the effectiveness of interventions, but many other research methodologies can illuminate other aspects of anaesthesia practice.
- Qualitative methods involve the analysis of words rather than numbers and typically balance data from a range of sources (“triangulation”) to increase the richness and validity of the findings.
- Surveys tend to have higher response rates if an initial postal approach is followed up by telephone and if the questionnaires are short and simple to complete.
- Systematic reviews differ from traditional ‘narrative’ reviews by pre-specifying key methodological stages such as inclusion criteria, data extraction, quality assessment of papers and statistical combination (meta-analysis).

References

7. Smith AF. Reaching the parts that are hard to reach: expanding the scope of professional education in anaesthesia. British Journal of Anaesthesia 2007; 99: 453-6.
18. Greenhalgh T. How to read a paper: papers that summarise other papers (systematic reviews and meta-analyses) British Medical Journal 1997; 315: 672-5.
