Special Article

Technology assessment of anaesthesia monitors: problems and future directions

Robert J. Byrick MD FRCPC,
Marsha M. Cohen MD MHSc FRCPC

Specific factors have limited the interpretation of studies regarding the efficacy, effectiveness and efficiency of technology in anaesthesia. Some of these problems are reviewed, including the lack of specific outcomes necessitating the use of intermediate measures (e.g., hypoxaemia, myocardial ischaemia), which are not necessarily related to ultimate patient outcomes. This emphasizes the need for anaesthesia investigators to define fundamental issues specifically and design studies accordingly. With respect to anaesthesia monitors, the "lead time" or early warning provided by a monitor relative to that required to alter therapy effectively needs to be defined better and compared with the "lead time" without the monitor. After defining the benefit of a monitor, investigators should analyze the cost relative to alternatives (cost-benefit and cost-effectiveness). A hierarchical model to guide technology assessment is presented that addresses in order, the scientific basis of the technology, and the influence on the patient followed by societal issues. Anaesthetists have relied on traditional methods of technology assessment adopted from other disciplines. These methodologies do not address specific issues related to anaesthesia practice (such as "lead time"). In defining problems specific to the specialty of anaesthesia, new outcome measures that focus on the human factors related to decision-making in the operating room need to be developed. Future evaluations of anaesthesia technology require innovative approaches that address specific anaesthesia-related problems. One such approach is the use of simulation-based studies of response patterns to critical incidents.

En anesthésie, certains facteurs spécifiques limitent l'interprétation des études portant sur l'efficacité, le rendement et la productivité de la technologie. Certains de ces problèmes sont passés en revue, en particulier l'absence de résultats spécifiques nécessitant l'utilisation de mesures intermédiaires (par ex., l'hypoxémie, l'ischémie myocardique) qui ne sont pas en rapport direct avec le devenir ultime des patients. Ceci souligne pour l'anesthésie le besoin de chercheurs orientés vers la définition des problèmes fondamentaux et l'élaboration subséquente d'études appropriées. Au regard des moniteurs, le délai de réaction (ou la rapidité de l'alarme produite) d'un moniteur comparé au temps requis pour modifier efficacement un traitement doit être mieux défini et comparé au délai de réaction enregistré sans moniteur. Après avoir défini le bienfait d'un moniteur, les chercheurs doivent en analyser le coût relatif avec ses alternatives (coût-bénéfice et coût-efficacité). Un modèle hiérarchique servant à guider l'évaluation de la technologie est présenté dans l'ordre suivant: la base scientifique de la technologie et son importance sur le patient suivies par des problèmes sociétaux. Traditionnellement, les anesthésistes ont eu recours aux méthodes d'évaluation technologique adoptées par d'autres disciplines. Cette méthodologie ne s'adresse pas aux problèmes spécifiques à la discipline d'anesthésie (comme le délai de réaction). En définissant les problèmes spécifiques à l'anesthésie, il faut développer des nouvelles mesures de résultats qui mettront l'accent sur les facteurs humains qui influencent la prise de décision en salle d'opération. A l'avenir, les évaluations de la technologie nécessiteront des approches innovatrices qui s'adresseront aux problèmes spécifiques à l'anesthésie. Une telle approche comprend l'utilisation de la simulation des types de réponses aux incidents graves.

Key words
MONITORING.

From the Department of Anaesthesia, St. Michael's Hospital and the Clinical Epidemiology Unit, Sunnybrook Health Science Centre, University of Toronto, Toronto Ontario.

Dr. Cohen is supported by a National Health Scholar Award from Health Canada.

Address correspondence to: Dr. R.J. Byrick, Department of Anaesthesia, University of Toronto, Room 131, FitzGerald Building, 150 College Street, Toronto, Ontario M5S 1A8.

Accepted for publication 15th October, 1994.

In today's climate of fiscal constraint, anaesthetists are increasingly being asked to justify the use and purchase

of new intraoperative monitoring devices. While some monitors, such as pulse oximeters, have relatively low unit costs, equipping all operating rooms, postanaesthesia care units and even wards with this technology results in a large expense. The anaesthetist-in-chief at a hospital must determine priorities based on fixed budgets. Can the department afford both the latest technological advance and promising new drug(s)? Anaesthetists are increasingly expected to show justification and cost-effectiveness before the acquisition of newer technologies.[1]

Terminology

While technology assessment can refer to any procedure, drug, device or process of care (preventive, diagnostic, rehabilitative), what is of particular interest among anaesthetists is the newer intraoperative monitors, which will be the main focus of this paper.

Technology assessment has been defined as the process of designing and conducting investigations which enable a reasoned judgment regarding the efficacy, the effectiveness and the efficiency of a monitor.[2] Efficacy has been referred to as the probability of benefit to individuals from the use of technology applied to a given medical problem under ideal conditions of use.[3] The key terms in this definition are "benefit to individuals" and "ideal conditions of use." Efficacy studies are generally carried out at tertiary care hospitals, by academic physicians, with strict protocols, and using a set of narrow inclusion criteria to a well-defined group of patients, say white men aged 50–65 yr. The study patients (who may not be representative of all patients with the condition) are carefully monitored by trained designated research staff and compliance is usually ensured by careful follow-up of patients.

In contrast, effectiveness refers to patient benefit among the general population, when the technology is used in routine, clinical practice by clinicians who are not academic or in large centres. While an efficacy study of a particular technology may show a positive benefit among a select group of patients, when used by community-based physicians on a wider group of patients, for example, women aged >65 yr, there may not be a benefit. There may be a variety of reasons to account for the differing results. The technology may only be of benefit to younger men perhaps due to physiological differences between men and women. The technology may not have been used or interpreted in a proper fashion; the settings may be incorrectly calibrated or the machine may not have been maintained regularly. The indications for the technology may be different from or widened from the efficacy studies or may not be applicable to the new indication. Alternatively, compliance with the technology may be poor as there is no systematic follow-up to its use (e.g., used for some patients but not others in a haphazard fashion.) Appropriate applications of a new technology in widespread practice may be delayed if practitioners do not receive feedback from efficacy studies on earlier prototypes. Thus, any new technology must be assessed in the real world situation, as well as at academic centres, to determine if there is any true net benefit to individuals for general medical use.

Outcome assessment

Most clinical anaesthetists intuitively believe that monitors have contributed to a reduction of perioperative morbidity and mortality. Yet proving this clinical impression with traditional techniques of epidemiological and scientific investigation has been difficult.[4,5] For example, with reference to the improvement in patient outcome attributable to pulse oximeters, Orkin demonstrated that the decrease in anaesthesia-related mortality began decades before the advent of these monitors[6] with most of the decline occurring before their introduction. Nonetheless, studies have attempted to evaluate monitors by using ingenious methodologies, for example, using the pulse oximeter but not telling the anaesthetist unless there was a problem.[7]

Two fundamental problems have been recognized in the conduct and design of studies to evaluate monitors. The first problem relates to the absence of a specific measurable outcome for monitor-related problems as a result of a lack of a clear definition of the clinical issue studied. This emphasizes a need for investigators of anaesthesia technology to define clearly the fundamental issue studied and to design the intervention to address this clinical problem specifically. Some examples will help to illustrate this point.

Consider death from anaesthesia as the main outcome measure of an assessment of a new technology. One would think that this is relatively straightforward, as deaths are easy to count. However, there are a number of problems. First, the definition of death is more complex than it used to be (e.g., by death do we mean brain death?). Secondly, how much later, after the surgical procedure, would a death be convincingly attributable to what the anaesthetist has done in the operating room? Third, what is the definition of anaesthetic-related mortality?[8] In the 1990s, there are so many other factors contributing to mortality that the anaesthetic component is extremely difficult to evaluate clearly.[9,10] Even if these considerations could be sorted out to everyone's satisfaction, there is still an obstacle in that death attributable to anaesthesia is extremely rare. Studies would be hopelessly complicated by the large patient groups required. For example, to detect a 20% improvement in anaesthesia mortality (a current estimate is 0.056 per 10,000 anaes-

thetics),[11] one would require more than 250,000 patients to show an effect of preventative technology. Even if one could show this effect, while the relative improvement may be interesting, the absolute difference is clinically meaningless. Such studies will not likely be completed.

The second problem relates to the use of intermediate outcomes which others have suggested be used instead of major outcomes such as death. Intermediate outcomes include hypoxaemia, ischaemia, and so forth, and have several advantages in that they can be clearly defined, occur more frequently, and can be more directly attributable in time to what the anaesthetist does. However, the use of intermediate outcomes is also problematic. There are two interrelated concerns about the use of intermediate outcomes. First, the relationship between intermediate outcomes and true outcomes (e.g., major morbidity and death) is not necessarily a direct or linear relationship. There may be many false alarms where the actual physiological value of the measurement falls below the threshold value, but the patient is not really in any danger. After all, we do not know what the age and sex-specific normal values are for pulse oximetry. There may be many individuals walking around with clearly abnormal values. The second relates to "lead time" or early warning provided by the monitor compared with that in routine clinical practice without the monitor or with alternative techniques.[4] The clinical importance of detecting a physiological event depends upon whether early detection results in the opportunity to apply a therapy early enough to be effective. If earlier detection using a monitor results in only a very brief increase in the early warning system and the clinician cannot institute effective therapy within this timeframe, the lead time of the monitor cannot be expected to effect an improved outcome. As well, if there is no therapy available to influence outcome, then the patient will not benefit even if the monitor does detect the physiological abnormality well in advance. In this circumstance, as in very brief lead time, there is no advantage in having the monitor (i.e., no benefit to patients). The challenge to the investigator is to determine how to demonstrate a benefit and then estimate the cost relative to any monitoring alternatives (cost-benefit or cost-effectiveness).

Pulse oximetry

If surveyed, most anaesthetists would agree that pulse oximetry has improved patient outcome. The Canadian Anaesthetists' Society (among others) has adopted guidelines[12] that mandated the routine use of oximetry without evidence from randomized clinical trials. Most anaesthetists believed then, and continue to believe, that oximetry prevents hypoxic injuries.

In 1993, Moller *et al.* from Denmark,[13,14] conducted what is likely to be the only large randomized evaluation of pulse oximetry in which over 20,000 patients were followed. Importantly, these investigators focused on the relationship between the detection of intraoperative and post-anaesthesia care unit (PACU) events by oximetry and postoperative complications. As expected, the incidence of intermediate outcomes such as respiratory events, including desaturation, was greater in the oximetry-monitored group than in the control (non-oximetry monitored) group. Yet, postoperative complications, such as pneumonia, atelectasis, myocardial infarction, were not different. Does this imply that pulse oximetry is not effective? The answer is that this study can neither confirm nor deny the effectiveness of pulse oximetry. Due to the relatively small sample size and other methodological concerns, one cannot rule out that pulse oximetry may be effective.

Technology assessment

We may have reached the point in the assessment of monitoring technology in anaesthesia where traditional approaches, such as the randomized controlled trial (RCT), useful as they have been, are too restrictive. Other techniques of technology assessment such as literature reviews, meta-analysis, consensus panels, cost-effectiveness analyses, database registries have been used in internal medicine, critical care and radiology to assess technology. Some of these techniques may also be applicable to anaesthesia. However, in anaesthesia, it appears that we need to adopt techniques that specifically address the issue of the warning function of our monitors. These studies should focus on the human factor, how we respond to warnings and the lead time factor. We need to define problems specific to the specialty of anaesthesia objectively and develop innovative outcome measures related to these problems.

An outline, or template, for technology assessment has been developed and used by radiologists[15] and this may be useful for the evaluation of anaesthesia technology. There are some similarities to the use of screening radiology such as mammograms and the use of anaesthesia monitors in that both serve an early warning function. Both have a technical component related to the measurement aspect of the technology. For example, the radiogram must be calibrated for exact penetration of the *x*-ray beam to maximize the image and minimize the "background," whereas the anaesthesia monitor must be set to maximize true physiological deviations and minimize false alarms.

The template proposed by Fryback and Thornbury[15] can be used to classify assessment studies. The model goes beyond the usual goal of providing "the best images and the best diagnosis" in radiology to be part of a larger

system, aimed at treating patients effectively and efficiently. The template is a "hierarchical model" as it addresses the scientific basis of the technology itself, then the influence on the patient, followed by societal issues. For a monitor to be efficacious at a higher level, it must be efficacious at a lower level, but not necessarily vice versa. We have modified this template for the potential assessment of anaesthesia technology as follows:

Basic science of the technology

Anaesthetists have been among medicine's most conscientious investigators in understanding the fundamental principles of applied monitoring technology. At this stage, a monitor is conceived, the prototypes developed, and the patents obtained. This is followed by an assessment of the technical aspects including the physical characteristics of the monitor, its ability to measure physiological variables, the various adjustments and sensitivities required for the technical functioning of the monitor. This assessment permits comparison with other techniques based on these characteristics. This assessment is usually carried out in the laboratory, often using animal models. However, while a monitor may be useful in a research laboratory, one cannot assume that it will be of ultimate benefit to patients in the clinical setting.

Site/indications for use

In anaesthesia practice, specific indications for monitoring have been rigorously identified in clinical trials for the use of some monitors. However, in practice, we have often extended the use of these monitors to other circumstances and other locations (e.g., pulse oximetry was developed for the operating room but now is used in the PACU and on the ward) without documented effectiveness of its use in these settings, a so-called "technology creep."[16] Thus, both the proposed site and the specific indications for use need to be rigorously delineated.

Efficacy

Fryback and Thornbury[15] discussed six levels of efficacy:

1 TECHNICAL EFFICACY

A monitor must be able to measure physiological variables in a reliable fashion under ideal conditions in clinical practice.

2 DIAGNOSTIC EFFICACY

This can be defined in various ways such as the number of abnormalities found, or predictive values such as sensitivity and specificity. However, as in other disciplines, establishing the "gold standard" for comparison is often not readily clear. This needs to be done in clinical practice under the actual circumstances of use. Thus, the characteristics of the monitor in practice that are not evaluated in the developmental phase can be assessed. For example, diagnostic efficacy of a monitor may be reduced in the operating room because of electromagnetic interference, which may not have been detected during development. We cannot assume that monitors will be used optimally in clinical practice, that the device will be properly applied, or that the information will be optimally interpreted in a timely manner. Even at this stage, one cannot assume that there will be benefit to patients.

3/4 DIAGNOSTIC THINKING EFFICACY AND THERAPEUTIC EFFICACY

These concepts are closely related so they will be discussed together. Diagnostic thinking efficacy refers to the change in the clinician's thought processes because of the information provided by the monitor. Therapeutic efficacy, on the other hand, refers to the impact of the information on implementing effective therapy. Patients' outcomes cannot be affected unless the anaesthetist is willing to act differently based on the information provided by the monitor. The information provided may provoke the anaesthetist to change a specific drug, delay extubation, or merely to continue on the present course (be reassured that the patient is not in danger). To illustrate this, a less celebrated outcome of the Danish pulse oximetry study[14] was that although there were no measurable differences in patient outcomes, anaesthetists in the study felt more reassured by having pulse oximeters. This "thinking and acting" process that anaesthetists use can be assessed methodologically for example by using patient scenarios in questionnaires. However, what anaesthetists say they would do is not necessarily what they actually do. Even what they actually do may not affect the ultimate patient outcome, so that again, patient benefit cannot be assessed at this level.

5 PATIENT OUTCOME

As noted above, the ultimate aim of anaesthesia care is benefit to the patient so that the determination of appropriate outcome measures is critical. It has been argued that the field of anaesthesia needs to broaden its concept of what is an outcome measure.

Anaesthetists are familiar with mortality or major morbidity as true endpoints and hypoxaemia, hypotension, hypothermia as intermediate outcomes. However, returning to the theme of "benefit to patients," newer concepts of outcome measures need to be considered. These include quality of life, psychological well-being, and time to return to usual activities.[4,17] Expansion of the "traditional" anaesthesia research methods is also needed to include cost-effectiveness, decision analysis, and patient preferences.

6 SOCIETAL EFFICACY

Sometimes there is a conflict between what is the "best for the patient" and what society is willing to pay. Here, resources allocated to anaesthesia care will have to compete not only with other medical or surgical specialties, but also with non-medical factors contributing to health of the population including recreation, roads, and the arts.[18] Thus, the introduction of new technology must be seen to be cost-effective not only in comparison with other competing technologies which purport to prevent the same disorders but also in relation to preventing or treating other non-related disorders. New, more cost-effective monitors may replace existing technologies, but unless the old technology is abandoned, newer technologies could be considered as "add-ons."

Effectiveness

The study of the use of medical technology in the every day world is known as "outcomes research." Outcomes research has been defined as "linking the type of care received by a variety of patients with a particular condition to positive and negative outcomes in order to identify what works best for which patients."[19] Guadagnoli and McNeil[19] suggest that there are a variety of factors which have led to a marked interest in this research. These include the costs of care, the need to make decisions about resource allocation, and the need to assess quality of care. While we are seeing more outcomes research in the field of anaesthesia, these studies are still rare.[20]

The future

The need for a rigorous scientific approach to technology assessment of anaesthesia monitors is clear. Innovative devices continue to be developed and traditional approaches to the assessment of monitors that clinicians intuitively believe are effective have been inadequate and narrowly focused. Anaesthetists must develop innovative techniques of assessment that specifically address problems related to anaesthesia outcome. We have learned that outcome measures used in internal medicine, surgery and radiology are good starting points, but are not specific enough for our needs. For example, clinical decision analyses often use utilities to measure quality of life in such ways as quality-adjusted life expectancy (QALE) or quality-adjusted life years (QALY) so that alternate therapies can be directly compared to one another. The conceptual framework behind such "endpoints" are years of life spent in various "health states" which range from perfect health, physical disability, pain, discomfort, emotional problems, social dysfunction, and so forth.[21] Since anaesthesia does not directly "treat" diseases, these concepts have limited applicability to the discipline of anaesthesia. Nonetheless, the development of a "single" quality endpoint for anaesthesia will be worthwhile so that the cost-effectiveness of various drugs, monitors, and techniques can be more easily compared.

Another area where research is needed is to address the "warning function" of monitors. Examples of innovations that may be useful in addressing this issue are simulation technology and prioritized monitoring alarms. Both of these techniques address issues related to the warning function. Simulation-based learning may help us understand how clinicians respond to warning signals and change treatment strategies when confronted with additional information. Experience with simulation-based research methods may assist in better defining the "lead time" required to effect changes in therapy.[22] Prioritized alarms may aid in establishing a sequential response pattern and minimize information overload from the cacophony of competing monitoring alarm systems.[6] Of course, these systems would need to be evaluated before widespread adoption.

Conclusion

Unless clinical investigators become more innovative and focus on specific anaesthesia problems, it will be difficult, if not impossible, to demonstrate improved patient outcome from our monitors. Without this supportive information, anaesthetists will be at a disadvantage when resource allocation decisions are made. It is essential that our investigators be trained in clinical epidemiology, the decision sciences, and biostatistics as well as clinical anaesthesia, to address creatively specific problems related to anaesthesia practice.

Acknowledgements

The authors wish to express our appreciation to Ms. Sharon Klimosco and Ms. Linda Leslie who prepared the manuscript for publication.

References

1 *Tuman KJ, Ivankovich AD.* High-cost, high-tech medicine: are we getting our money's worth? J Clin Anesth 1993: 5: 168–77.

2 *Rettig RA.* Technology assessment – an update. Invest Radiol 1991; 26: 165–73.

3 *Brook RH, Lohr KN.* Efficacy, effectiveness, variations, and quality: boundary-crossing research. Med Care 1985; 23: 710–22.

4 *Orkin FK, Cohen MM, Duncan PG.* The quest for meaningful outcomes (Editorial). Anesthesiology 1993; 78: 417–22.

5 *Duncan PG, Cohen MM.* Pulse oximetry and capnography in anaesthetic practice: an epidemiological appraisal. Can J Anaesth 1991; 38: 619–25.

6 *Orkin FK.* Patient monitoring during anesthesia as an ex-

ercise in technology assessment. *In*: Saidman LJ, Smith NT (Eds.). Monitoring in Anesthesia, 3rd ed. Boston: Butterworth-Heinemann, 1993; 439–55.

7 *Coté CJ, Goldstein EA, Coté MA, Hoaglin DC, Ryan JF.* A single-blind study of pulse oximetry in children. Anesthesiology 1988; 68: 184–8.

8 *Keats AS.* Anesthesia mortality in perspective. Anesth Analg 1990; 71: 113–9.

9 *Cohen MM, Duncan PG, Tweed WA, et al.* The Canadian four-centre study of anaesthetic outcomes: I. Description of methods and populations. Can J Anaesth 1992; 39: 420–9.

10 *Cohen MM, Duncan PG, Tate RB.* Does anesthesia contribute to operative mortality? JAMA 1988; 260: 2859–63.

11 *Lunn JN, Devlin HB.* Lessons from the confidential inquiry into perioperative deaths in three NHS regions. Lancet 1987; 2: 1384–6.

12 *Canadian Anaesthetists' Society, Members' Guide 1993.* Guidelines to the Practice of Anaesthesia: as recommended by the Canadian Anaesthetists' Society. Can J Anaesth 1993; 40: 10.

13 *Moller JT, Pederson T, Rasmussen LS, et al.* Randomized evaluation of pulse oximetry in 20,802 patients. I. Design, demography, pulse oximetry failure rate, and overall complication rate. Anesthesiology 1993; 78: 436–44.

14 *Moller JT, Johannessen NW, Espersen K, et al.* Randomized evaluation of pulse oximetry in 20,802 patients. II. Perioperative events and postoperative complications. Anesthesiology 1993; 78: 445–53.

15 *Fryback DG, Thornbury JR.* The efficacy of diagnostic imaging. Med Decis Making 1991; 11: 88–94.

16 *Naylor CD, Williams JI, Basinski A, Goel V.* Technology assessment and cost-effectiveness analysis: misguided guidelines? Can Med Assoc J 1993; 148: 921–4.

17 *White PF, Watcha MF.* Are new drugs cost-effective for patients undergoing ambulatory surgery? (Editorial) Anesthesiology 1993; 78: 2–5.

18 Premier's council on health, well-being and social justice. Health for all Ontarians. A provincial dialogue on the determinants of health. Queen's printer for Ontario, 1994.

19 *Guadagnoli E, McNeil BJ.* Outcomes research: hope for the future or the latest rage? Inquiry 1994; 31: 14–24.

20 *Hudson RJ, Friesen RM.* Health care 'reform' and the costs of anaesthesia (Editorial). Can J Anaesth 1993; 40: 1120–5.

21 *Sandberg SI, Barnes BA, Weinstein MC, Braun P.* Elective hysterectomy: benefits, risks, and costs. Med Care 1985; 23: 1067–85.

22 *Gaba D, DeAnda A.* A comprehensive anesthesia simulation environment: re-creating the operating room for research and training. Anesthesiology 1988; 69: 387–94.