

EDITORIAL

Prognostic studies of perioperative risk: robust methodology is needed

M. P. W. Grocott^{1,2,3*} and R. M. Pearse⁴

¹ Surgical Outcomes Research Centre (SOuRCe), Joint UCLH/UCL Comprehensive Biomedical Research Centre, London, UK

² General Intensive Care Unit, Southampton University Hospitals NHS Trust, Southampton, UK

³ NIAA Health Services Research Centre, Royal College of Anaesthetists, London, UK

⁴ Barts and The London School of Medicine and Dentistry, Queen Mary's University of London, London, UK

* Corresponding author. E-mail: mike.grocott@ucl.ac.uk

Estimation of the likelihood of adverse outcome after surgery is a central objective of preoperative assessment. However, it remains uncertain which technique provides the most accurate prediction of perioperative risk. Limitations in the design and conduct of many studies evaluating methods of risk assessment contribute to this uncertainty. Further studies of preoperative risk assessment are urgently required and should meet several essential criteria.

Tissue injury and physiological disturbance associated with major surgery can have long-lasting consequences and may sometimes be immediately life-threatening. Low overall mortality rates conceal the existence of a subgroup of high-risk patients who accounts for more than 80% of postoperative deaths.^{1–2} In the UK alone, more than 170 000 high-risk non-cardiac surgical procedures are performed each year, after which 100 000 patients develop complications resulting in more than 25 000 deaths before hospital discharge.^{1–3} Importantly, those patients who develop complications but survive to leave hospital suffer a substantial reduction in functional independence and long-term survival.^{4–6} Some have suggested that among developed nations, poor surgical outcomes are a particular problem in the UK.⁷ However, published data suggest that the issue is widespread.^{5–6–8–9} Recent figures suggest that more than 230 million surgical procedures are performed worldwide each year.¹⁰ Assuming a mortality rate of between 0.5% and 2.0%,^{6–11} surgery will be associated with 1.25–5 million deaths annually worldwide. Morbidity rates

are likely to be between five and 10 times this figure. Clearly, major surgery represents an extremely important cause of death and disability, even in those parts of the world where mortality rates are low.¹⁰

The evidence of the impact of poor surgical outcomes is growing. However, we appear to be failing to accurately identify high-risk surgical patients and allocate them to an appropriate level of perioperative care. Recent UK data indicate that only a minority of high-risk patients are admitted directly to critical care after surgery, and that many postoperative deaths occur following delayed admission to critical care with initial treatment on a standard surgical ward.^{1–2} Precise evaluation of an individual's risk of death or complications is fundamental to the process of improving postoperative outcomes, but remains an elusive goal.¹² Such information would allow more effective allocation of resources and targeting of specific interventions to those patients most likely to benefit. Potentially, beneficial strategies include preoperative specialist medical review, specific perioperative interventions, such as β -blockade or flow-guided fluid therapy and elective critical care admission. Greater precision in the assessment of risk would also improve the quality of information that can be given to patients, many of whom simply rely on the guidance of surgeons and anaesthetists when consenting for surgery. Objective, accurate information balancing the risks and benefits of surgery should improve the quality, and experience, of patient decision-making.

Although a variety of methods are currently used to provide objective assessment of perioperative risk, none has yet been subjected to the type of robust evaluation required for routine clinical use. Cohort studies are a valuable and commonly used method of evaluating prognostic tools but when the test under investigation is allowed to influence treatment decisions, 'confounding by indication' may complicate interpretation, particularly in un-blinded studies.¹³ In other words, if clinicians are aware that a given result may indicate a poor prognosis, it is likely that they will alter their practice accordingly, with corresponding effects on outcome. A prognostic test might precisely and reliably identify patients at high risk who are then appropriately admitted to critical care after surgery. If one accepts the assumption that critical care may improve outcome, when compared with ward care, then the prognostic value of the test will be underestimated in this situation. Importantly, this methodological problem distorts not only the measurement of predictive accuracy but also the optimal threshold values for a given test used to categorize patients into high- and low-risk groups.

This issue is particularly relevant to the evaluation of cardiopulmonary exercise testing (CPET) as a means of assessing the risks of non-cardiac surgery. This prognostic test involves breath-by-breath expired gas analysis in patients undergoing a graded exercise challenge to derive objective measures of functional cardiorespiratory capacity (anaerobic threshold and peak oxygen consumption). Several studies have reported an association between postoperative mortality and poor cardiorespiratory reserve as determined by CPET.¹⁴⁻²¹ However, in all of these studies, clinicians were aware of, and their clinical decisions influenced by, the results of CPET. Inevitably, this has obscured the true relationship between CPET-derived measures of risk and patient outcome. Such confounding by indication in cohort studies can be mitigated to some extent by a variety of methods. These include restriction of subjects, adjustment using propensity scores, blinded prospective review, ecological analysis, and the instrumental variable approach.²² Unfortunately, none of these methods was used in the studies cited above. The recent findings of a small, single-centre study, with some clinician blinding of CPET data, hint at a stronger association between low anaerobic threshold and postoperative morbidity than described previously.²³ The hypothesis that CPET-derived variables could be used to accurately predict surgical risk is intuitively appealing, but is only now, finally, being robustly tested. In contrast, most of the studies that have evaluated the predictive accuracy of plasma biomarkers such as B-type natriuretic peptide (BNP) have been blinded,^{24, 25} reducing the chance of significant confounding. Nonetheless, even in the case of plasma BNP measurement, large multi-centre trials are required to confirm predictive accuracy and identify the optimal threshold values for allocating patients into high- and low-risk categories.

In order to confirm the validity of a particular test as a means of assessing risk in individual patients, it is not sufficient simply to demonstrate that poor clinical outcomes

are more frequently associated with an abnormal result. Although such association studies can identify promising candidate methods of risk assessment, more rigorous evaluation is required to establish their utility in clinical practice. Further studies must assess predictive accuracy, establish the optimal discriminatory threshold for categorizing patients according to surgical risk and demonstrate that the test has incremental value over and above existing methods. Where possible, such investigations should be followed by randomized trials to confirm whether the use of the test to triage patients to specific interventions improves outcome. Randomized trials also allow health economic evaluations that can provide powerful arguments for or against implementation. The current evidence base for perioperative risk assessment falls far short of this ideal standard. It is also essential that our evaluation of risk assessment technologies is objective. A lack of equipoise among clinicians can represent a formidable barrier to the robust evaluation of health-care technology. This in turn hampers widespread implementation into routine practice.

There is a clear and urgent need for large, well-designed clinical investigations to define the optimal approach to risk assessment before non-cardiac surgery. Until such studies have been performed, we must recognize the important limitations of the available evidence, make best use of currently available tools, and be prepared to adapt our practice as new data become available.

Conflict of interest

None declared.

Funding

M.P.W.G. (SOuRCe) is funded in part by the University College London Hospital—University College London Comprehensive Biomedical Research Centre which received a portion of its funding from the United Kingdom Department of Health's National Institute of Health Research Comprehensive Biomedical Research Centre funding scheme. R.M.P. is a National Institute for Health Research (UK) Clinician Scientist.

References

- 1 Jhanji S, Thomas B, Ely A, Watson D, Hinds CJ, Pearse RM. Mortality and utilisation of critical care resources amongst high-risk surgical patients in a large NHS trust. *Anaesthesia* 2008; **63**: 695–700
- 2 Pearse RM, Harrison DA, James P, et al. Identification and characterisation of the high-risk surgical population in the United Kingdom. *Crit Care* 2006; **10**: R81
- 3 Cullinane M, Gray AJ, Hargraves CM, Lansdown M, Martin IC, Schubert M. *The 2003 Report of the National Confidential Enquiry into Peri-Operative Deaths*. London: NCEPOD, 2003
- 4 Head J, Ferrie JE, Alexanderson K, Westerlund H, Vahtera J, Kivimaki M. Diagnosis-specific sickness absence as a predictor of mortality: the Whitehall II prospective cohort study. *Br Med J* 2008; **337**: a1469
- 5 Khuri SF, Henderson WG, DePalma RG, Mosca C, Healey NA, Kumbhani DJ. Determinants of long-term survival after major

- surgery and the adverse effect of postoperative complications. *Ann Surg* 2005; **242**: 326–41
- 6 Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. *N Engl J Med* 2009; **360**: 1418–28
 - 7 Bennett-Guerrero E, Hyam JA, Shaefi S, et al. Comparison of P-POSSUM risk-adjusted mortality rates after surgery between patients in the USA and the UK. *Br J Surg* 2003; **90**: 1593–8
 - 8 Juul AB, Wetterslev J, Gluud C, et al. Effect of perioperative beta blockade in patients with diabetes undergoing major non-cardiac surgery: randomised placebo controlled, blinded multicentre trial. *Br Med J* 2006; **332**: 1482
 - 9 Rigg JR, Jamrozik K, Myles PS, et al. Epidural anaesthesia and analgesia and outcome of major surgery: a randomised trial. *Lancet* 2002; **359**: 1276–82
 - 10 Weiser TG, Regenbogen SE, Thompson KD, et al. An estimation of the global volume of surgery: a modelling strategy based on available data. *Lancet* 2008; **372**: 139–44
 - 11 NHS Executive. *HSC 2000/023. Quality and performance in the NHS: NHS Performance Indicators; July 2000*. Leeds: Department of Health, 2001
 - 12 Reilly CS. Can we accurately assess an individual's perioperative risk? *Br J Anaesth* 2008; **101**: 747–9
 - 13 Johnston SC. Identifying confounding by indication through blinded prospective review. *Am J Epidemiol* 2001; **154**: 276–84
 - 14 Carlisle J, Swart M. Mid-term survival after abdominal aortic aneurysm surgery predicted by cardiopulmonary exercise testing. *Br J Surg* 2007; **94**: 966–9
 - 15 Epstein SK, Freeman RB, Khayat A, Unterborn JN, Pratt DS, Kaplan MM. Aerobic capacity is associated with 100-day outcome after hepatic transplantation. *Liver Transpl* 2004; **10**: 418–24
 - 16 Forshaw MJ, Strauss DC, Davies AR, et al. Is cardiopulmonary exercise testing a useful test before esophagectomy? *Ann Thorac Surg* 2008; **85**: 294–9
 - 17 McCullough PA, Gallagher MJ, Dejong AT, et al. Cardiorespiratory fitness and short-term complications after bariatric surgery. *Chest* 2006; **130**: 517–25
 - 18 Nagamatsu Y, Shima I, Yamana H, Fujita H, Shirouzu K, Ishitake T. Preoperative evaluation of cardiopulmonary reserve with the use of expired gas analysis during exercise testing in patients with squamous cell carcinoma of the thoracic esophagus. *J Thorac Cardiovasc Surg* 2001; **121**: 1064–8
 - 19 Nagamatsu Y, Yamana H, Fujita H, et al. The simultaneous evaluation of preoperative cardiopulmonary functions of esophageal cancer patients in the analysis of expired gas with exercise testing. *Nippon Kyobu Geka Gakkai Zasshi* 1994; **42**: 2037–40
 - 20 Nugent AM, Riley M, Megarry J, O'Reilly MJ, MacMahon J, Lowry R. Cardiopulmonary exercise testing in the pre-operative assessment of patients for repair of abdominal aortic aneurysm. *Ir J Med Sci* 1998; **167**: 238–41
 - 21 Older P, Hall A, Hader R. Cardiopulmonary exercise testing as a screening test for perioperative management of major surgery in the elderly. *Chest* 1999; **116**: 355–62
 - 22 McMahon AD. Approaches to combat with confounding by indication in observational studies of intended drug effects. *Pharmacoevidenciol Drug Saf* 2003; **12**: 551–8
 - 23 Snowden CP, Prentis JM, Anderson HL, et al. Submaximal cardiopulmonary exercise testing predicts complications and hospital length of stay in patients undergoing major elective surgery. *Ann Surg* 2010; **251**: 535–41
 - 24 Sear JW, Howard-Alpe G. Preoperative plasma BNP concentrations: do they improve our care of high-risk non-cardiac surgical patients? *Br J Anaesth* 2007; **99**: 151–4
 - 25 Cuthbertson BH, Amiri AR, Croal BL, et al. Utility of B-type natriuretic peptide in predicting perioperative cardiac events in patients undergoing major non-cardiac surgery. *Br J Anaesth* 2007; **99**: 170–62