Thorax Online First, published on July 13, 2010 as 10.1136/thx.2009.133181

Cardiopulmonary exercise testing for the selection of patients undergoing surgery for lung cancer: friend or foe?

Eric Lim,[1] Michael Beckles,[2] Chris Warburton,[3] David Baldwin[4]

The contribution of exercise testing for risk assessment for lung resection is well established and has been embedded in international guidelines from Europe[1] and the USA.[2] There are many forms of exercise tests (6 min walk, 12 min walk, shuttle walk, stair climbing), but the most established investigation is formal assessment of maximum oxygen consumption during exercise (V_{O_2} max). British and American (American College of Chest Physicians (ACCP)) guidelines use V_{O_2} max as the ultimate assessment of operative risk, positioned at or near the end of the functional algorithm,[3] whereas European guidelines recommend the use of this test much earlier in patients with a forced expiratory volume in 1 s (FEV_1) or carbon monoxide transfer factor (TlCO) <80% predicted.[1 2]

Numerous cohort studies and a meta-analysis report the association of low V_{O_2} max and 'high risk' lung resection.[4–18] However 'high' is not quantified and 'risk' is not defined, two fundamentally important definitions if guidelines that use these terms are to be applied clinically. Here we focus on validity of the V_{O_2} max studies and the clinical utility of the available evidence with respect to individual interpretation of risk.

SAMPLE SIZE AND PRECISION OF RISK ESTIMATION OF DEATH

Arguably, the most important outcome when considering surgery for lung cancer is the ability to survive the procedure. The most apparent limitation of the currently available evidence is the lack of appropriately powered studies to address this. The precision of a risk model is not specifically dependent on sample size, but rather the number of events—that is deaths—an uncommon outcome in thoracic surgery. In the UK, lobectomy, the most common procedure for lung cancer, carried an operative mortality of ~2% in 2004–2005,[19] and in the USA the mortality rate has been reported to range from 2.3% to 4.1%.[20] Reflective of this, the largest study in this context on V_{O_2} max (422 patients) had only 15 deaths. What is clearly more disconcerting is that publications for which recommendations on estimation of operative mortality risk have been based have sample sizes ranging from 8 to 160.[21]

UPPER LIMITS OF UNCERTAINTY FOR SAFE CUT-OFF VALUES

Many studies have defined arbitrary cut-off values ranging from 15 to 20/ml/kg/min as a 'safe' cut-off value[4 12–14] because above these levels no patient experienced an adverse event. What is the validity of this type of recommendation?

The answer lies in the uncertainly that surrounds the observation of no events (ie, upper 95% CI), a function of the sample size. For a standard binomial distribution, the upper limit of the CI of zero events with the sample size of 8–160 corresponds to 42–2.7%, respectively (figure 1), illustrating high limits of uncertainty in the majority of studies with smaller sample sizes.

ALTERNATIVE MODELS FOR RISK ESTIMATION OF DEATH

Given these limitations, are there any other alternatives for the risk assessment for operative morality? Thoracoscore is a composite scoxring system that can be used to quantify risk. It is a logistic regression-derived model with coefficients provided for individual risk factors, calculated to provide a percentage probability of death. It is currently the best model and was developed from a sample size of 15 183 patients with 338 deaths, and provides excellent discrimination with an area under the curve of 0.82.[22] Furthermore, it has been validated in different populations.[23] Apart from superior statistical power, much larger sample size, external validity and excellent performance, the logistic risk model carries two further attractive advantages compared with V_{O_2} max assessment: it is cost free and can be universally available.

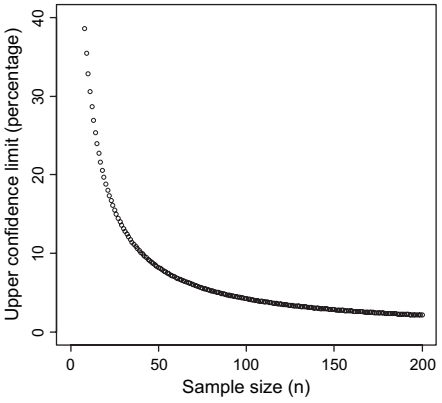

Figure 1 Upper binomial confidence limit for 'no observations'.

OTHER OUTCOMES AND COMPOSITE END POINTS

The consistent message that lower values of V_{O_2} max are associated with higher risk of complications is to be expected as a measure of cardiovascular fitness. From a patient's and clinician's perspective, however, the nature of the complications is of central importance. All studies to date have used composite end points and, when multiple outcomes are combined, it becomes difficult to interpret the impact of each individual component. It has been recommended that each outcome should have a similar weighting or clinical importance to facilitate clinical interpretation.[24] For example, death and myocardial infarction would be combined to estimate the total number of patients that may have experienced a myocardial infarction and survived, added to those that have (presumably) experienced a myocardial infarction and died.

Researchers, however, may use composite outcomes to increase the power of the study (by increasing event rate) and therefore increasing the chances of

[1]Imperial College and Academic Division of Thoracic Surgery, The Royal Brompton Hospital, London, UK; [2]Department of Acute Medicine, The Royal Free Hospital, London, UK; [3]Department of Respiratory Medicine, Aintree Chest Centre, Liverpool, UK; [4]Respiratory Medicine Unit, David Evans Research Centre, Nottingham University Hospitals, Nottingham, UK

Correspondence to Eric Lim, Academic Division of Thoracic Surgery, Royal Brompton Hospital, Sydney Street, London SW3 6NP, UK; e.lim@rbht.nhs.uk

Copyright Article author (or their employer) 2010. Produced by BMJ Publishing Group Ltd (& BTS) under licence.

Thorax: first published as 10.1136/thx.2009.133181 on 13 July 2010. Downloaded from http://thorax.bmj.com/ on April 19, 2024 by guest. Protected by copyright.

achieving statistically significant results.[25] The corollary is that important outcomes such as death can be piggybacked within the pool of less important outcomes such as atelectasis[5 7–9 11–14 16] or purulent sputum,[4] giving rise to considerable difficulties for the clinician and, more importantly, the patient to evaluate the importance of the overall result. We believe that most patients would not consider readmission to the intensive care unit,[11] atelectasis,[5 7–9 11–14 16] arrhythmia[4 6–8 10 13 14] or postoperative CO_2 retention[6 8–10 13 16] as 'prohibitive' complications leading them to refuse surgery.

QUANTIFICATION AND INTERPRETATION OF RISK

A clear explanation of risks and benefits is central to good consenting practice when offering treatment options to our patients. Dichotomous categorisation of 'high' and 'standard' risk using V_{O_2} max for risk assessment, accompanied by a combination of varied outcomes (some of which have little influence on patient decision making) renders the information difficult to apply in practice. The lack of a numerical estimate leads to subjective interpretation of 'high'; moreover, many studies do not document the uncertainty (confidence limits) that surround their estimates. As there is no accepted level of baseline risk, it is not possible to quantify the relative magnitude of 'high' to facilitate the interpretation.

COST OF GETTING IT WRONG

It is intuitive that clinicians seek to protect the interests of their patients, and some may wonder if a discussion of the quantification and interpretation of risk is relevant as opposed to acceptance and avoidance of risk based on published values. In the CALGB 9238 study, the largest in the series (with 422 patients), physicians were allowed to offer surgical treatment of patients with 'very high risk', defined as FEV_1 <900 ml and V_{O_2} max of <15/ml/kg/min. Of the 68 patients in the 'very high risk' group, there was only one postoperative death within 30 days and a total of three in-hospital deaths.[17] More importantly, on follow-up, the operated patients in the very high risk group had more than double the median survival compared with the non-operated patients (36.0 months vs 15.8 months, $p<0.001$), illustrating acceptable procedural mortality and morbidity with twice the median survival with case selection on parameters independent of V_{O_2} max. Denying patients with 'prohibitive' values of V_{O_2} max the opportunity to consider surgery as a management option may in fact be against their best interests. As the study was not randomised, it is important to bear in mind the invariable presence of selection bias, and the possibility that a better result was achieved by offering surgery to fitter patients with less comorbidity. Our point is more to question the 'conventional' lower limit of safety and the results that can be achieved by further selection.

THE FUTURE

We acknowledge the consistent message that low levels of V_{O_2} max are associated with increased complications from surgery. However, we believe current recommendations are flawed by small sample sizes, resulting in imprecise risk estimates. Moreover, the lack of numerical quantification leads to difficulties in defining the level of acceptable risk. Furthermore, the use of composite outcomes leads to a lack of agreement on the importance of the risks, and the incongruence limits the clinical applicability to inform patients on the decision to undergo surgery. We believe that management options should be discussed at a multidisciplinary level but decisions should be undertaken at patient level. This is because patients are heterogeneous, with individual perceptions on the value of benefit and risk. As the lower limits of safety remain imprecisely defined, patients with multidisciplinary team-defined 'prohibitive' levels of risk may not be offered the opportunity to consider surgery as an option and denied the possibility of increased life expectancy.

There may also be a degree of concern if postoperative quality of life may be a trade-off for any increase in life expectancy in the high risk cohort; however, prospective studies indicated that patients traditionally considered at higher risk of lung resection had postoperative physical and emotional quality of life scores similar to those observed in younger and fitter patients.[26]

Before widespread use, further work needs to be performed to determine if cardiopulmonary exercise testing is an independent predictor of mortality (eg, above and beyond that of Thorascore), to relate the study to individual outcomes that would influence the decision to undergo surgery, to provide numerical quantification of risk with an estimate of uncertainty and to demonstrate validity in different cohorts.

Competing interests None.

Provenance and peer review Commissioned; externally peer reviewed.&online_st;

Thorax 2010;■:1—3.
doi:10.1136/thx.2009.133181

REFERENCES

1. **Brunelli A,** Charloux A, Bolliger CT, *et al*. ERS/ESTS clinical guidelines on fitness for radical therapy in lung cancer patients (surgery and chemo-radiotherapy). *Eur Respir J* 2009;**34**:17—41.
2. **Colice GL,** Shafazand S, Griffin JP, *et al*. Physiologic evaluation of the patient with lung cancer being considered for resectional surgery: ACCP evidenced-based clinical practice guidelines (2nd edition). *Chest* 2007;**132**(3 Suppl):161S—77S.
3. **British Thoracic Society; Society of Cardiothoracic Surgeons of Great Britain and Ireland Working Party.** BTS guidelines: guidelines on the selection of patients with lung cancer for surgery. *Thorax* 2001;**56**:89—108.
4. **Bechard D,** Wetstein L. Assessment of exercise oxygen consumption as preoperative criterion for lung resection. *Ann Thorac Surg* 1987;**44**:344—9.
5. **Bolliger CT,** Jordan P, Soler M, *et al*. Exercise capacity as a predictor of postoperative complications in lung resection candidates. *Am J Respir Crit Care Med* 1995;**151**:1472—80.
6. **Boysen PG,** Clark CA, Block AJ. Graded exercise testing and postthoracotomy complications. *J Cardiothorac Anesth* 1990;**4**:68—72.
7. **Brunelli A,** Belardinelli R, Refai M, *et al*. Peak oxygen consumption during cardiopulmonary exercise test improves risk stratification in candidates to major lung resection. *Chest* 2009;**135**:1260—7.
8. **Brutsche MH,** Spiliopoulos A, Bolliger CT, *et al*. Exercise capacity and extent of resection as predictors of surgical risk in lung cancer. *Eur Respir J* 2000;**15**:828—32.
9. **Epstein SK,** Faling LJ, Daly BD, *et al*. Predicting complications after pulmonary resection. Preoperative exercise testing vs a multifactorial cardiopulmonary risk index. *Chest* 1993;**104**:694—700.
10. **Richter Larsen K,** Svendsen UG, Milman N, *et al*. Exercise testing in the preoperative evaluation of patients with bronchogenic carcinoma. *Eur Respir J* 1997;**10**:1559—65.
11. **Markos J,** Mullan BP, Hillman DR, *et al*. Preoperative assessment as a predictor of mortality and morbidity after lung resection. *Am Rev Respir Dis* 1989;**139**:902—10.
12. **Morice RC,** Peters EJ, Ryan MB, *et al*. Exercise testing in the evaluation of patients at high risk for complications from lung resection. *Chest* 1992;**101**:356—61.
13. **Smith TP,** Kinasewitz GT, Tucker WY, *et al*. Exercise capacity as a predictor of post-thoracotomy morbidity. *Am Rev Respir Dis* 1984;**129**:730—4.
14. **Torchio R,** Gulotta C, Parvis M, *et al*. Gas exchange threshold as a predictor of severe postoperative complications after lung resection in mild-to-moderate chronic obstructive pulmonary disease. *Monaldi Arch Chest Dis* 1998;**53**:127—33.
15. **Villani F,** Busia A. Preoperative evaluation of patients submitted to pneumonectomy for lung carcinoma: role of exercise testing. *Tumori* 2004;**90**:405—9.
16. **Wang J,** Olak J, Ultmann RE, *et al*. Assessment of pulmonary complications after lung resection. *Ann Thorac Surg* 1999;**67**:1444—7.
17. **Loewen GM,** Watson D, Kohman L, *et al*. Preoperative exercise V_{O_2} measurement for lung resection candidates: results of Cancer and Leukemia Group B Protocol 9238. *J Thorac Oncol* 2007;**2**:619—25.

Thorax: first published as 10.1136/thx.2009.133181 on 13 July 2010. Downloaded from http://thorax.bmj.com/ on April 19, 2024 by guest. Protected by copyright.

18. **Bobbio A,** Chetta A, Internullo E, *et al*. Exercise capacity assessment in patients undergoing lung resection. *Eur J Cardiothorac Surg* 2009;**35**:419—22.
19. **Page R,** Keogh B. *National Thoracic Surgery Activity and Outcomes Report*. Oxford; Dendrite Clinical Systems Ltd, 2008.
20. **Schipper PH,** Diggs BS, Ungerleider RM, *et al*. The influence of surgeon specialty on outcomes in general thoracic surgery: a national sample 1996 to 2005. *Ann Thorac Surg* 2009;**88**:1566—72;discussion 1572—63.
21. **Benzo R,** Kelley GA, Recchi L, *et al*. Complications of lung resection and exercise capacity: a meta-analysis. *Respir Med* 2007;**101**:1790—7.
22. **Falcoz PE,** Conti M, Brouchet L, *et al*. The Thoracic Surgery Scoring System (Thoracoscore): risk model for in-hospital death in 15,183 patients requiring thoracic surgery. *J Thorac Cardiovasc Surg* 2007;**133**:325—32.
23. **Chamogeorgakis TP,** Connery CP, Bhora F, *et al*. Thoracoscore predicts midterm mortality in patients undergoing thoracic surgery. *J Thorac Cardiovasc Surg* 2007;**134**:883—7.
24. **Montori VM,** Permanyer-Miralda G, Ferreira-Gonzalez I, *et al*. Validity of composite end points in clinical trials. *Br Med J* 2005;**330**:594—6.
25. **Lim E,** Brown A, Helmy A, *et al*. Composite outcomes in cardiovascular research: a survey of randomized trials. *Ann Intern Med* 2008;**149**:612—17.
26. **Brunelli A,** Socci L, Refai M, *et al*. Quality of life before and after major lung resection for lung cancer: a prospective follow-up analysis. *Ann Thorac Surg* 2007;**84**:410—16.

Thorax: first published as 10.1136/thx.2009.133181 on 13 July 2010. Downloaded from http://thorax.bmj.com/ on April 19, 2024 by guest. Protected by copyright.