Review Article

Risk prediction for lung cancer surgery in the current era

Marcus Taylor[1], George Whittaker[1], Stuart W. Grant[2], Richard Booton[3], Rajesh Shah[1]

[1]Department of Cardiothoracic Surgery, Manchester University Hospital NHS Foundation Trust, Wythenshawe Hospital, Manchester, UK; [2]Division of Cardiovascular Sciences, University of Manchester, ERC, Manchester University Hospital NHS Foundation Trust, Manchester, UK; [3]Department of Respiratory Medicine, Manchester University Hospital NHS Foundation Trust, Wythenshawe Hospital, Manchester, UK

Contributions: (I) Conception and design: M Taylor, R Shah; (II) Administrative support: G Whittaker; (III) Provision of study materials or patients: M Taylor; (IV) Collection and assembly of data: M Taylor, G Whittaker; (V) Data analysis and interpretation: M Taylor, SW Grant; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Marcus Taylor. Department of Cardiothoracic Surgery, Manchester University Hospital NHS Foundation Trust, Wythenshawe Hospital, Southmoor Road, Manchester, M23 9LT, UK. Email: Marcus.taylor1@nhs.net.

Abstract: Patients with lung cancer being considered for lung resection should undergo a tripartite assessment to ensure that the risks associated with surgery are not prohibitively high. A key component of the tripartite assessment is estimating the risk of peri-operative mortality which is normally calculated using a clinical prediction model. Clinical prediction models are mathematical equations that use information from patient risk factors to predict the risk of a healthcare outcome. When developed using appropriate statistical methodology and externally validated with acceptable rigour, their role in clinical practice is important. However, if inaccurately developed or incorrectly applied, their use can lead to patients being erroneously accepted for or denied surgery according to their perceived risk. Although several different models have been developed to predict operative risk in patients with lung cancer, contemporary validation studies demonstrate that no single model can currently be recommended for use in routine clinical practice. This is due to either inadequate statistical performance, inclusion of only a limited number of relevant clinical variables or utilisation of an outcome that does not adequately capture all deaths related to the index surgical procedure. Although surgery remains the gold standard approach for the treatment of lung cancer, an increasing number of non-surgical treatments for lung cancer with ever-improving outcomes are available. As a result, accurate risk stratification of patients with lung cancer has assumed an even greater significance. The absence of well-performing models designed to predict appropriate outcome metrics demonstrates the need for further work in this important area of lung cancer surgery.

Keywords: Risk prediction; risk model; 90-day mortality; RESECT-90; lung cancer surgery

Received: 29 December 2021; Accepted: 12 February 2022; Published: 15 March 2022.

doi: 10.21037/vats-21-47

View this article at: https://dx.doi.org/10.21037/vats-21-47

Background

Lung cancer remains the leading cause of cancer death worldwide (1) and is associated with particularly poor outcomes. Only around one third of patients are still alive 1 year after diagnosis, with less than 15% of patients surviving to 5 years (2). These poor survival statistics are largely due to the fact that more than half of all patients already have metastatic disease at the time of initial diagnosis (3). There are several different treatment options available for non-small cell lung cancer (NSCLC), including surgical resection, chemotherapy, radiotherapy, immunotherapy and a combination of two or more of these modalities (4).

Although for early-stage lung cancer stereotactic ablative body radiotherapy (SABR) has produced promising results (5), guidelines continue to advocate that all suitable patients with NSCLC should be offered surgical resection as a first-line treatment (6). In broad terms, in order to be deemed "suitable" for lung resection, patients must be fit enough to undergo surgery and must have a pattern of

disease for which the operating surgeon is confident of achieving a complete resection. These two factors underpin the selection process for lung resection.

There is clear evidence that the number of patients undergoing lung cancer surgery has increased over recent years, with around 7,000 resections for lung cancer currently being performed each year across the United Kingdom (7). The introduction of lung cancer screening programmes is one initiative designed to further increase the number of patients with lung cancer who are suitable for surgical resection. Results from initial screening programmes demonstrated that lung cancer was detected in 2–3% of patients, of which more than 80% were early stage (stage I/II) lung cancers (8,9). Early-stage lung cancers normally represent less than 25% of all lung cancers at the time of initial diagnosis (10). Widespread adoption of screening is therefore highly likely to lead to further increases in surgical activity.

With a rising number of operations being performed and potential alternative treatment options more readily available, the ability to accurately assess the risks of surgical resection is increasingly important. This review considers the pre- and peri-operative factors underpinning risk prediction in contemporary thoracic surgery.

The tripartite approach

For patients with NSCLC anatomically suitable for surgical resection, a comprehensive assessment to ensure that they are physiologically robust enough to undergo surgery and will recover an appropriate quality of life once treatment has been completed is required. In order to ensure that pre-operative decision making is robust and reproducible, a standardised approach should be applied to all patients. Guidelines from the British Thoracic Society (BTS) focus on three key elements of peri-operative risk, advocating a tripartite approach designed to simultaneously assess a patient's risk of experiencing post-operative dyspnoea, peri-operative cardiac event and peri-operative mortality (6).

Determining the extent of resection

Determining the extent of resection required to adequately treat lung cancer is a key first step in assessing the risk of surgery. Unlike other surgical procedures, where a return to baseline function can be expected after a period of post-operative recovery, a resection for NSCLC involves the removal of a substantial portion of healthy lung tissue. From an oncological perspective, the gold standard operation for NSCLC is an anatomical resection (11). This is usually a lobectomy, although for larger and more central tumours, a bi-lobectomy or pneumonectomy may be required. Anatomical resection is associated with a lower risk of locoregional recurrence and improved cancer-free survival but also inevitably involves the removal of healthy lung tissue (12). There is increasing evidence that minimising the amount of healthy lung tissue removed by performing sublobar anatomical resections is associated with acceptable oncological outcomes (13). The extent of resection required influences the risk of all three elements of the tripartite assessment and is therefore a principal component of the pre-operative risk assessment process.

Assessing post-operative dyspnoea

The assessment of lung function is a pre-requisite for all patients being considered for lung resection. The principal spirometry value traditionally used as part of the assessment for patients being considered for lung resection is forced expiratory volume in 1 second (FEV1). A further useful value is the percentage post-operative predicted FEV1, obtained by calculating the number of remaining bronchopulmonary segments after lung resection and multiplying this value (expressed as a fraction) by the percentage predicted FEV1. Studies have previously demonstrated that lower values of FEV1 and post-operative predicted FEV1 are associated with worse overall outcomes (14-18). However, more recently, studies have suggested that the association between spirometry values and outcomes has become weaker over time, particularly in patients with chronic obstructive pulmonary disease (COPD) (19,20).

Another important metric in the assessment of the risk of post-operative dyspnoea is predicted and percentage post-operative predicted diffusion capacity of the lung for carbon monoxide (DLCO) values. A number of studies have identified DLCO as a predictor of peri-operative mortality and morbidity after lung resection (21-25). DLCO has also been shown to be associated with adverse outcomes in patients with COPD where the impact of FEV1 and other spirometry values diminished over time (26). Although guidelines differ, the overall consensus is that no patient should be declined surgery based solely on lung function tests. Instead, either percentage predicted, or percentage post-operative predicted values below a defined threshold (dependent upon whichever guideline is being followed)

necessitates the patient undergoing additional objective functional testing (27).

Objective functional testing

Comprehensive physiological assessment with objective functional testing is deemed necessary when lung function results are suboptimal. Again, various different tests are available, including the stair climb, the six-minute walk, the shuttle walk test (SWT) and the cardiopulmonary exercise test (CPEX). With the exception of the CPEX, these measures of functional testing require minimal technical expertise and equipment. Despite CPEX being recognised as the gold standard physiological investigation, the SWT is included in both European and American guidelines (28,29). Studies investigating the correlation between the SWT, and peri-operative outcomes have found that patients walking <250 m are at prohibitively high risk and those walking >400 m are at very low risk of adverse outcomes (30). A SWT of less than 400 m has also been shown to be associated with an increased risk of adverse outcomes (31).

Although perhaps not reflective of real-world practice, guidelines suggest that all patients walking <400 m should be considered for CPEX. There is evidence demonstrating good correlation between distance walked on the SWT and maximal oxygen consumption (VO_2 max), one of the key variables measured as part of the CPEX test (32). Indeed, the study from Win *et al.* demonstrated that all patients who walked in excess of 400 m had a VO_2 max of >15 mL/kg/min (32). Given that previous evidence has shown that VO_2 max of ≥15 mL/kg/min is associated with a low rate of peri-operative morbidity and mortality (33), it is this study which underpins the guidelines where CPEX is deemed unnecessary for patients who have walked >400 m on the SWT.

Assessing cardiac risk

The assessment of cardiac risk is a key component of almost all non-cardiac surgery risk stratification pathways. A Revised Cardiac Risk Index score was devised in 1999 by Lee *et al.* (34) and subsequently revised in 2010 by Brunelli *et al.* (35), who published the Thoracic Revised Cardiac Risk Index (ThRCRI), to be used solely for patients being considered for lung resection. The four variables included in the model are serum creatinine >2 g/dL (1 point), undergoing pneumonectomy (1.5 points), presence of cerebrovascular disease (1.5 points) and history of coronary artery disease (1.5 points). The patient cohort from which this revised score was developed consisted of 1,696 patients, stratified into four different groups based on cardiac morbidity. Risk of cardiac morbidity ranged from 1.5% (group A, 0 points) to 5.8% (group B, 1.0–1.5 points) to 19% (group C, 2.0–2.5 points) to 23% (group D, >2.5 points). Two additional studies have externally validated these results, both of which demonstrated that a score of >2.5 was associated with significantly higher post-operative cardiac morbidity (36,37).

Assessing peri-operative mortality risk

Whilst the previous studies and scores cited have attempted to risk stratify according to the degree of cardiorespiratory comorbidity burden and the likelihood of experiencing post-operative cardiopulmonary complications, the principal component of any risk assessment process is accurate prediction of mortality directly attributable to the index surgical procedure. In order to ensure all peri-operative deaths are captured and to minimise the number of deaths included which are not related to surgery, the outcome metric for measuring peri-operative mortality has traditionally been either in-hospital mortality, 30-day mortality or a composite endpoint combining both. Although extensive research has been conducted regarding the influence of individual variables on peri-operative mortality, in order to produce a useful estimate of risk for each patient, a clinical risk prediction model is required. Both the BTS guidelines and the National Institute for Clinical Excellence (NICE) guidelines (38) advocate the use of a specific clinical prediction model (39).

What is a clinical prediction model?

In mathematical terms, a risk model is an equation which utilises patient risk factor information to provide an estimate of the probability of an individual patient experiencing a healthcare outcome. Whilst risk prediction models can be useful adjuncts to support clinical decision making, if they are used incorrectly or developed inaccurately, their use can lead to patients being inappropriately accepted for or denied surgery. Clinicians should be aware of how models are developed and validated in order to determine whether individual models can be applied to their own patient cohorts. Whilst statistical analysis may demonstrate that a model performs well overall, no information is provided as

to the reliability of the estimated mortality when applied to each individual within the cohort. Consequently, no patient should be denied surgery solely as a result of the estimated risk produced by a clinical risk prediction model.

How are clinical prediction models developed?

Developing a risk prediction model is both an art and a science. The overall aim is to produce an accurate model which can be readily used in clinical practice. Model accuracy must be balanced against model parsimony. A model containing a large number of variables, many of which are not routinely collected, may be highly accurate, but is unlikely to be regularly used in clinical practice if it cannot be easily calculated. Conversely, a model with fewer variables, all of which are routinely collected, may be less accurate but has a greater chance of being incorporated into routine clinical practice. Although a simple model may be easy to calculate, if it does not include key risk factors widely known to influence outcomes then the clinical validity of the model is questionable.

Clinical prediction models can be developed from clinical trials, observational studies or clinical registries. Most are developed using logistic regression although other methods are available. It is important that the objective of the model is clearly defined and that this is considered throughout the development process. Data used for development needs to be good quality, contemporary and representative of the population in which the model is intended to be used. The sample needs to be adequately powered for model development (40). Both predictors and outcome should be objectively measured, easily available, clearly defined and have minimal measurement error. The outcome should be important for both patients and clinicians. All of these model development features should be carefully considered with the aim of producing a clinically valid model with good statistical performance.

How are clinical prediction models assessed?

Clinical prediction model performance is assessed using measures of discrimination and calibration. Discrimination refers to the ability of the model to differentiate between those patients who experience the event and those who do not and is commonly measured using the area under the curve (AUC). An AUC of 1 represents perfect discrimination, whilst a value of 0.5 signifies that the model is no better than chance (i.e., flipping a coin) at predicting who will experience the outcome. An AUC of 0.7 is felt to represent acceptable model performance, and an AUC ≥0.8 represents excellent discriminatory ability.

Calibration is an assessment of how closely predicted outcomes match observed outcomes. The simplest measure of calibration is the observed to expected (O:E) ratio, which is calculated by dividing the mean observed and expected outcome rates. The drawback with an overall O:E ratio is that an under-estimation in one part of the data may be cancelled out by over-estimation in another, producing a ratio close to 1. A calibration plot is a more detailed measure and is produced by dividing the dataset into ten centiles, and plotting the mean observed and expected outcomes onto a graph, which is overlaid with a line representing perfect calibration. The Hosmer-Lemeshow (H-L) test is frequently used in isolation but provides no information as to either the extent or direction of the miscalibration and as such is no longer considered by statisticians to be an acceptable method of assessing calibration (41).

Clinical prediction model validation

Model performance should be assessed in two settings. Internal validation refers to assessment of model performance in the same cohort from which the model was developed. In this setting, model performance is likely to be overly optimistic. Although statistical approaches such as bootstrapping provide a degree of adjustment for optimism, internal validation in isolation is generally considered inadequate before a model is utilised in clinical practice. The strongest test of the performance of a model is external validation, which is performed on a different patient cohort, separated from the development cohort by either time or geographical location. A model which has been developed in accordance with statistical principles and has performed well on external validation is likely to be suitable for use in practice.

Mortality outcomes after thoracic surgery

The low and declining rates of in-hospital mortality after thoracic surgery in contemporary practice means that large multi-centre datasets are generally required for model development and validation. It is also thought that traditional measures of peri-operative mortality, such as in-hospital or 30-day mortality, do not adequately reflect the period of increased peri-operative mortality risk. Patients transferred outside of the centre in which they

Table 1 Variables included in the Thoracoscore model

Variables	Categories
Age	<55
	55–65
	>65
Gender	Male
	Female
ASA score	<3
	≥3
PS score	<3
	≥3
NYHA dyspnoea score	<3
	>3
Priority of surgery	Elective
	Urgent/emergency
Procedure class	Pneumonectomy
	Non-pneumonectomy
Diagnosis	Benign
	Malignant
Comorbidity score	Smoking
	History of cancer
	COPD
	Hypertension
	Heart disease
	Diabetes
	Peripheral vascular disease
	Obesity
	Alcoholism
	Hypercholesterolaemia

ASA, American Society of Anesthesiologists; PS, performance status; NYHA, New York Heart Association; COPD, chronic obstructive pulmonary disease.

were operated or who experience a particularly protracted post-operative course may not be identified using more traditional metrics. These measures would also fail to identify patients discharged from hospital who subsequently die at home shortly afterwards.

Several large-scale studies from across North America (42,43), Europe (44) and the UK (45) have shown that 90-day mortality after lung resection is twice the rate of 30-day mortality. This essentially demonstrates that the number of people dying between post-operative days 31 and 90 is the same as the number of patients dying within the first thirty days. Whilst the counter-argument is that selecting a longer time-period as an endpoint for measuring procedure-related mortality will unavoidably include patients whose deaths were not related to surgery, it is difficult to label an operation for lung cancer undertaken with curative intent a success if the patient survived longer than 30 days but died within 90 days of surgery. Consequently, as a result of these findings 90-day mortality is now advocated as a more appropriate measure of peri-operative mortality (46,47).

Clinical prediction models in thoracic surgery

The most commonly used clinical risk prediction model in the UK and across Europe is the Thoracoscore model (39). Indeed, the model is specifically referred to in both the NICE and BTS guidelines. Given its singular recognition by national and international guidelines, multiple validation studies have been undertaking, allowing for a detailed overview of the model's performance to be presented.

The Thoracoscore

The Thoracoscore was designed to predict in-hospital mortality and was developed from 15,183 patients undergoing all thoracic surgery procedures for both benign and malignant disease between 2002 and 2005 in multiple hospitals across France. The variables included in the final model are listed in *Table 1* (39).

The model was internally validated using a split sample approach and demonstrated excellent performance. The AUC was 0.85 for the training set (n=10,122) and 0.86 for the test set (n=5,061). Calibration, assessed using the H-L test, was also acceptable in both the training and testing sets. Based on these results, which were published in 2007, the Thoracoscore became well established in routine thoracic surgical practice.

Multiple validation studies of Thoracoscore have subsequently been undertaken and published. A UK-based single centre & single surgeon validation incorporating 290 patients undergoing lung resection between 2008 and 2011 was published in 2012 (48). In-hospital mortality was 3.1% (n=9). The AUC was 0.60, indicating poor discrimination.

No assessment of calibration was undertaken. Another UK-based single-centre analysis (49), which assessed 703 patients who underwent open lung resection between 2007 and 2010 had an in-hospital mortality rate of 2.0% (n=16). Model discrimination was poor (AUC 0.68), and model calibration was adequate when assessed using the H-L test (P=0.18) but demonstrated poor predictive ability when assessed using the R^2 method (P=0.028). Poullis *et al.* (50) also validated the model, in an additional single-centre UK-based study. In total, 2,574 patients underwent open lung resection for NSCLC between 2001 and 2011. In-hospital mortality was 2.3% (n=59). Model discrimination was poor (AUC 0.69), and model calibration was adequate (P=0.46). Qadri *et al.* (51) validated the model in a population of 243 patients who underwent pneumonectomy between 1998 and 2008 in a single UK centre. In-hospital mortality was 4.5% (n=11). Discrimination of the model in this cohort of patients was extremely poor, with an AUC of 0.44. Calibration of the model was not assessed.

Two multi-centre validation assessments of Thoracoscore have been undertaken. The study by Sharkey *et al.* (52) comprised 2,245 patients who underwent lung resection in six centres in the UK between 2011 and 2012. In-hospital mortality was 1.38% (n=31). Model discrimination was found to be acceptable (AUC 0.71) although model calibration was not assessed. O'Dowd *et al.* (53) also undertook a multi-centre validation of Thoracoscore. A total of 3,222 patients were included from the National Lung Cancer Audit (NLCA) database (lung resection for NSCLC between 2004 and 2012 in multiple UK centres). The AUC was 0.60, indicating poor discriminatory ability of the model for this cohort. Calibration was not assessed.

More recently, Die Loucou *et al.* (54) undertook an external validation of the Thoracoscore model and went on to develop an updated model, using the same predictors but altering the coefficients to reflect a difference in the distribution of contemporary patient characteristics in comparison to the characteristics of the cohort used for development of the baseline model. The cohort was comprised of 56,279 patients undergoing surgery for mediastinal, chest wall, pleural or lung disease across several European centres between 2016 and 2017. External validation of the original Thoracoscore model in this population of patients demonstrated excellent discrimination (AUC 0.80) but poor calibration. Performance of the updated Thoracoscore model when internally validated against the same cohort of patients was also excellent, with an AUC of 0.83 and acceptable calibration. This model has not yet been externally validated.

Additional thoracic surgery risk prediction models

A recent systematic review of models specifically developed to predict short-term mortality after thoracic surgery has been performed and identified 20 such models (55) of which 11 were designed specifically to predict mortality after lung resection for lung cancer. Overall, flaws in model development (assessed using the PROBAST risk of bias tool) (56) were identified in 17 of the 20 models. These drawbacks included inappropriate handling of missing data, conversion of continuous variables to categorical (thereby weakening the statistical power of the variable), an insufficient number of events relative to the number of variables, insufficient/inappropriate analysis of performance and lack of external validation.

Eight models had undergone external validation in ten different studies. Only one model reported acceptable measures of both discrimination and calibration. However, this was a single-centre study comprising only 155 patients with just eight deaths (57). An external validation study of six models identified from the systematic review has been performed in a cohort of 6,600 patients undergoing lung resection in two large UK centres between 2012 and 2018 (58). This validation study found that model performance was inadequate in five of the six models. The sixth model, the modified Eurolung model, lacked key clinically relevant variables. As a result, the study concluded that none of the models validated could be recommended for use in contemporary UK thoracic surgical practice.

Two risk prediction models identified in the systematic review were developed to predict 90-day mortality after lung resection. The NLCA model was developed from 10,991 patients who underwent lung resection for stage I–IIIA NSCLC between 2004 and 2010 in multiple UK centres (59). The 90-day mortality rate was 5.9% (n=647). No measures of discrimination or calibration were reported. In 2016, O'Dowd *et al.* (53) undertook internal and external validations of this model, both of which demonstrated inadequate discrimination with calibration not assessed. The video assisted thoracoscopic surgery (VATS) model was developed from 733 patients who underwent anatomical lung resection (lobectomy or segmentectomy) for stage I-II NSCLC via a VATS approach in a single UK centre between 2012 and 2016 with a 90-day mortality rate of 2.5% (n=18) (60). Discrimination was excellent (AUC 0.85) and

Table 2 Variables included in the RESECT-90 model

Variables	Categories
Age	Continuous variable
Gender	Male
	Female
PS score	Continuous variable
% predicted DLCO	Continuous variable
BMI	Continuous variable
Serum creatinine	Continuous variable
Pre-operative anaemia*	Yes
	No
Pre-operative arrhythmia	Yes
	No
Laterality	Left
	Right
Number of resected bronchopulmonary segments	Continuous variable
Approach	Open
	Minimally invasive
Diagnosis	Benign
	Malignant

*, classified according to the World Health Organisation definitions. PS, performance status; DLCO, diffusion capacity of the lung for carbon monoxide; BMI, body mass index.

calibration was adequate (H-L test, P=0.99).

RESECT-90

The NLCA model demonstrated poor statistical performance on external validation and the VATS model is both not applicable to other patient cohorts and statistically underpowered (only 18 deaths within 90 days of surgery). Additional models externally validated for their ability to predict 90-day mortality also failed to demonstrate acceptable model performance (58). However, this is not a particularly surprising finding, as this was not the endpoint for which these models were originally developed to predict. In order to address this shortcoming, the RESECT-90 model has been developed (61).

Derived from a contemporary population of 6,600 patients undergoing lung resection in two UK centres between 2012 and 2018, the 90-day mortality rate was 3.1% (n=204). The model was sufficiently powered and was developed using multivariable logistic regression. Twelve variables were included in the final model, as listed in *Table 2* (61).

Following internal validation with adjustment for optimism, the model performed well, with an AUC of 0.74 and acceptable measures of calibration (assessed using flexible calibration plots, calibration-in-the-large and calibration slope). Whilst these results are encouraging, external validation is required prior to recommending the model for use in routine clinical practice. A pan-UK project to externally validate the model is currently underway.

Conclusions

In the modern era, whilst surgery for lung cancer remains the best treatment modality for those patients who are suitable, the increasing number of effective non-surgical treatments means that risk prediction for lung cancer surgery has become increasingly important. Individual variables, such as lung function, comorbidity burden and functional status have all been demonstrated to be predictive of adverse outcomes after lung resection, and hence are well-represented in many models developed to predict peri-operative mortality.

Ninety-day mortality is emerging as a superior outcome metric for capturing peri-operative deaths and should be considered instead of traditional measures such as 30-day or in-hospital mortality. Of the many risk models developed to predict short-term mortality after lung resection, none can currently be recommended for use in clinical practice. This represents a major weakness in the pre-operative assessment of thoracic surgery patients.

Encouragingly, the recently published RESECT-90 model has been developed with appropriate statistical rigour and has demonstrated acceptable performance in predicting 90-day mortality after lung resection. If the current pan-UK external validation of the model upholds these encouraging preliminary results, the RESECT-90 model may emerge as a useful tool for risk prediction in patients undergoing lung cancer surgery in the UK.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Michael Shackcloth and Amer Harky) for the series "Lung Cancer Surgery" published in *Video-Assisted Thoracic Surgery*. The article has undergone external peer review.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://vats.amegroups.com/article/view/10.21037/vats-21-47/coif). The series "Lung Cancer Surgery" was commissioned by the editorial office without any funding or sponsorship. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Dela Cruz CS, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. Clin Chest Med 2011;32:605-44.
2. Arnold M, Rutherford MJ, Bardot A, et al. Progress in cancer survival, mortality, and incidence in seven high-income countries 1995-2014 (ICBP SURVMARK-2): a population-based study. Lancet Oncol 2019;20:1493-505.
3. Zhang C, Wang L, Li W, et al. Surgical outcomes of stage IV non-small cell lung cancer: a single-center experience. J Thorac Dis 2019;11:5463-73.
4. Jones GS, Baldwin DR. Recent advances in the management of lung cancer. Clin Med (Lond) 2018;18:s41-6.
5. Bahig H, Chen H, Louie AV. Surgery versus SABR for early stage non-small cell lung cancer: the moving target of equipoise. J Thorac Dis 2017;9:953-6.
6. Lim E, Baldwin D, Beckles M, et al. Guidelines on the radical management of patients with lung cancer. Thorax 2010;65 Suppl 3:iii1-27.
7. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018;391:1023-75.
8. Field JK, Duffy SW, Baldwin DR, et al. The UK Lung Cancer Screening Trial: a pilot randomised controlled trial of low-dose computed tomography screening for the early detection of lung cancer. Health Technol Assess 2016;20:1-146.
9. Crosbie PA, Balata H, Evison M, et al. Implementing lung cancer screening: Baseline results from a community-based "Lung Health Check" pilot in deprived areas of Manchester. Thorax 2019;74:405-9.
10. Blandin Knight S, Crosbie PA, Balata H, et al. Progress and prospects of early detection in lung cancer. Open Biol 2017;7:170070.
11. Ginsberg RJ, Rubinstein LV. Randomized trial of lobectomy versus limited resection for T1 N0 non-small cell lung cancer. Ann Thorac Surg 1995;60:615-23.
12. Chamogeorgakis T, Ieromonachos C, Georgiannakis E, et al. Does lobectomy achieve better survival and recurrence rates than limited pulmonary resection for T1N0M0 non-small cell lung cancer patients? Interact Cardiovasc Thorac Surg 2009;8:364-72.
13. Bedetti B, Bertolaccini L, Rocco R, et al. Segmentectomy versus lobectomy for stage I non-small cell lung cancer: A systematic review and meta-analysis. J Thorac Dis 2017;9:1615-23.
14. Nakahara K, Ohno K, Hashimoto J, et al. Prediction of postoperative respiratory failure in patients undergoing lung resection for lung cancer. Ann Thorac Surg 1988;46:549-52.
15. 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.
16. Magdeleinat P, Seguin A, Alifano M, et al. Early and long-term results of lung resection for non-small-cell lung cancer in patients with severe ventilatory impairment. Eur J Cardiothorac Surg 2005;27:1099-105.
17. Licker MJ, Widikker I, Robert J, et al. Operative Mortality and Respiratory Complications After Lung Resection for Cancer: Impact of Chronic Obstructive Pulmonary Disease and Time Trends. Ann Thorac Surg 2006;81:1830-7.

18. Ferguson MK, Siddique J, Karrison T. Modeling major lung resection outcomes using classification trees and multiple imputation techniques. Eur J Cardiothorac Surg 2008;34:1085-9.
19. Brunelli A, Refai M, Salati M, et al. Predicted versus observed FEV1 and DLCO after major lung resection: a prospective evaluation at different postoperative periods. Ann Thorac Surg 2007;83:1134-9.
20. Brunelli A, Sabbatini A, Xiume' F, et al. A model to predict the decline of the forced expiratory volume in one second and the carbon monoxide lung diffusion capacity early after major lung resection. Interact Cardiovasc Thorac Surg 2005;4:61-5.
21. Ferguson MK, Reeder LB, Mick R. Optimizing selection of patients for major lung resection. J Thorac Cardiovasc Surg 1995;109:275-81; discussion 281-3.
22. Ferguson MK, Gaissert HA, Grab JD, et al. Pulmonary complications after lung resection in the absence of chronic obstructive pulmonary disease: the predictive role of diffusing capacity. J Thorac Cardiovasc Surg 2009;138:1297-302.
23. Santini M, Fiorello A, Vicidomini G, et al. Role of diffusing capacity in predicting complications after lung resection for cancer. Thorac Cardiovasc Surg 2007;55:391-4.
24. Berry MF, Villamizar-Ortiz NR, Tong BC, et al. Pulmonary function tests do not predict pulmonary complications after thoracoscopic lobectomy. Ann Thorac Surg 2010;89:1044-51; discussion 1051-2.
25. Brunelli A, Refai MA, Salati M, et al. Carbon monoxide lung diffusion capacity improves risk stratification in patients without airflow limitation: evidence for systematic measurement before lung resection. Eur J Cardiothorac Surg 2006;29:567-70.
26. Ferguson MK, Vigneswaran WT. Diffusing capacity predicts morbidity after lung resection in patients without obstructive lung disease. Ann Thorac Surg 2008;85:1158-64; discussion 1164-5.
27. Salati M, Brunelli A. Risk Stratification in Lung Resection. Curr Surg Rep 2016;4:37.
28. Brunelli A, Kim AW, Berger KI, et al. Physiologic evaluation of the patient with lung cancer being considered for resectional surgery: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 2013;143:e166S-90S.
29. 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.
30. Fennelly J, Potter L, Pompili C, et al. Performance in the shuttle walk test is associated with cardiopulmonary complications after lung resections. J Thorac Dis 2017;9:789-95.
31. Win T, Jackson A, Groves AM, et al. Relationship of shuttle walk test and lung cancer surgical outcome. Eur J Cardiothorac Surg 2004;26:1216-9.
32. Win T, Jackson A, Groves AM, et al. Comparison of shuttle walk with measured peak oxygen consumption in patients with operable lung cancer. Thorax 2006;61:57-60.
33. 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.
34. Lee TH, Marcantonio ER, Mangione CM, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. Circulation 1999;100:1043-9.
35. Brunelli A, Varela G, Salati M, et al. Recalibration of the revised cardiac risk index in lung resection candidates. Ann Thorac Surg 2010;90:199-203.
36. Ferguson MK, Celauro AD, Vigneswaran WT. Validation of a modified scoring system for cardiovascular risk associated with major lung resection. Eur J Cardiothorac Surg 2012;41:598-602.
37. Brunelli A, Cassivi SD, Fibla J, et al. External validation of the recalibrated thoracic revised cardiac risk index for predicting the risk of major cardiac complications after lung resection. Ann Thorac Surg 2011;92:445-8.
38. NICE Guideline [NG122] Lung Cancer: diagnosis and management 2019.
39. 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.
40. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020;368:m441.
41. Van Calster B, Nieboer D, Vergouwe Y, et al. A calibration hierarchy for risk models was defined: from utopia to empirical data. J Clin Epidemiol 2016;74:167-76.
42. Pezzi CM, Mallin K, Mendez AS, et al. Ninety-day mortality after resection for lung cancer is nearly double 30-day mortality. J Thorac Cardiovasc Surg 2014;148:2269-77.
43. McMillan RR, Berger A, Sima CS, et al. Thirty-day mortality underestimates the risk of early death after major

resections for thoracic malignancies. Ann Thorac Surg 2014;98:1769-74; discussion 1774-5.
44. Damhuis RA, Wijnhoven BP, Plaisier PW, et al. Comparison of 30-day, 90-day and in-hospital postoperative mortality for eight different cancer types. Br J Surg 2012;99:1149-54.
45. NLCA. National Lung Cancer Audit Lung cancer clinical outcomes publication. 2019.
46. Bryant AS, Rudemiller K, Cerfolio RJ. The 30- versus 90-day operative mortality after pulmonary resection. Ann Thorac Surg 2010;89:1717-22; discussion 1722-3.
47. Taylor M, Grant SW, West D, et al. 90-day mortality: re-defining the peri-operative period after lung resection. Clinical Lung Cancer 2021;22:e642-5.
48. Barua A, Handagala SD, Socci L, et al. Accuracy of two scoring systems for risk stratification in thoracic surgery. Interact Cardiovasc Thorac Surg 2012;14:556-9.
49. Bradley A, Marshall A, Abdelaziz M, et al. Thoracoscore fails to predict complications following elective lung resection. Eur Respir J 2012;40:1496-501.
50. Poullis M, McShane J, Shaw M, et al. Prediction of in-hospital mortality following pulmonary resections: improving on current risk models. Eur J Cardiothorac Surg 2013;44:238-42; discussion 242-3.
51. Qadri SSA, Jarvis M, Ariyaratnam P, et al. Could Thoracoscore predict postoperative mortality in patients undergoing pneumonectomy? Eur J Cardiothorac Surg 2014;45:864-9.
52. Sharkey A, Ariyaratnam P, Anikin V, et al. Thoracoscore and European Society Objective Score Fail to Predict Mortality in the UK. World J Oncol 2015;6:270-5.
53. O'Dowd EL, Lüchtenborg M, Baldwin DR, et al. Predicting death from surgery for lung cancer: A comparison of two scoring systems in two European countries. Lung Cancer 2016;95:88-93.
54. Die Loucou J, Pagès PB, Falcoz PE, et al. Validation and update of the thoracic surgery scoring system (Thoracoscore) risk model. Eur J Cardiothorac Surg 2020;58:350-6.
55. Taylor M, Hashmi SF, Martin GP, et al. A systematic review of risk prediction models for perioperative mortality after thoracic surgery. Interact Cardiovasc Thorac Surg 2021;32:333-42.
56. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. Ann Intern Med 2019;170:W1-W33.
57. Chamogeorgakis T, Toumpoulis I, Tomos P, et al. External validation of the modified Thoracoscore in a new thoracic surgery program: prediction of in-hospital mortality. Interact Cardiovasc Thorac Surg 2009;9:463-6.
58. Taylor M, Szafron B, Martin GP, et al. External validation of six existing multi-variable clinical prediction models for short-term mortality in patients undergoing lung resection. Eur J Cardiothorac Surg 2021;59:1030-6.
59. Powell HA, Tata LJ, Baldwin DR, et al. Early mortality after surgical resection for lung cancer: an analysis of the English National Lung cancer audit. Thorax 2013;68:826-34.
60. Brunelli A, Dinesh P, Woodcock-Shaw J, et al. Ninety-Day Mortality After Video-Assisted Thoracoscopic Lobectomy: Incidence and Risk Factors 2017. Ann Thorac Surg 2017;104:1020-6.
61. Taylor M, Martin GP, Abah U, et al. Development and internal validation of a clinical prediction model for 90-day mortality after lung resection: the RESECT-90 score. Interact Cardiovasc Thorac Surg 2021;33:921-7.

doi: 10.21037/vats-21-47
Cite this article as: Taylor M, Whittaker G, Grant SW, Booton R, Shah R. Risk prediction for lung cancer surgery in the current era. Video-assist Thorac Surg 2022;7:10.