



Big data, big contributions: outcomes research in thoracic surgery

Benjamin J. Resio, Andrew P. Dhanasopon, Justin D. Blasberg

Section of Thoracic Surgery, Department of Surgery, Yale School of Medicine, New Haven, CT, USA

Contributions: (I) Conception and design: All authors; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: None; (V) Data analysis and interpretation: None; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Justin D. Blasberg, MD. 330 Cedar St. BB205, PO Box 208062, New Haven, CT, USA. Email: justin.blasberg@yale.edu.

Abstract: In recent years, analysis of registry data has defined clinically significant practice patterns and treatment strategies that optimize cancer care for thoracic surgery patients. These higher-order outcome studies rely on large patient cohorts that minimize the risk of selection bias and allow for a powered analysis that is not achievable with single- or multi-institutional data. This review uses recent study examples to highlight important contributions to our knowledge of thoracic surgery and describes how outcomes research using large data can address high impact clinical questions.

Keywords: Thoracic surgery; database; outcomes research

Submitted Dec 04, 2018. Accepted for publication Dec 31, 2018.

doi: 10.21037/jtd.2019.01.04

View this article at: <http://dx.doi.org/10.21037/jtd.2019.01.04>

Introduction

Lung cancer is the leading cause of cancer death in the United States, with approximately 200,000 new cases diagnosed each year and 175,000 deaths (1). Over the past two decades, significant improvements in preoperative imaging, minimally invasive techniques for resection, and expanded indications for adjuvant therapy have, stage for stage, improved survival and quality of life associated with surgical resection (1,2). As clinicians strive to improve care for lung cancer patients, large databases have emerged as important resources to provide insights and engender opportunities to ask clinically meaningful research questions. These efforts have resulted in improved surgical outcomes associated with data driven practice. While institutional studies of the past have been useful to define outcomes that were possible, large registries have become popular as a resource to describe outcomes that are current and probable.

To define best practice, utilization of cancer registries that provide insight into the practice patterns and associated outcomes for thoracic surgery patients are necessary. A handful of frequently utilized databases each have distinct advantages for thoracic surgery research (*Table 1*). The

National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program database provides a population-based sample of the United States and can be used to estimate changes in diagnosis, treatment, and prognosis over time (4). The American Cancer Society and the American College of Surgeon's National Cancer Database (NCDB) captures approximately 70% of cancer cases in the United States and provides patient level, granular data on the treatment and outcomes representing nationwide practice (3). The Society of Thoracic Surgeons general thoracic surgery database (STS-GTSD) focuses on patient specific metrics such as preoperative comorbidities, perioperative complications, and mortality (5). The STS-GTSD has also been combined with Medicare claims to analyze long term outcomes associated with treatment (6,7).

Recent studies using these registries have highlighted the strengths and limitations of retrospective outcomes research and have made significant contributions to our understanding of best practices for lung cancer. Much of this work relies on unique attributes associated with large data, including case diversity and large sample size. These data sources have proven useful in answering important clinical questions and confirming conclusions derived from smaller studies, such as the benefit of anatomic resection

Table 1 Advantages and disadvantages of discussed databases for thoracic surgery research

Database	Advantages	Disadvantages
National Cancer Database (NCDB) (3)	Captures ~70% of incident cancers in the US Granular tumor specific and treatment data including staging, chemotherapy, radiation, surgery Long term survival outcomes Includes comprehensive sociodemographic information	Only tracks overall survival No recurrence information Only tracks first course of treatment Only comorbidity information is Charlson comorbidity index No smoking status No performance status No pulmonary function tests
Surveillance Epidemiology and End Results (SEER) (4)	Captures ~34% of cancers in the US currently Cohort is designed to be nationally representative Tumor and staging characteristics Tracks overall and cancer specific long-term survival Cause of death is available	No smoking status No comorbidity data No chemotherapy data No pulmonary function tests Smaller sample than NCDB
SEER—Medicare	Longitudinal claims available, which can provide pre-op imaging, invasive mediastinal staging, readmission, death, chemotherapy or radiation, associated costs, additional comorbidities and performance status	Only patients with Medicare (mostly patients >65)
Society of Thoracic Surgeons General Thoracic Surgery (STS-GTSD) (5)	Includes performance status Includes pulmonary function Includes smoking status Detailed comorbidities Tumor staging information Surgeon level data Includes prior surgery, prior chemotherapy, prior radiation Includes surgical complications	No longitudinal follow up data (only 30-day outcomes)
STS—Medicare (6)	Longitudinal claims available, which can provide pre-op imaging, invasive mediastinal staging, readmission, death, chemotherapy or radiation, associated costs, additional comorbidities and performance status	Only patients with Medicare (mostly patients >65)

for early stage disease, the impact of adjuvant therapy on survival, and the association of patient specific metrics on outcomes. Large data has also been used to track changes in utilization, identify practice patterns over time, uncover disparities in care, and describe outcomes associated with rare events that would otherwise be difficult to study with smaller samples.

This review uses recent study examples to describe how outcomes research can address high impact clinical questions, demonstrates the methods employed by these

studies to minimize bias, and highlights recent contributions to our knowledge of thoracic surgery.

Review

Addressing important clinical questions

The adoption of minimally invasive lung resection has increased significantly over the past two decades, likely due to the established benefits of small incision surgery including reduced postoperative pain, shorter hospital

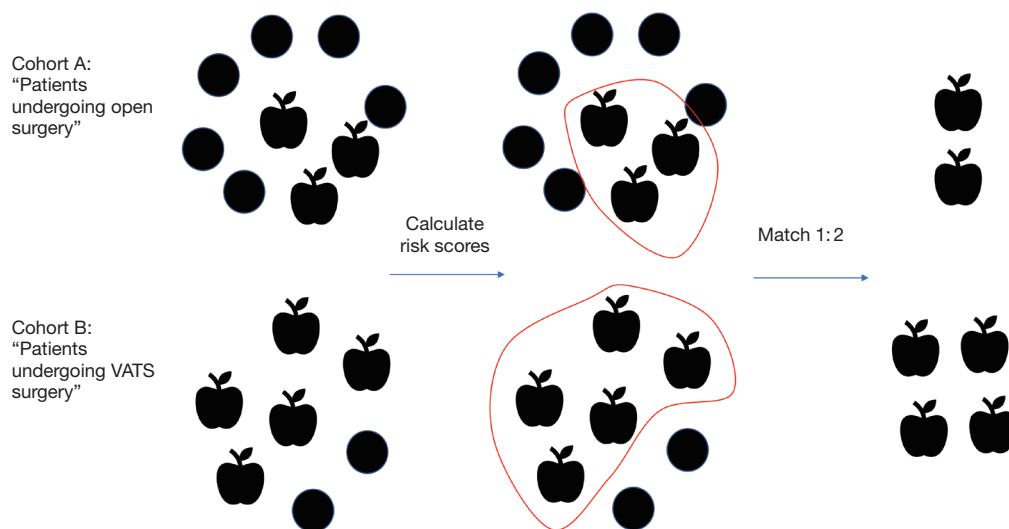


Figure 1 Propensity matching schematic. Propensity matching starts with two cohorts with uneven characteristics (VATS patients and open surgery patients in this example). Risk scores are calculated for each subject based on their characteristics and then patients with similar scores are matched in a specified ratio for analysis. Patients who do not match are excluded. VATS, video-assisted thoracoscopic surgery.

length of stay, and earlier resumption of daily activities (8). Concurrently, the oncologic equivalence and safety of minimally invasive resection has been a heavily debated topic, impacting surgeon preference and patient access to these less invasive surgical techniques. These differences in practice patterns have been particularly apparent in the Midwest and rural communities (8). Additionally, critics of minimally invasive resection have challenged the oncologic equivalence of video-assisted thoracoscopic surgery (VATS), including the thoroughness of lymph node dissection and the accuracy of pathologic staging (9,10). Initially founded in single institution data, the advantages of minimally invasive resection were not fully realized until the outcomes associated with large data were available to support a shift in practice. In the absence of large randomized trials, three studies derived from the STS-GTSD demonstrated that minimally invasive resection by VATS yielded comparable outcomes to open resection with regards to postoperative complication risk, the completeness of lymph node dissection, and long term survival (7,11,12). Although these studies have limitations inherent to their retrospective design, investigators gave careful thought to study methodology with the goal of reducing bias, thereby increasing the applicability of their findings.

The most recent of these retrospective studies (Minimally Invasive Lung Cancer Surgery Performed by Thoracic Surgeons as Effective as Thoracotomy) (7), used

a combination of STS-GTSD and Medicare claims to gather meaningful surgical characteristics linked to long term outcomes and demonstrated a common method to mitigate selection bias. A specific criticism of retrospective studies that compare surgical technique is that surgeons consciously selected distinct groups of patients for different procedures. Therefore, comparing the equivalence or superiority of a particular technique is not methodologically sound. Likewise, in the aforementioned study, investigators found that patients who underwent open resection tended to be worse surgical candidates and had comorbidities associated with reduced long-term survival (higher BMI, higher ASA class, marginal pulmonary function, and larger tumors). Comparing patient groups that are fundamentally different would undoubtedly bias the results of any comparison of outcomes. However, investigators were able to select comparable paired cohorts from each arm by using propensity matching. Propensity matching selects subjects of similar risk for a certain outcome from distinct populations by estimating and compiling the effects of all individual characteristics linked to that outcome, therefore pairing subjects to make them exchangeable (propensity matching schematic, *Figure 1*). Thinking carefully about possible confounders, these authors also identified surgeon training status (board certified thoracic surgeon versus general surgeon) as an important determinant of choice for operative approach, restricting their study group to patients

managed by board certified thoracic surgeons. In effect, this reduced the impact of general surgeon preference for open surgery and eliminated potential confounding.

After propensity matching, two cohorts of more than 1,000 patients were defined from which a Kaplan-Meier survival analysis demonstrated the benefits of VATS resection (7). Importantly, these conclusions implied that any possible deficits of lymph node dissection previously associated with VATS did not independently impact long-term outcomes, therefore supporting further adoption of minimally invasive resection.

Another important clinical question that has been evaluated by registry data is whether extent of resection (lobectomy versus sub-lobar resection) is associated with equivalent long-term outcomes (13-16). Veluswamy *et al.* leveraged the large sample size and tumor characteristics defined within the SEER-Medicare database to address this question, identifying over 3,000 eligible non-small cell lung cancer (NSCLC) patients with 10-year outcome data (13). Linkage of SEER with Medicare was crucial to this analysis as claims data were used to capture important patient specific metrics including comorbidity and performance status data. This study confirmed previous work that limited resection was inferior to lobectomy, particularly after adjusting for risk factors using propensity scoring (13). An additional follow up study in SEER clarified the survival advantage associated with anatomic resection, demonstrating that outcomes of ~500 segmentectomy patients propensity matched to ~1,000 lobectomy patients were similar for stage 1A disease (tumors between 1 and 2 cm in size) (16). Clinically, these retrospective studies have served as the foundation for current clinical trials, including the randomized controlled trial of lobar versus sub-lobar resection (CALGB 140503), (17). Additionally, these results have helped to inform surgeons about the risks and benefits associated with sub-lobar resection, and ultimately have changed consensus guidelines regarding the restricted use of wedge for select high-risk patients (18).

Registry data has also addressed important questions within healthcare delivery. As rates of readmission and hospital length of stay have become important metrics to define quality, critics have questioned whether fast-track discharge protocols are truly associated with equivalent short-term outcomes and reduced costs. Recent studies of both the Healthcare Cost and Utilization Project Nationwide Readmission Database (NRD) and the NCDB have evaluated the impact of expedited discharge on these important outcome metrics (19,20). Specifically, Jean and

colleagues evaluated over 100,000 patients who underwent lobectomy from 2010 to 2014, demonstrating that expedited discharge (1–3 days) had a 3% risk adjusted decrease in readmission and a \$4,000 lower index hospitalization cost compared to routine (4–7 days) or delayed (8+ days) discharge (20). These results and other similar studies reinforce the need for healthcare systems to evaluate cost and utilization when addressing quality associated with best practice guidelines.

The “Silver Standard”: the importance of big data when there is no RCT

The infrastructure, costs, and equipoise needed to study and define best practice is challenging, particularly because many clinically relevant questions cannot be answered with a randomized controlled trial (RCT). Retrospective studies fill a critical knowledge gap for these questions, and new hypotheses derived from large data can become important starting points for clinical trials. This is particularly true when attempting to define treatment recommendations for clinical scenarios without strong prospective data, including indications for adjuvant therapy in patients with T3 satellite nodules or the comparability of SBRT versus surgical resection in high-risk patients (21-24). Variability in treatment within cancer registries provides an opportunity to review data in these important patient subsets and to define best practice.

The 7th edition of the lung cancer staging system redefined NSCLC patients with a satellite nodule within the same lobe as T3 disease, down-staged from T4 in the prior edition (25). Unlike other T3 subtypes (tumor size or local invasion) where consensus recommendations for adjuvant therapy are clear, the role of chemotherapy to improve long-term survival in T3 satellite disease was not well defined. Salazar and colleagues attempted to clarify these indications from retrospective data, thoughtfully creating multiple cohorts to study the benefit of chemotherapy in T3 satellite patients (*Figure 2*) (22). Investigators started with over 1,000 treatment naïve patients who underwent surgical resection for pT3 NSCLC. The authors were careful to exclude tumor subtypes associated with a favorable prognosis such as bronchioloalveolar carcinoma and carcinoid tumors, competing malignancies, treatment with non-standard chemotherapy, and patients who received delayed chemotherapy. In this first cohort, a 10% relative survival advantage over 3 years was identified for T3 satellite patients treated with chemotherapy compared to

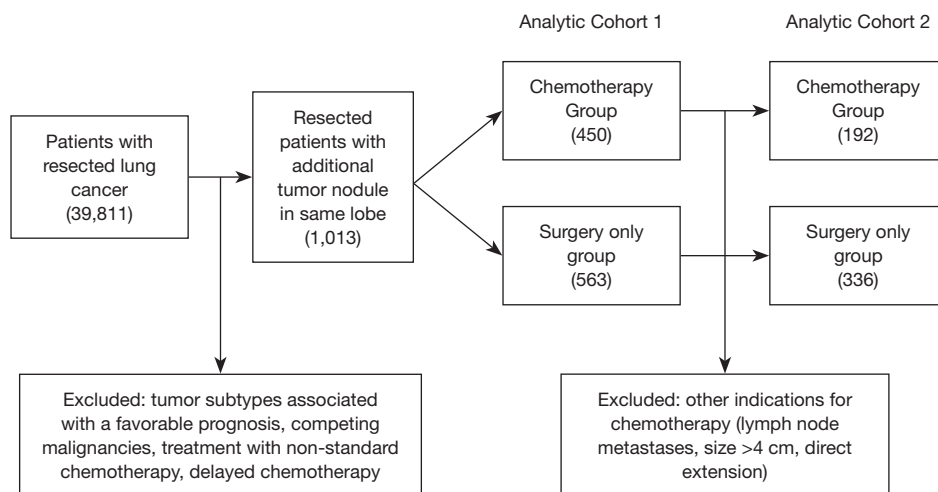


Figure 2 Schematic of study cohort selection steps from Salazar *et al.* (22).

those who were not.

Salazar and colleagues also recognized that there may be additional tumor-specific factors that impact the decision to treat with chemotherapy, creating bias in the analysis. For example, patients with tumors within 2 cm of the carina may receive adjuvant chemotherapy, independent of the presence of a satellite nodule. Therefore, investigators created a more selective cohort excluding all other indications for chemotherapy (lymph node metastases, size >4 cm, direct extension, etc.). This more focused analysis confirmed the broader findings of a survival benefit associated with chemotherapy, giving strength to the argument that adjuvant therapy was advantageous for T3 satellite nodule patients (22).

Another example where retrospective studies have filled the knowledge gap is in the domain of optimal treatment for high-risk patients. Investigators have previously designed RCT's to define the best local treatment option for medically inoperable patients with tumors <3 cm (specifically, radiation versus sublobar resection). All of these trials were closed due to poor accrual. Insight into the clinical utility of stereotactic body radiation therapy (SBRT) versus resection was therefore derived from three studies performed using SEER-Medicare. The first showed that propensity matched elderly patients treated with lobectomy or SBRT had equivalent long-term survival (14). A subsequent study also using propensity matching confirmed similar survival for SBRT versus wedge resection, and additionally found a survival benefit for anatomic resection (segmentectomy) (26). The third study clarified these

two analyses by stratification according to predicted life expectancy, demonstrating that patients at risk for mortality within 5 years of their cancer diagnosis (defining the high-risk cohort) had similar survival regardless of treatment (27). These studies demonstrate an important consideration when performing retrospective research: within a population, sub-groups may have different outcomes that are not identifiable until the data is stratified by clinically relevant factors potentially associated with outcome (i.e., healthier, elderly patients benefit differently from surgery versus SBRT than elderly patients as a whole). Clinically, these conclusions also support 2 important points: (I) high risk patients that have competing causes of death likely benefit from a local treatment strategy of lesser risk given they are unlikely to die from their lung cancer and (II) less invasive local therapy (SBRT) is associated with a survival benefit that is likely similar to surgical resection in the setting of competing causes of mortality. These ideals have changed clinical practice and optimized treatment paradigms for select patients previously managed by surgery alone.

Tracking utilization and identifying disparities

The ability to track treatments and outcomes over time and across populations, identifying changing trends in practice and disparities in care, is a fundamental advantage associated with large broadly representative data (28,29). A notable example comes from the role of surgical resection for small cell lung cancer (SCLC). Although SCLC is often metastatic at presentation, specific patients with

limited disease may benefit from surgical resection. These recommendations are based on NCCN guidelines, defining the role of surgery for clinically node-negative T1/T2 disease (30). However, the utilization of surgical resection for early stage SCLC is variable based on patient and hospital characteristics. Wakeam *et al.* evaluated SCLC practice patterns using the NCDB and identified that surgery is underutilized (29). The authors demonstrated that although the number of surgical procedures for SCLC increased from 2004 to 2013, approximately 2/3 of eligible patients did not undergo resection (29). These findings and similar studies provide a broad view of care across the country that may be overlooked by institutional series derived from well-resourced but biased data. Specifically, the Wakeam study provided critical insight into the systematic challenges associated with complex lung cancer care, as well as opportunities to improve the availability of evidence-based practice, quality, and efficiency.

Providing insight into rare events

Rare events are, by definition, too infrequent to study from institutional data given the power needed to obtain statistically valid conclusions. Registry data has been successfully used to define outcomes associated with rare lung cancer types that have unique biology and treatment considerations (31-34). For example, Mucoepidermoid Carcinoma (MEC) is a rare indolent lung cancer thought to arise from submucosal glands within the airways. Small institutional studies have variously identified risk factors associated with this cancer subtype and particularly debated the prognostic significance of tumor grade on survival (35,36). However, until recently, no large series were available to define optimal treatment due to the rarity of this disease. Additionally, previous data were confounded by the fact that, over time, practice patterns and treatments have evolved. Recent studies from the NCDB have defined expectations for long-term survival and the optimal management of MEC patients by utilizing 'relative survival analysis', a methodology that helps investigators define the natural history of indolent cancers (31).

This use of 'relative survival' becomes important in the study of indolent tumors because the NCDB tracks overall survival but not cancer specific survival. Therefore, non-cancer related causes of death can be important analysis confounders in patients with indolent tumors. In other words, because cancer is the most likely cause of death for patients with aggressive tumors, overall survival is likely

to be the same as cancer specific survival. However, in less aggressive (indolent) tumors, non-cancer causes of death predominate and overall survival becomes an inferior (potentially biased) surrogate for cancer specific survival. In analyzing the long-term survival of MEC, Resio and colleagues used overall survival data from the NCDB to calculate a more relevant 'relative survival analysis' (31), derived by comparing patients with MEC to age and sex matched counterparts based on US life tables from the National Institute of Health. This data yielded a ratio describing the likelihood that a patient with MEC would survive compared to the likelihood that a similar person without MEC would survive, which was used as a surrogate for cancer specific mortality. Using 'relative survival', Resio and colleagues demonstrated that low grade MEC tumors had a favorable 8-year survival after resection (90%) compared to high grade tumors (50%). Overall this analysis demonstrates the value and limitations of working with large datasets to analyze rare events, and that unique methods to compare patients thought to be relatively similar may identify clinically relevant attributes that impact care.

Conclusions

Large patient registries have emerged as a valuable tool to study outcomes associated with the surgical management of patients with lung cancer. Advantages of these databases include well-powered and diverse samples and inclusion of a wide spectrum of clinically impactful data. However, there are many challenges and limitations inherent to retrospective data that should be carefully considered. The above examples demonstrate how thoughtfully planned methodology can help authors address the limitations and bias which can confound analyses and subsequently derive clinically relevant conclusions. Indeed, studies from large data registries have provided insight into critical questions that have defined what is considered standard of care in thoracic surgery (minimally invasive versus open thoracic surgery) and inspired important clinical trials (lobar versus sub-lobar resection, SBRT versus resection). These studies have also defined variations in care and enlightened our understanding of rare events. Looking to the future, there is an expectation that thoracic surgery will continue to evolve with higher rates of adoption of minimally invasive techniques (VATS and robotics), greater penetrance of neoadjuvant therapy for patients with earlier stage disease, expanded use of immunotherapy, and increased identification of driver mutations with actionable targets.

As these practices evolve, the role of cancer registry data will be an important resource from which to evaluate and redefine best practice for lung cancer patients.

Acknowledgements

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* 2018;68:7-30.
2. Kapadia NS, Valle LF, George JA, et al. Patterns of Treatment and Outcomes for Definitive Therapy of Early Stage Non-Small Cell Lung Cancer. *Ann Thorac Surg* 2017;104:1881-8.
3. Boffa DJ, Rosen JE, Mallin K, et al. Using the National Cancer Database for Outcomes Research: A Review. *JAMA Oncol* 2017;3:1722-8.
4. SEER Incidence Database - SEER Data & Software [Internet]. SEER. [cited 2018 Nov 28]. Available online: <https://seer.cancer.gov/data/index.html>
5. STS National Database | STS [Internet]. [cited 2018 Nov 28]. Available online: <https://www.sts.org/registries-research-center/sts-national-database>
6. Fernandez FG, Furnary AP, Kosinski AS, et al. Longitudinal Follow-up of Lung Cancer Resection From the Society of Thoracic Surgeons General Thoracic Surgery Database in Patients 65 Years and Older. *Ann Thorac Surg* 2016;101:2067-76.
7. Boffa DJ, Kosinski AS, Furnary AP, et al. Minimally Invasive Lung Cancer Surgery Performed by Thoracic Surgeons as Effective as Thoracotomy. *J Clin Oncol* 2018;36:2378-85.
8. Blasberg JD, Seder CW, Levenson G, et al. Video-Assisted Thoracoscopic Lobectomy for Lung Cancer: Current Practice Patterns and Predictors of Adoption. *Ann Thorac Surg* 2016;102:1854-62.
9. Licht PB, Jørgensen OD, Ladegaard L, et al. A national study of nodal upstaging after thoracoscopic versus open lobectomy for clinical stage I lung cancer. *Ann Thorac Surg* 2013;96:943-9; discussion 949-50.
10. Medbery RL, Gillespie TW, Liu Y, et al. Nodal Upstaging Is More Common with Thoracotomy than with VATS During Lobectomy for Early-Stage Lung Cancer: An Analysis from the National Cancer Data Base. *J Thorac Oncol* 2016;11:222-33.
11. Boffa DJ, Kosinski AS, Paul S, et al. Lymph Node Evaluation by Open or Video-Assisted Approaches in 11,500 Anatomic Lung Cancer Resections. *Ann Thorac Surg* 2012;94:347-53.
12. Boffa DJ, Dhamija A, Kosinski AS, et al. Fewer complications result from a video-assisted approach to anatomic resection of clinical stage I lung cancer. *J Thorac Cardiovasc Surg* 2014;148:637-43.
13. Veluswamy RR, Ezer N, Mhango G, et al. Limited Resection Versus Lobectomy for Older Patients With Early-Stage Lung Cancer: Impact of Histology. *J Clin Oncol* 2015;33:3447-53.
14. Shirvani SM, Jiang J, Chang JY, et al. Lobectomy, sublobar resection, and stereotactic ablative radiotherapy for early-stage non-small cell lung cancers in the elderly. *JAMA Surg* 2014;149:1244-53.
15. Khullar OV, Liu Y, Gillespie T, et al. Survival After Sublobar Resection versus Lobectomy for Clinical Stage IA Lung Cancer: An Analysis from the National Cancer Data Base. *J Thorac Oncol* 2015;10:1625-33.
16. Zhao ZR, Situ DR, Lau RWH, et al. Comparison of Segmentectomy and Lobectomy in Stage IA Adenocarcinomas. *J Thorac Oncol* 2017;12:890-6.
17. Fox N, Bauer T. CALGB 140503: A Randomized Phase III Trial of Lobectomy versus Sublobar Resection for Small (< 2cm) Peripheral Non-Small Cell Lung Cancer. *Oncol Issues* 2008;23:20-1.
18. Donington J, Ferguson M, Mazzone P, et al. American College of Chest Physicians and Society of Thoracic Surgeons Consensus Statement for Evaluation and Management for High-Risk Patients With Stage I Non-small Cell Lung Cancer. *Chest* 2012;142:1620-35.
19. Rosen JE, Salazar MC, Dharmarajan K, et al. Length of Stay From the Hospital Perspective: Practice of Early Discharge Is Not Associated With Increased Readmission Risk After Lung Cancer Surgery. *Ann Surg* 2017;266:383-8.
20. Jean RA, Chiu AS, Boffa DJ, et al. Delayed discharge does not decrease the cost of readmission after pulmonary lobectomy. *Surgery* 2018;164:1294-9.
21. Hancock JG, Rosen JE, Antonicelli A, et al. Impact of adjuvant treatment for microscopic residual disease after non-small cell lung cancer surgery. *Ann Thorac Surg* 2015;99:406-13.
22. Salazar MC, Rosen JE, Arnold BN, et al. Adjuvant

- Chemotherapy for T3 Non-Small Cell Lung Cancer with Additional Tumor Nodules in the Same Lobe. *J Thorac Oncol* 2016;11:1090-100.
23. Boffa DJ, Hancock JG, Yao X, et al. Now or Later: Evaluating the Importance of Chemotherapy Timing in Resectable Stage III (N2) Lung Cancer in the National Cancer Database. *Ann Thorac Surg* 2015;99:200-8.
 24. Gao SJ, Corso CD, Blasberg JD, et al. Role of Adjuvant Therapy for Node-Negative Lung Cancer Invading the Chest Wall. *Clin Lung Cancer* 2017;18:169-77.e4.
 25. Goldstraw P, Crowley J, Chansky K, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. *J Thorac Oncol* 2007;2:706-14.
 26. Ezer N, Veluswamy RR, Mhango G, et al. Outcomes after Stereotactic Body Radiotherapy versus Limited Resection in Older Patients with Early-Stage Lung Cancer. *J Thorac Oncol* 2015;10:1201-6.
 27. Yu JB, Soulos PR, Cramer LD, et al. Comparative effectiveness of surgery and radiosurgery for stage I non-small cell lung cancer. *Cancer* 2015;121:2341-9.
 28. Ost DE, Niu J, Elting LS, et al. Determinants of practice patterns and quality gaps in lung cancer staging and diagnosis. *Chest* 2014;145:1097-113.
 29. Wakeam E, Varghese TK, Leighl NB, et al. Trends, practice patterns and underuse of surgery in the treatment of early stage small cell lung cancer. *Lung Cancer* 2017;109:117-23.
 30. Kalemkerian GP, Akerley W, Bogner P, et al. Small Cell Lung Cancer. *J Natl Compr Canc Netw* 2011;9:1086-113.
 31. Resio BJ, Chiu AS, Hoag J, et al. Primary Salivary Type Lung Cancers in the National Cancer Database. *Ann Thorac Surg* 2018;105:1633-9.
 32. Arnold BN, Thomas DC, Rosen JE, et al. Lung Cancer in the Very Young: Treatment and Survival in the National Cancer Data Base. *J Thorac Oncol* 2016;11:1121-31.
 33. Johnson KK, Rosen JE, Salazar MC et al. Outcomes of a Highly Selective Surgical Approach to Oligometastatic Lung Cancer. *Ann Thorac Surg* 2016;102:1166-71.
 34. Rahouma M, Kamel M, Narula N, et al. Pulmonary sarcomatoid carcinoma: an analysis of a rare cancer from the Surveillance, Epidemiology, and End Results database. *Eur J Cardiothorac Surg* 2018;53:828-34.
 35. Hsieh CC, Sun YH, Lin SW et al. Surgical outcomes of pulmonary mucoepidermoid carcinoma: A review of 41 cases. *PLoS One* 2017;12:e0176918.
 36. Komiya T, Perez RP, Yamamoto S, et al. Primary lung mucoepidermoid carcinoma: Analysis of prognostic factors using Surveillance, Epidemiology, and End Results Program (SEER). *Clin Respir J* 2017;11:847-53.

Cite this article as: Resio BJ, Dhanasopon AP, Blasberg JD. Big data, big contributions: outcomes research in thoracic surgery. *J Thorac Dis* 2019;11(Suppl 4):S566-S573. doi: 10.21037/jtd.2019.01.04