Quality and patient safety in surgery: Clinical applications and critical appraisal of a prospective, standardized, and comprehensive system for monitoring and reporting post-operative adverse events

Jelena Ivanovic

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the Doctorate of Philosophy degree in Epidemiology

Department of Epidemiology and Community Medicine

Faculty of Medicine

University of Ottawa

THESIS ADVISORY COMMITTEE

Thesis Supervisor:	Dr. Andrew J.E. Seely, MD, PhD, FRCSC
Thesis Co-Supervisor:	Dr. Timothy Ramsay, PhD
Thesis Advisory Committee:	Dr. Doug Coyle, PhD
	Dr. Donna E. Maziak, MD, MSc, FRCSC, FACS
	Dr. David Schramm, MD, MSc, FRCSC
	Dr. R. Sudhir Sundaresan, MD, FRCSC, FACS
Thesis Examination Panel:	Dr. Rodney Breau, MD, MSc, FRCSC
	Dr. Jamie Brehaut, PhD
	Dr. Lisa Calder, MD, MSc, FRCPC
	Dr. Liane S. Feldman, MD, FRCSC, FACS

TABLE OF CONTENTS

ABSTRACT

Evaluation of quality of surgical care begins with the Donabedian triad focusing on structure, process, and outcomes. Outcomes, which are inherently patient-centered, are most easily and commonly measured, and are indeed fundamental to evaluating the quality of surgical care. Specifically, post-operative adverse events (AEs) remain the most frequently measured and reported outcomes, as they represent harm to the patient; and thus, are often used as a means for comparing institutional, as well as, individual surgeon performance. The importance of rigorous recording of clearly defined AEs, although widely recognized, is poorly performed in practice.

In previous work, created in accordance to the Clavien-Dindo classification, we developed and integrated a classification of Thoracic Morbidity & Mortality (TM&M) within The Ottawa Hospital's Division of Thoracic Surgery allowing objective and standardized assessment of all post-operative AEs following all surgeries. In this thesis, the complementary studies that were conducted surrounding the continued clinical application and critical appraisal of the TM&M classification system as a means toward quality improvement are described.

Using standardized reporting of both incidence and severity of post-operative complications, we first provide an overview of the burden and distribution that the two most pervasive post-operative AEs have on the thoracic surgical patient population, including prolonged alveolar air leak and atrial fibrillation (**Chapter I and II**). Next, we explore the inter-system reliability of reported AEs following thoracic surgery from the American College of Surgeons-National Surgical Quality Improvement Program (ACS-NSQIP), which is widely considered the most prominent surgical quality improvement effort, and the TM&M classification system in order to better understand to what extent the methods used to collect data may be impacting results (**Chapter III**). The disparity between the two systems and the duplicate participation indicates distinct value to the two quality reporting systems.

An absence of evidence in the literature regarding individual surgeon outcome reporting and its impact on the quality of care prompted us to create risk-adjusted, surgeon-specific outcome reports to enable individualized performance measurement and feedback (**Chapter IV**). A priority for the division has been to ensure such measurement translates into reproducible improvements in surgical performance. To do so, we implemented complementary continuous quality improvement seminars to provide an additional forum for discussion regarding collective results, utilizing positive deviance, to unmask best performers as a catalyst for discussing practice measures to improve specific AEs.

Lastly, an evolutionary understanding of the heterogeneity of TM&M data was considered as a critical next step to following improvements in care (**Chapter V**). Recognizing that software was necessary to efficiently record and review TM&M data, iterative development led to an evolution of a real-time, web-based, point-of-care Thoracic Surgery Quality monitoring, Information management, and Clinical documentation (TSQIC) software system. The TSQIC system has enabled bedside data recording and storage, and automated dynamic analysis and reporting of surgical volume and quality.

We observe that measurement of TM&M data alone, while necessary, is not sufficient for quality improvement. We suggest that in addition to implementing a complementary point-of-care, interactive, web-based quality monitoring system, key factors for improving quality and patient safety include a combination of temporal analyses of AEs, effective surgeon-specific feedback mechanisms, actionable information based on best practice measures, standardization of case reviews, and a unit-based approach conducive of team-work and safety culture, led by open and collegial dialogue.

ABBREVIATIONS & ACRONYMS

% DLCO	Diffusing capacity of the lung for carbon monoxide
% FEV_1	Forced expiratory volume exhaled at the end of first second
ACS-NSQIP	American College of Surgeons-National Surgical Quality Improvement Program
AEs	Adverse Events
CI	Confidence Interval
CQI	Continuous Quality Improvement
CT	Computed Tomography
EVAD Score	Expiratory Volume, Age, and Diffusing capacity Score
FEV_1/FVC	Forced expiratory volume in 1 second/Forced vital capacity ratio
FVC	Forced Vital Capacity
IT	Information Technology
LOS	Length of Stay
M&M	Morbidity & Mortality
O/E	Observed/Expected Ratio
OR	Odds Ratio
PAAL	Prolonged Alveolar Air Leak
PAF	Post-operative Atrial Fibrillation
PD	Positive Deviance
PFT	Pulmonary Function Test
POD	Post-operative Day
SSORs	Surgeon-Specific Outcome Reports
TM&M	Thoracic Morbidity & Mortality
TOH	The Ottawa Hospital

TABLES

FIGURES

INTRODUCTION & BACKGROUND

Reliable and reproducible evaluation of quality of surgical care is of utmost importance for governments, hospitals, clinicians, and patients. Given the elective nature of the majority of surgical care, and the immediacy of its impact, surgeons have built a strong culture of patient safety, and a longstanding tradition of quality assessment and peer-review. Building upon this foundation, there is increasing focus on standardization of means to evaluate surgical quality, developed by associations, institutions, and individual centers.

The overall aim of this thesis is to review these means, focusing on an individual division of thoracic surgery, its efforts to standardize the evaluation of surgical quality, and develop a system to monitor it over time. While the tools developed may or may not be immediately relevant, it is hoped that the resources and principles outlined are useful in the widespread development of surgical quality assessment programs.

Defining quality of care

Before evaluation can begin, it is necessary to first have a clear definition of quality care. Quality of care has been defined in a number of ways. Definitions of quality are either generic or multidimensional. Using a generic definition, the Institute of Medicine defines quality as the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge.[1] Generic definitions are not easily actionable and trade both sensitivity and specificity for generalizability.[2]

Multidimensional approaches, on the other hand, recognise that quality is made up of a complex interplay of different domains.[3,4] For example, Campbell and colleagues propose that there are only two domains of quality - access and effectiveness. First, do patients get the care they need, and second, once care is received, is it effective?[2]

Framework for the evaluation of surgical quality

Donabedian advocates that the evaluation of quality in health care entails the assessment of specific and measurable elements of practice that are amenable to quantification.[3] Donabedian has further opined that the information from which assumptions can be made regarding the quality of care can be arranged under a multidimensional framework consisting of: structure, process, and outcome. The three domains are interrelated, as quality assessment is only possible because good structure increases the likelihood of good process, and good process increases the likelihood of good outcome.[3] This triad approach provides a useful means to classify, evaluate, and monitor the quality of surgical care (**Figure 0.1**), and provides a foundation upon which this thesis was built.

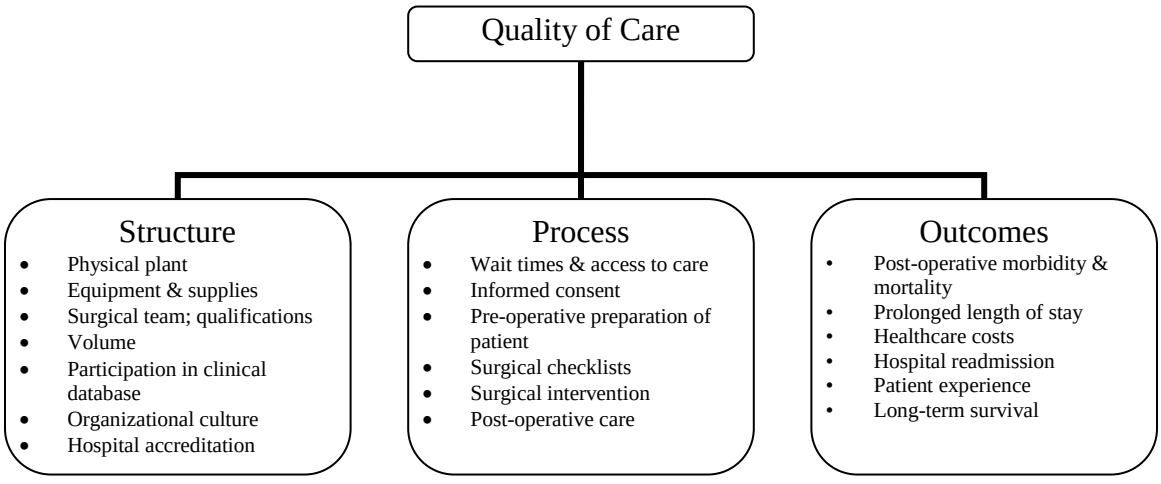


Figure 0.1 Adapting the Donabedian triad: Dimensions of quality of surgical care.

i. Structure

Structure refers to the characteristics of the settings in which health care occurs.[3] The organization's structure can be rigid or dynamic, depending on the culture, the need for change and improvement, and the availability of resources.[5] Structural measures include the physical plant, the equipment and supplies, knowledge and organization of healthcare personnel, and hospital accreditation.[6] Participation in a clinical database with state/provincial, regional, or national representation that provides regular performance reports based on benchmarked data has also garnered increasing attention as a structural measure of quality. The effect of surgical specialization and surgical case volume on outcomes have been of special interest to surgeons. The volume/outcome relationship postulates that a higher caseload and specialization results in an improved outcome.[7] It is important to note that many structural measures are not readily actionable, which limit their ultimate effectiveness as a means toward quality improvement.[7]

ii. Process

Process refers to the act of giving and receiving care.[3] In surgery, important process measures include: wait times and access to care, informed consent, the preoperative preparation of the patient, the choice of the surgical intervention and its execution, use of preoperative checklists, and routine post-operative care.[5-7] Many processes of care are strongly associated with improved patient outcomes because they allow for more specific actions geared towards quality improvement.[7] When strictly applied, such processes can have a significant effect on large patient populations.[5] Although process measures are actionable, they are however difficult to measure reliably on a routine basis given the diversity of surgical procedures.[5]

iii. Outcomes

Outcomes measurement has been fundamental in evaluating the quality of surgical care.[8] The measurement of surgical outcomes dates back to the early 20th century from the work of Codman.[9] Surgical outcomes are the most discussed, most cited, and arguably the most relevant of the three domains of care. Essentially, outcome signifies the effects of care on the health status of patients.[3] Examples of commonly used outcomes in surgery include post-operative adverse events (AEs), such as post-operative morbidity & mortality (M&M), prolonged hospital length of stay, hospital readmission, long-term survival, healthcare costs, and patient experience – including patient satisfaction, functional health status, and other measures of health-related quality of life. Of these, post-operative AEs rates remain the most frequently measured and reported outcomes. As such, data on post-operative AEs rates are often used as a means of comparing surgical techniques, individual surgeon outcomes, and institutional performance.[8] In this thesis, emphasis is placed on the evaluation of surgical AEs, with a focus on the standardized reporting and monitoring of post-operative M&M.

Defining adverse events and other negative outcomes of surgical procedures

Before evaluation can begin, it is important to first define and differentiate all negative outcomes throughout the perioperative process, as strategies to prevent them will vary widely. An *error* thus is defined as the failure of a planned action to be completed as originally intended (error of execution) or the use of the wrong plan to achieve an aim (error of planning).[10] Errors can further be classified as either latent or active. *Active errors* are those committed by individual practitioners at the point-of-care. *Latent errors* are circumstances established by policies and practices of an institution, culture, or society that predispose clinicians to errors. Errors can

however be intercepted by appropriate action that minimizes the threat to patient safety.[11] A *close call* then is an event that almost leads to patient harm but is prevented as a result of timely interception.[12]

An AE, or complication, is any unintended result of medical treatment that results in prolonged hospital stay, morbidity, or mortality; it may also be an injury caused by individual error or medical management rather than by the underlying condition of the patient.[13] If an AE is caused by an error, it is preventable.[14] Post-operative mortality is defined as either in-hospital mortality, 30-day mortality, or a combination of the two.[15] Post-operative mortality is the easiest variable to quantify; however, it can only be used as a measure of quality if it occurs frequently enough to yield statistical power.[16] In contrast to post-operative mortality, the rate of post-operative morbidity is high enough to allow for its use as a measure of quality of care in all specialities and after all major operations.[15]

A sequela, or an after-effect, results from the surgical procedure, but appears after recovery and is persistent.[17] Examples of surgical sequela are scars, or the reduction in pulmonary function after pneumonectomy. However, some AEs may contain elements of both sequela and complications, such as severe post-operative pain, post-operative paralytic ileus after abdominal surgery, and need for intensive care.[17]

Surgical procedures may be well executed and uncomplicated, yet result in *failure*. The most common example of a failure of surgical therapy is recurrence of a tumour.[17]

Although all negative outcomes should be recorded and reported to assess the quality of surgical care and enhance patient safety, the remainder of this thesis will be limited to post-operative AEs.

Pervasive problem of post-operative adverse events and their relation to quality and patient safety

Patient safety refers to minimizing the harm to patients arising from administered treatments, or simply, freedom from accidental injury.[13] According to the 1999 Institute of Medicine Report, *To Err is Human*, errors in health care are the eighth leading cause of death in the United States and account for up to 100,000 deaths annually.[13] Highest rates of AEs with serious consequences are most likely to occur in intensive care units, emergency departments, and *operating rooms*.[13] In support of this, Gawande and colleagues examined 15,000 patient charts and found that 66% of all AEs were surgical in nature, and of these, 54% were preventable.[18] Similarly, a Canadian study found that between 37%–51% of reported AEs have been evaluated as preventable.[19] Of those preventable AEs, 51% have been a result of the surgical care provided.[19] In addition, post-operative AEs are costly,[20] cause pain and suffering,[21] extend hospital length of stay,[22] and are the most important determinant of decreased post-operative survival.[23]

Importance of rigorous and standardized recording of post-operative adverse events

The importance of rigorous recording of clearly defined post-operative AEs is widely recognized, yet is rarely and poorly performed in practice. Martin and colleagues[8] conducted a systematic review on the quality of AE reporting in the surgical literature, and found that the lack of uniformity in documenting and reporting AEs results in incomplete capture of data, making comparisons among different surgical approaches and institutional series difficult. The authors

strongly argue for the creation, and generalized use of standards for reporting this information, in order to assess the impact of therapeutic changes on patient outcomes, and to follow improvements in quality and safety over time.[8]

Methods for reporting and assessing post-operative adverse events and sources of bias

In the following section, the most common methods for reporting post-operative AEs are reviewed and potential sources of bias presented. It is important to note that the choice of reporting method can significantly affect the detection and reporting rates of post-operative AEs.

i. Morbidity & mortality conference

The most common strategy for quality reporting and assessment in surgery is the morbidity & mortality (M&M) conference. This approach uses peer-review of cases resulting in AEs to identify inappropriate care. M&M conferences have a number of drawbacks, focusing on: outliers and assigning blame, individual performance rather than system processes, individual events rather than patterns of outcomes, immediate complications rather than long-term results.[24] In addition, the total number of cases discussed at the M&M conferences is too low to accurately measure incidence or prevalence rates of AEs.

M&M conferences may also be limited by both hindsight and outcome bias. With respect to post-operative AEs reporting, hindsight bias refers to the tendency for individuals with outcome knowledge to exaggerate the extent to which they would have predicted the event beforehand.[11] Outcome bias refers to the influence of outcome knowledge upon evaluations of decision quality. Highly unfavourable outcomes might predispose a reviewer toward harsher judgments, while

minor injuries might elicit a less critical response.[11] This bias may be avoided by using evaluators blinded to outcome.

ii. Adverse event reporting systems

AEs may be reported using structured and objective data collection systems. This method of AE reporting is also referred to as voluntary reporting or incident reporting.[14] AE reporting systems are a way to involve surgeons in research and quality improvement. However, AE reporting systems cannot reliability measure incidence and prevalence rates because numerous factors may affect whether AEs are reported.[25] Surgeons may not report AEs because they are busy, afraid of lawsuits, or worried about their reputation (reporting bias). On the contrary, high reporting rates may indicate an organizational culture committed to identifying and reducing AEs rather than a truly high rate (biased over-reporting).[26]

iii. Administrative data review

Administrative data review is used commonly for public reporting and for inter-institution comparisons.[27] Administrative data review uses hospital billing data to screen for events containing International Classification of Diseases codes indicating hospital complications. This approach to quality assessment does not involve peer-review. There are several limitation to this method, including very high proportions of false-positive and false-negative cases resulting from coding mistakes in the billing data, and the method does not track all AE types, as codes are not currently adapted to capture diagnostic, system or management errors. Together, these errors can amount to up to 30% of AEs.[14,27]

iv. Medical record review or chart review

Reviews of patients' medical records have been the foundation of research into errors and AEs.[28] However, inter-rater agreements about the presence of AEs have low to moderate reliability or precision. Medical record review is also limited by poor documentation, presenting a source of hindsight bias.[28]

v. Electronic health record surveillance

Surveillance of the electronic health record may improve detection of errors and AEs by monitoring in near real-time and by integrating multiple data sources.[25] This method of detecting AEs has high sensitivity, but variable specificity, meaning it identifies most AEs but at the expense of a high false-positive rate. The high false-positive rate is a result of underdeveloped health information management systems in most centers and can be improved by including a peer-review process.[14,29]

vi. Prospective clinical surveillance of patient care

Prospective clinical surveillance is the most precise and accurate method of reporting AEs. Prospective clinical surveillance involves the direct observation of clinical care, including attending clinical rounds, daily review of charts, and participating in interviews with caregivers.[12] Prospective clinical surveillance is ideally suited for assessing the effectiveness of specific interventions to decrease explicitly defined AEs. However, prospective clinical surveillance is limited by practical and methodological issues, including the requirement for an observer who clearly understands clinical processes to ensure reliability. Furthermore, if observers are not blinded to patient outcome, outcome bias may once again be present.[25] Of the

above mentioned, prospective clinical surveillance, however, is the only method for which there are strong correlations to support the systematic measurement and monitoring of AEs with improvement in patient outcomes.[30]

vii. Large scale quality initiatives

There are several ongoing, large-scale initiatives aimed specifically at measuring and improving surgical outcomes. For example, the American College of Surgeons-National Surgical Quality Improvement Program (ACS-NSQIP)[31] and the Society for Thoracic Surgeons (STS) database,[32] provide hospitals and surgeons with information on their risk-adjusted M&M rates. The ACS-NSQIP was created as a program that originated in the Veteran's Administration Hospitals in the United States as a quality improvement tool for surgical care.[31] The ACS-NSQIP uses clinical information from medical records to risk-adjust hospital M&M rates. One of the advantages of this type of system includes extensive clinical information on over one million patients for risk-adjusted analyses on 30-day outcomes of surgical care. Benefit from participating in the ACS-NSQIP continues to be documented. Hall and colleagues have demonstrated that on average, the outcomes of all hospitals participating in ACS-NSQIP improved over time.[33] Specifically, since ACS-NSQIP has been implemented, marked improvement in surgical quality has been documented – M&M rates have declined, patient satisfaction has improved, and lengths of stay have decreased.[33]

Similarly, the STS database was initiated in 1989 and has become the premier clinical data registry for cardiothoracic surgery.[32] It was designed to capture detailed clinical data and outcomes related to the immediate postoperative period of 30 days. In 1999, the STS database

was expanded to include data on general thoracic surgical procedures. Risk-adjusted short-term results are provided to participating institutions on a biannual basis. The STS database provides participating members with risk-adjusted national thoracic surgical benchmarks for lung and esophageal cancer resections. Since its inception, the STS database has grown as a powerful source of risk-adjusted outcomes, large scale scientific contributions, and invaluable information for healthcare policy making.[32]

Together, the ACS-NSQIP and STS databases independently offer inter-institutional benchmarking, however, they are less applicable as a continuous quality improvement measure for an individual surgical program, as understanding and improving the delivery of a particular operation may require measures tailored to that operation,[7] such as proper evaluation of the burden of illness of individual complications and subsequent patient impact. Second, the tracking of post-operative morbidity is less successful than of mortality due to the lack uniform definitions and severity grading between and within the ACS-NSQIP and STS databases.[34] Third, without a precise review of 100% a surgeon's case portfolio, the use of ACS-NSQIP data could theoretically result in inaccurate reporting and ranking of individual surgeon performance.[35]

Clavien-Dindo classification system of post-operative adverse events

In 1992, Clavien and colleagues were the first to introduce the concept of severity grading of post-operative AEs (**Table 0.1**).[17] The principle underlying this classification is that the severity of an AE is proportional to its impact on a patient and the degree of effort to rectify it.[17] The rationale for this approach was to eliminate subjective interpretation of serious AEs and any

tendency to down-grade complications, as it is based on data that are well documented and easily verified. The system, now known as the Clavien-Dindo classification of AEs, was validated in 2004 in a large cohort of general surgical patients,[36] and has since been used in a number of different surgical sub-specialities.[37-40]

Table 0.1 Classification of post-operative adverse events.

Grade		Definition
Complication		Any deviation from the normal post-operative course.
Minor	Grade I	Any complication without need for pharmacologic treatment or other intervention.
	Grade II	Any complication that requires pharmacological treatment or minor intervention only.
Major	Grade III	Any complication that requires surgical, radiological, endoscopic intervention, or multi-therapy.
	Grade IIIa	Intervention does not require general anaesthesia.
	Grade IIIb	Intervention requires general anaesthesia.
	Grade IV	Any complication requiring ICU management and life support.
	Grade IVa	Single organ dysfunction.
	Grade IVb	Multi-organ dysfunction.
Mortality	Grade V	Any complication leading to the death of the patient.

Systematic classification of Thoracic Morbidity & Mortality

In 2008, modeled on the Clavien-Dindo system, The Ottawa Hospital's Division of Thoracic Surgery developed a standardized approach to identify both presence and severity of Thoracic Morbidity & Mortality (TM&M), with the aim to capture all AEs after all surgeries.[41] Definitions of surgical AEs were modified according to complications in patients following thoracic surgery through peer review and questionnaire, and adjusted based on surgeons' experience. A complication was defined as any deviation from the normal post-operative course. For each of the following systems – pulmonary, pleural, cardiac, renal, anastomotic, gastrointestinal, neurological, wound, and other – complications were defined and classified according to their

specific gradation. The Common Terminology Criteria for Adverse Events (version 3.0) was also used to further refine specific definitions.[42]

The TM&M classification system is an in-hospital prospectively collected monitoring system. Our staff and residents are trained to proactively monitor patients for post-operative AEs. Post-operative AEs are chosen from a series of standardized definitions and complications are recorded in real time on a daily basis by thoracic surgical fellows. Weekly review by staff surgeons allows for affirmation of complications. Ongoing feedback in the process of quality reporting plays an essential role in maintaining the accuracy and completeness of data. The TM&M classification system allows users to analyze outcomes in the full census of thoracic surgical patients and data can be subdivided by priority of surgery, disease diagnosis, procedure class, and surgical approach/incision. The system can be used to evaluate severity and burden of post-operative AEs, and represents a continuous and divisional approach to surgical quality assessment.

Testing the reliability and reproducibility of Thoracic Morbidity & Mortality classification system

The rigorous implementation in clinical practice and evaluation of such a system requires time. Since its inception, our division has demonstrated that the TM&M classification is feasible, facilitates objective comparison, identifies burden of individual complications, and provides an effective method for continuous surgical quality assessment.[41]

We evaluated the TM&M classification system for its reliability and reproducibility through a questionnaire of the Canadian Association of Thoracic Surgeons (CATS) (n = 95 members at the

time of study), in which respondents were asked to classify 20 case-based examples of post-operative AEs in accordance to the proposed classification system.[43] The membership of CATS includes full-time practitioners of general (non-cardiac) thoracic surgery, along with qualified general and cardiovascular surgeons whose practice includes more than 50% thoracic surgery.[44] Inter-rater agreement among the survey respondents was evaluated with the use of Kappa statistics.

Results of membership survey revealed that that the TM&M system offers high inter-rater reliability or agreement on the classification of complications: 87% of kappa statistics were >0.81, a range that is interpreted as "almost perfect agreement;" and the remaining 13% ranged between 0.61 and 0.8, interpreted as "substantial agreement."[43] The classification system was further regarded as straightforward by 98% of the respondents, reproducible by 94%, logical 92%, and useful 98% of respondents.

More recently, the TM&M classification system has been adopted by numerous surgical groups internationally.[45-46] Salatia and colleagues have demonstrated the usefulness of the TM&M classification system in auditing the quality of care within a single surgical unit. The authors concluded that the TM&M classification system revealed a decline in quality of care within their unit otherwise undetected by applying traditional outcomes measures, and that the system can be used as an additional graded outcomes end point to refine internal audit of performance.[45-46] We are also aware of randomized trials, which have included this classification in their end point evaluation.[47]

The above noted studies[41, 43] were conducted at the University of Ottawa, in conjunction with the Clinical Epidemiology Program at the Ottawa Hospital Research Institute, in partial fulfillment of the requirements for the Master of Science degree in Epidemiology, under the supervision of Dr. Andrew Seely and Dr. Tim Ramsay.

THESIS OBJECTIVES

The purpose of the following doctoral thesis is to further test the utility of the TM&M classification system in various clinical scenarios and to critically evaluate its use as a means for quality improvement in the thoracic surgical care setting at The Ottawa Hospital. Specifically, the objectives are:

1. To use the prospective AE monitoring system to **determine** the incidence, severity, and risk factors of the most common complication overall (i.e. prolonged alveolar air leak) following pulmonary resection (**Chapter I**)
2. To use the prospective AE monitoring system to **determine** the incidence, severity, and risk factors of the most common cardiovascular complication, and the second most common AE overall (i.e. atrial fibrillation) following pulmonary resection (**Chapter II**)
3. To **compare** post-operative AE reporting on the same patients utilizing two classification systems: the retrospectively-recorded ACS-NSQIP and the prospectively-collected TM&M classification system (**Chapter III**)
4. To **develop**, **implement** and **evaluate** the feasibility of surgeon-specific outcome reports and of a surgeon-led, continuous quality improvement program as a means to improve outcomes following thoracic surgical care (**Chapter IV**)

5. To **analyze** the impact of a prospective, standardized, and comprehensive post-operative AE monitoring and reporting system as a means for quality improvement (**Chapter V**)

REFERENCES

1. Lohr KN, Donaldson MS, Harris-Wehling J. Medicare: A strategy for quality assurance: Quality of care in a changing health care environment. QRB Qual Rev Bull 1992;18:120−26.
2. Campbell SM, Roland MO, Buetow SA. Defining quality of care. Soc Sci Med 2000;51:1611−25.
3. Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260:1743−48.
4. Brook RH, McGlynn EA, Shekelle PG. Defining and measuring quality of care: A perspective from US researchers. Int J Qaul Health Care 2000;12:281−95.
5. Fink AS, Itani KM, Campbell DC Jr. Assessing the quality of surgical care. Surg Clin N Am 2007;87:837−52.
6. Gruen RL. Monitoring and improving quality in surgery. ANZ J Surg 2008;78:940−42.
7. Birkmeyer JD, Dimick JB, Birkmeyer NJ. Measuring the quality of surgical care: structure, process, or outcomes? J Am Coll Surg 2004;198:626−32.
8. Martin RC 2nd, Brennan MF, Jaques DP. Quality of complication reporting in the surgical literature. Ann Surg 2002;235:803−13.
9. Neuhauser D. Ernest Amory Codman MD. Qual Health Care 2002;11:104−05.
10. Reason JT. Human Error, Cambridge University Press, Cambridge, UK, 1990.
11. Calland JF, Guerlain S, Adams RB, et al. A systems approach to surgical safety. Surg Endosc 2002;16:1005−14.
12. Etchells E, O'Neill C, Bernstein M. Patient safety in surgery: Error detection and prevention. World J Surg 2003;27:936−42.
13. Kohn LT, Corrigan JM, Donaldson MS. To err is human: Building a safer health system. Washington (DC): Institute of Medicine. National Academy Press; 1999.
14. Forster AJ, Dervin G, Martin C, et al. Improving patient safety through the systematic evaluation of patient outcomes. Can J Surg 2012;55:418−25.
15. Khuri SF, Daley J, Henderson W. The measurement of quality in surgery. Advan Surg 1999;33:113−40.
16. Khuri SF, Daley J, Henderson W, et al. Risk adjustment of the post-operative mortality rate for the comparative assessment of the quality of surgical care: Results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:315−27.
17. Clavien PA, Sanabria JR, Strasberg SM. Proposed classification of complications of surgery with examples of utility incholecystectomy. Surg 1992;111:518−26.
18. Gawande AA, Thomas EJ, Zinner MJ, et al. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. Surg 1999;126:66−5.
19. Baker GR, Norton PG, Flintoft V et al. The Canadian Adverse Events Study: The incidence of adverse events among hospital patients in Canada. CMAJ 2004;170:1678−86.
20. Vonlanthen R, Slankamenac K, Breitenstein S, et al. The impact of complications on costs of major surgical procedures: A cost analysis of 1200 patients. Ann Surg 2011;254:907–13.
21. Chorney JM, McGrath P, Finley GA. Pain as the neglected adverse event. CMAJ 2010;182: 732.
22. Khan NA, Quan H, Bugar JM, et al. Association of post-operative complications with hospital costs and length of stay in a tertiary care center. J Gen Intern Med 2006;21:177–80.

23. Khuri SF, Henderson WG, DePalma RG, et al. Determinants of long-term survival after major surgery and the adverse effect of post-operative complications. Ann Surg 2005;242: 326–43.
24. Feldman L, Barkun J, Barkun A, et al. Measuring post-operative complications in general surgery patients using an outcomes-based strategy: Comparison with complications presented at morbidity and mortality rounds. Surgery 1997;122:711−19.
25. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. J Gen Intern Med 2003;18:61−7.
26. Edmonson AC. Learning from mistakes is easier said than done: Group and organizational influences on the detection and correction of human error. J Appl Behav Sci 1996;32:5−28.
27. Weingart SN, Iezzoni LI, Davis RB, et al. Use of administrative data to find substandard care: Validation of the complications screening program. Med Care 2000;38:796−806.
28. Luck J, Peabody JW, Dresselhaus TR, et al. How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. Am J Med 2000;108:642−49.
29. Murff HJ, Patel VL, Hripcsak G, et al. Detecting adverse events for patient safety research: A review of current methodologies. J Biomed Inform 2003;36:131−43.
30. Neumayer L, Mastin M, Vanderhoof L, et al. Using the Veterans Administration National Surgical Quality Improvement Program to improve patient outcomes. J Surg Res 2000;88:58−61.
31. Khuri SF, Daley J, Henderson W, et al. The Department of Veterans Affairs' NSQIP: The first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. Ann Surg 1998;228:491−507.
32. Caceres M, Braud RL, Garrett HE Jr. A short history of the society of thoracic surgeons national cardiac database: Perceptions of a practicing surgeon. Ann Thor Surg 2010;89:332−39.
33. Hall BL, Hamilton BH, Richards K, et al. Does surgical quality improve in the American College of Surgeons National Surgical Quality Improvement Program: An evaluation of all participating hospitals. Ann Surg 2009;250:363−76.
34. Grover FL, Shroyer AL, Hammermeister KE. Calculating risk and outcome: The veterans affairs database. Ann Thor Surg 1996;62(5 Suppl):S6−11.
35. Kuhnen AH, Marcello PW, Roberts PL, et al. Can the National Surgical Quality Improvement Program Provide Surgeon-Specific Outcomes? Dis Colon Rectum 2015;58:247–253.
36. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: A new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205−13.
37. DeOliveira ML, Winter JM, Schafer M, et al. Assessment of complications after pancreatic surgery: A novel grading system applied to 633 patients undergoing pancreaticoduodenectomy. Ann Surg 2006;244:931–39.
38. Gonzalgo ML, Pavlovich CP, Trock BJ, et al. Classification and trends of perioperative morbidities following laparoscopic radical prostatectomy. J Urol 2005;174:135–39.
39. Kocak B, Koffron AJ, Baker TB, et al. Proposed classification of complications after live donor nephrectomy. Urology 2006;67:927–31.

40. Tamura S, Sugawara Y, Kaneko J, et al. Systematic grading of surgical complications in live liver donors according to Clavien's system. Transpl Int 2006;19:982–87.
41. Seely AJE, Ivanovic J, Threader J, et al. Systematic classification of morbidity and mortality after thoracic surgery. Ann Thor Surg 2010:90:36−42.
42. Trotti A, Colevas AD, Setser A, et al. CTCAE v3.0: Development of a comprehensive grading system for the adverse effects of cancer treatment. Semin Radiat Oncol 2003;13:176−81.
43. Ivanovic J, Al-Hussaini A, Al-Shehab D, et al. Evaluating the reliability and reproducibility of the Ottawa Thoracic Morbidity and Mortality classification system. Ann Thorac Surg 2011;91:387−93.
44. Darling GE, Maziak DE, Clifton JC, et al. The practice of thoracic surgery in Canada. Can J Surg 2004;47:438−45.
45. Salati M, Pompili C, Refai M, et al. The use of the Thoracic Morbidity and Mortality system for the internal analysis of performance: A case-matched temporal audit. Eur J Cardiothorac Surg 2014;45:859–63.
46. Salati M, Refai M, Pompili C, et al. Major morbidity after lung resection: A comparison between the European Society of Thoracic Surgeons Database system and the Thoracic Morbidity and Mortality system. J Thorac Dis 2013;5:217–22.
47. Infante M, Chiesa G, Solomon D, et al. Surgical procedures in the DANTE trial: A randomized study of lung cancer early detection with spiral computed tomography: Comparative analysis in the screening and control arm. J Thorac Onc 2011;6:327–35.

Commentary

Upon our initial introduction of the systematic classification of Thoracic Morbidity & Mortality (TM&M), post-operative prolonged alveolar air leak (PAAL) was identified as the most common, and arguably, the most specialty-relevant, and impactful complication following lung resection.[1] The impact that PAAL has on patient recovery and hospital resources is substantial. The high rate of post-operative PAAL has thus prompted significant interest in measures designed to prevent this common complication. However, these measures are costly and may not be beneficial if used routinely. Identification of patients at highest risk might allow for more effective use of these preventative measures. In the following study, we sought to determine the incidence and severity of PAAL using the TM&M classification system, and to develop a simple score to predict PAAL. We defined PAAL as a forced expiratory air leak present on post-operative day 5. The 5-day definition is also consistent with that used in the European Society of Thoracic Surgeons and the American Society of Thoracic Surgeons databases.

Since the publication of the following study, several members from our division have issued a follow-up review chapter on the current available evidence of management modalities for persistent post-operative PAAL with the goal of optimal individualized patient care plans. These management modalities include: outpatient chest tube drainage, intra-pleural chemical sclerosis, intra-pleural blood patch, topical sealants, pneumoperitoneum, endo-bronchial valves and surgical repair.[2]

REFERENCES

1. Seely AJ, Ivanovic J, Threader J, et al. Systematic classification of morbidity and mortality after thoracic surgery. Ann Thorac Surg 2010;90:936−42.
2. McGuire AL, Sundaresan RS. Management of Persistent Post-operative Alveolar Air Leak. Difficult Decisions in Surgery: An Evidence-Based Approach 2014:1;207−28.

Chapter I

Quantifying the incidence and impact of post-operative prolonged alveolar air leak after pulmonary resection

J Thorac Cardiovasc Surg 2013;145:948–954

Shuyin Liang, MD;[1]
Jelena Ivanovic, MSc;[2,3,4]
Sebastien Gilbert, MD, FRCSC;[3]
Donna E. Maziak, MDCM, MSc, FRCSC, FACS;[2,3,4]
Farid M. Shamji, MBBS, FRCSC;[3]
R. Sudhir Sundaresan, MD, FRCSC, FACS;[3,4]
and Andrew J. E. Seely, MD, PhD, FRCSC.[2,3,4]

Faculty of Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Canada[1]
Department of Epidemiology and Community Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Canada[2]
Division of Thoracic Surgery, Department of Surgery, The Ottawa Hospital, 501 Smyth Road, Ottawa, Canada[3]
Clinical Epidemiology Program, Ottawa Hospital Research Institute, 501 Smyth Road, Ottawa, Canada[4]

ABSTRACT

OBJECTIVE: Prolonged alveolar air leak (PAAL) is a frequent occurrence after lobectomy or lesser resections. The resulting complications and their impact are not well understood. Our aims are to prospectively determine the incidence and severity of PAAL after pulmonary resection using the Thoracic Morbidity & Mortality (TM&M) classification system and to identify risk factors.

METHODS: A prospective collection of TM&M data was performed for all consecutive pulmonary resections (n=380; January 2008 to April 2010). Demographics, comorbidities, and preoperative cardiopulmonary assessment were retrospectively identified. The incidence and severity (grades I-V) of burden from PAAL were quantified using the TM&M classification system. Risk factors for PAAL and severe PAAL (defined as leading to major intervention, organ failure, or death) were sought with univariate and multivariate analyses.

RESULTS: The incidences of PAAL and severe PAAL were 18% and 4.8%, respectively. PAAL prolonged the median hospital stay by 4 days. The majority of complications associated with PAAL were limited to pulmonary and pleural categories (90%). Significant predictors of PAAL from multivariate analysis include severe radiologic emphysema (odds ratio [OR], 2.8; confidence interval [CI], 1.2−6.2), histopathologic emphysema (OR, 1.9; CI, 1.1−3.6), percentage of predicted value for FEV_1 less than 80% (OR, 1.9; CI, 1.1−3.3), and lobectomy (OR, 4.9; CI, 1.7−14.1). Risk factors for severe PAAL include radiologic emphysema, percentage of predicted value for forced expiratory volume in 1 second less than 80%, FEV_1/FVC ratio less than 70%, and intraoperative difficulties (p<0.05).

CONCLUSIONS: PAAL leads to longer hospital stays, and approximately 4.8% of patients undergoing pulmonary resection experience PAAL that necessitates placement of additional chest drains, bronchoscopy, reoperation, or life support. Further study is required to assess the cost-effectiveness of measures to reduce PAAL.

INTRODUCTION

Prolonged alveolar air leak (PAAL) is the most common complication and reason for increased hospital length of stay (LOS) after elective lobectomy or lesser lung resections.[1,2] PAAL is defined as air leakage that lasts more than 3 to 7 days,[3-6] and its reported incidence ranges from 8% to 26%.[3-5,7-9] Given the important clinical impact of PAAL, attempts to delineate specific risk factors for PAAL have been reported in previous series with variable consistency. The most consistent risk factor is chronic obstructive pulmonary disease, reflected by preoperative pulmonary function test (PFT): forced expiratory volume in 1 second (FEV_1)/forced vital capacity (FVC) ratio less than 70%, FEV_1 less than 1.5 liters, FEV_1 less than 79% predicted, and diffusing capacity of carbon monoxide (DLCO) less than 80% predicted.[3-9] Other potential risk factors, including radiologic and pathologic findings of chronic obstructive pulmonary disease, have not been studied.

The impact that PAAL has on patient recovery and hospital resources is significant. It increases LOS by 5 to 13 days[2,10] and leads to additional complications in both the lung and the pleural space, such as atelectasis, pneumonia, empyema, and prolonged need for chest drains.[3,11] There has been some difficulty in quantifying what constitutes severe PAAL. As a result, the incidence, predictors, and burden of illness from severe PAAL remain elusive.

The current study addresses the issues of stratification of burden of illness by using an adverse event (AE) reporting and monitoring system, the Thoracic Morbidity & Mortality (TM&M) system. The TM&M system was derived from the Clavien-Dindo classification,[12-14] which classifies the severity of a complication on the basis of the impact it has on the patient, namely, a

complication that occurs leading to no change in management (I), new medical therapy (II), major intervention (III), organ failure (IV), or death (V). We developed definitions of thoracic complications listed by system and stratified by severity.[12,13] Furthermore, the origin and classification of each complication were reviewed and discussed weekly over several years, helping to refine the TM&M system. The objectives of the current study were 2-fold: to systematically quantify the burden of illness from PAAL using the TM&M system and to stratify risk factors for non-severe PAAL (grades I and II) and severe PAAL (grades III, IV, and V). Secondary outcomes of interest include the presence of additional AEs, LOS, and rates of readmissions in patients with PAAL.

MATERIALS & METHODS

Patients

The data for 380 consecutive pulmonary resections for malignant and benign disease within The Ottawa Hospital from January 2008 to May 2010 were prospectively collected by The Ottawa Hospital's Division of Thoracic Surgery, approved by the Ottawa Hospital Research Ethics Board. Prospective TM&M data were initially recorded by the chief resident, reviewed weekly by thoracic staff surgeons, and presented monthly at morbidity and mortality rounds. Patients with pancoast tumors (n=0), patients with tumors requiring pneumonectomy (n=24), or patients who did not have any preoperative evaluations before surgery on record (n=4) were excluded. The records of 352 pulmonary resections remained for analysis. Four patients had 2 separate pulmonary resections during the study period. The data for each surgery were considered as an independent entry in the analysis.

Collection of preoperative data

The preoperative evaluations included complete history, physical examination, PFT, arterial blood gas, computed tomography (CT) scan of the chest, echocardiography or cardiac stress tests, and biopsy. The severity of emphysema was graded by a chest radiologist and recorded in the radiologic reports. Operative, radiologic, and pathology reports and the TM&M database were reviewed to document the procedure, intraoperative complications, and pathologic stage. Data were collected on paper case report forms and entered into a Microsoft Excel computer database (Microsoft Corp, Redmond, Wash).

Intra-operative and post-operative collection of prolonged alveolar air leak and Thoracic Morbidity & Mortality data

The techniques of pulmonary resection, chest tube placement, and management were not controlled; however, general principles guided surgical intraoperative and post-operative practice. Mechanical staplers were mostly used to complete incomplete fissures; however, in open cases, cautery often was used to develop fissures overlying the pulmonary artery. All bronchial stumps were verified to be airtight before closure. If vigorous air leaks were identified intra-operatively, the parenchymal source of bubbling was repaired with sutures. In general, patients who underwent lobectomy received two 28F chest tubes or one 28F chest tube and one 14F pigtail pleural catheter. Those patients who underwent segmentectomy or wedge resection received one 28F chest tube. Immediately after the surgery, the chest tubes were attached to the Sahara S-11000 (Teleflex, Research Triangle Plus, Durham, NC) analogue chest drainage system and placed on −10 to −20 cm H_2O suction. The tubes were converted to water seal on the morning of post-operative day (POD) 1 after chest radiography. The forced expiratory air leak was determined by visualizing bubbles in the analogue drainage system while the patient

coughed in an upright sitting position. Patients remained on water seal unless they had an enlarging symptomatic pneumothorax or subcutaneous emphysema developed. When no air leak was detected, the chest tube was removed. If the air leak was equivocal, the tube was clamped and chest radiography was performed, followed by removal of the chest tube if no new pneumothorax or subcutaneous emphysema was identified. If the patient continued to have an air leak on the day of discharge, the patient was discharged with the chest tube attached to a Pneumostat Chest Drain Valve (Atrium Medical Corp, Hudson, NH) and re-evaluated 5 to 7 days later.

Classification of post-operative complications

In the current study, PAAL is defined as a forced expiratory air leak present on POD 5. The 5-day definition is also consistent with that used in the European Society of Thoracic Surgeons and the American Society of Thoracic Surgeons research databases. The presence and severity of post-operative PAAL were classified using the validated TM&M schema,[12,13] developed in accordance with the Clavien-Dindo classification system.[14] The types of complications are pulmonary, pleural, anastomotic, cardiac, renal, gastric, neurologic, wound, and other. In the context of PAAL, the complication grade starts at II, which requires a chest tube for more than 5 days after surgery; and proceeds to grade III, which requires the insertion of an additional chest tube (grade IIIA) or reoperation (grade IIIB); grade IV, which requires intensive care and life support; and grade V, which results in mortality within 30 days. The use of the Pneumostat Chest Drain Valve (Atrium Medical Corp) by itself was considered a grade II complication, as long as the patient did not require interventions listed in the higher grades.

Statistical analysis

Data were collected as categoric variables and converted to binary numeric data where applicable. Univariate analysis using chi-square tests was carried out between the control (no PAAL) and PAAL groups, and between the control and severe PAAL groups. Variables with a *p* value <0.05 in the univariate analysis were used as independent variables in forward logistic regression analysis. In 46 cases, there were missing values in PFT values or radiologic grading of emphysema. Therefore, imputational statistics were used to replace the missing values. Discrimination and calibration of the model were assessed using the C-statistic and the Hosmer–Lemeshow goodness-of-fit test. For variables containing multiple categories, such as procedure performed, a reference category was chosen. The model satisfied the convergence criteria. Data were analyzed using SAS version 9.2 (SAS Institute Inc, Cary, NC).

RESULTS

Study population demographics

A total of 352 pulmonary resections met the inclusion criteria: 13 resections for benign disease and 339 resections for malignant disease (270 non–small cell lung cancers and 63 other malignancies). There was no significant difference between study groups regarding age greater than 70 years (p=0.12), male sex ($p=0.25$), and body mass index greater than 25 kg/m ^{2} (p=0.31).

Univariate and multivariate analyses to identify risk factors associated with prolonged alveolar air leak

The incidence of PAAL was 18% (n=65). **Table 1.1** lists the significant preoperative and intraoperative risk factors analyzed in this study (p<0.05). The PAAL group had higher pack-years of smoking (41.5 vs. 38.5 pack-years) and self-reported diagnosis of bronchitis (10.8% vs.

4.2%). Patients with PAAL were more likely to have undergone pulmonary resections for non–small cell lung cancers (92.3% vs. 73.1%), to display severe emphysema on CT scan (20.6% vs. 8.0%) and on histopathology (67.7% vs. 44.3%), and to have an obstructive pattern on PFT (predicted FEV_1<80%: 58.6% PAAL vs. 38.2% control; FEV_1/FVC<70%: 64.4% PAAL vs. 45.8% control). An additional *t*-test for PFT items showed that patients with PAAL have a significantly lower percentage of predicted value for FEV_1 (%FEV_1) and FEV_1/FVC ratio (p<0.05), but there was no difference in percentage of predicted values for DLCO for age, gender, and height. Patients with PAAL were more likely to have undergone lobectomy (92.3% of PAAL cases vs. 67.2% of control cases) and have pleural adhesions requiring lysis or decortication (40.4% vs. 26.5%). There was no difference in the rate of PAAL between minimally invasive or open approaches (p=0.83). Other nonsignificant variables from the univariate analyses include history of coronary artery disease (p=0.34), asthma (p=0.15), pulmonary hypertension (p=0.26), lung cancer stage (p=0.20), chemotherapy (p =0.28), radiotherapy (p=0.63), percent of residual volume (p=0.32), %DLCO (p=0.88), intraoperative complications (p=0.58), and atelectasis on POD 1 (p=0.18).

Table 1.1 Univariate analyses to identify peri-operative risk factors associated with PAAL.

Patient characteristics	Control, n (%)	PAAL, n (%)	*p* Value
Bronchitis	12/287 (4.2%)	7/65 (10.8%)	0.0338*
Smoking Status	n=285	n=65	- -
Never	43 (15.1%)	4 (6.2%)	0.0566
Current	94 (33.0%)	23 (35.4%)	0.7110
Pack Year	20.2	28.8	0.0209*
Past	139 (48.8%)	38 (58.4%)	0.1585
Pack Year	39.0	41.5	0.0033**
Diagnosis	n = 287	n = 65	- -
NSCLC	210 (73.1)	60 (92.3)	0.0010**
Other Malignant	66 (23.0)	3 (4.6)	0.0007**
Benign	11 (3.8)	2 (3.1)	0.7705
Emphysema on CT	n = 275	n = 63	- -
None	203 (73.8%)	32 (50.8%)	0.0003**
Not severe	50 (18.2%)	18 (28.6%)	0.0635
Severe	22 (8.0%)	13 (20.6%)	0.0030**
FEV1 Actual	2.21	2.13	0.1169
% FEV1	n = 275	n = 58	- -
# Patients <80	105 (38.2%)	34 (58.6%)	0.0041**
Numerical value	Mean 85.3, SE 1.2	Mean 78.9, SE 2.6	0.029*
FEV_1/FVC	n = 271	n = 59	- -
# Patients <70	124 (45.8%)	38 (64.4%)	0.0094**
Numerical value	Mean 68.8, SE 0.58	Mean 64.2, SE 1.5	0.005**
Lobectomy	193/287 (67.2%)	60/65 (92.3%)	<.0001**
Right upper	58/193 (30.1%)	33/60 (55%)	0.0004**
Right middle	17/193 (8.8%)	1/60 (1.7%)	0.0602
Right lower	31/193 (16.1%)	9/60 (15.0%)	0.8439
Left upper	50/193 (25.9%)	15/60 (25.0%)	0.8883
Left lower	37/193 (19.2%)	2/60 (3.3%)	0.0030**
Wedge Resection	171/216 (88.6%)	30/54 (46.2%)	0.0004**
Lobectomy+Wedge	92/287 (32.1%)	25/64 (38.5%)	0.2823
Lobectomy Alone	38/287 (13.2%)	24/65 (36.9%)	<.0001**
Wedge Alone	79/287 (27.5%)	4/65 (6.2%)	0.0002**
Bilobectomy	2/291 (0.7%)	0/65 (0%)	0.5027
Extended Lobectomy	3/287 (1.0%)	3/65 (4.6%)	0.0447*
Segmental Resection	16/287 (5.6%)	1/65 (1.5%)	0.1705
Pleural Adhesions	76/286 (26.5%)	26/65 (40.0%)	0.0314*
Emphysema on Path	127/287 (44.3%)	44/64 (67.7%)	0.0006**

Statistically significant risk factors for prolonged alveolar air leak (PAAL) from univariate analysis. The results are expressed as count/total population of the collected data, followed by percentage of total of the group in brackets, except in rows with continuous variable statistics. "*" denotes *p* value of <0.05, and "**" denotes *p* value of <0.01. Lung cancers are staged according to American Joint Committee on Cancer 7th edition. NSCLC=non-small cell lung cancer; CT= computed tomography; FEV_1= forced expiratory volume in 1 second; $\%FEV_1$= percentage of predicted value for FEV_1; FVC= forced vital capacity; %DLCO= percentage of predicted values for DLCO for age, gender and height.

The results of the multivariate analysis are shown in **Table 1.2**. For the purpose of multivariate analysis, the category "procedure performed" was reclassified into 3 subcategories: wedge, segmental resection, and lobectomy. Lobectomy includes all the different lobes resected and extended lobectomies. Significant predictors of PAAL are radiologic findings of severe emphysema (odds ratio [OR], 2.8; confidence interval [CI], 1.2−6.2), histopathologic finding of emphysema (OR, 1.9; CI, 1.1−3.6), %FEV_1 less than 80% (OR, 1.9; CI, 1.1−3.3), and lobectomy (OR, 4.8; CI, 1.7−14.1). Of note, smoking history, self-reported history of bronchitis, type of tumor, pleural adhesions, and FEV_1/FVC less than 70% were not significant predictors of PAAL. The final model had a C-statistic of 0.74, Hosmer–Lemeshow chi-square value of 4.1, and *p* value of 0.77.

Table 1.2 Multivariate analyses to identify peri-operative risk factors associated with PAAL.

Variable	*p* value	Point Estimate of Odds Ratio	95% Confidence Interval	
Radiological emphysema	0.035*			
None (reference)				
Not severe	0.15	1.7	0.83	3.3
Severe	0.014*	2.8*	1.2	6.2
Pathological emphysema	0.032*	1.9*	1.1	3.6
Bronchitis	0.058*	2.9	0.97	8.5
%FEV_1	0.035*			
≥ 80 (reference)				
≤ 80	0.035*	1.9*	1.1	3.3
Procedure performed	0.0097**			
Wedge (reference)				
Segmental resection	0.72	1.52	0.15	15.4
Lobectomy	0.0038**	4.8*	1.7	14.1

The final model has a C statistic value of 0.74, Hosmer–Lemeshow chi-square of 4.1, and *p* value of 0.77. OR, Odds ratio; CI, confidence interval; %FEV_1, percentage of predicted value for FEV_1. Results of multivariate analysis on risk factors for PAAL. *p<0.05. **p<0.01.

Univariate analysis to identify risk factors associated with severe prolonged alveolar air leak

The incidence of severe PAAL was 4.8% among all patients and 26% in the PAAL group. Risk factors for severe PAAL are shown in **Table 1.3**. Significant risk factors ($p<0.05$) included abnormal right ventricular function seen on echocardiogram (12.5% severe PAAL vs. 2.2% control), radiologic finding of emphysema on CT scan (23.5% vs. 8.0%), FEV_1 less than 80% predicted (71.4% vs. 38.2%), FEV_1/FVC less than 70 (80% vs. 45.8%), and occurrence of intraoperative difficulties (17.6% vs. 4.5%), such as intraoperative hypoxemia, bleeding, and hypotension. An additional *t*-test for PFT items showed that patients with severe PAAL had significantly lower $\%FEV_1$ and FEV_1/FVC ratio ($p<0.05$), but there was no difference in %DLCO. Patients with severe PAAL were less likely to have undergone wedge resection (53.8% severe PAAL vs. 79.2% control). Of note, patients with severe PAAL did not display a significant difference in smoking history ($p=0.65$) or pathologic findings of emphysema ($p=0.24$) compared with the control group.

Table 1.3 Univariate analyses to identify peri-operative risk factors associated with severe PAAL.

	Control (%)	Severe PAAL (%)	*p* Value
Smoking Status	n = 278	n = 17	- -
Never	44 (15.8%)	2 (11.8%)	0.6540
Current	95 (34.2%)	6 (35.3%)	0.9246
Past	139 (50.0%)	9 (52.9%)	0.8139
Abnormal right echo	5/225 (2.2%)	2/16 (12.5%)	0.0180*
Emphysema on Path	127/287 (44.3%)	10/17 (58.8%)	0.2404
Emphysema on CT	n = 275	n = 17	- -
None	201 (73.1%)	6 (35.3%)	0.0009**
Not severe	50 (18.2%)	7 (41.2%)	0.0203*
Severe	22 (8.0%)	4 (23.5%)	0.0291*
% FEV_1	n = 275	n = 14	- -
# Patients <80	105 (38.2%)	10 (71.4%)	0.0132*
Numerical value	Mean 85.3, SE 1.2	Mean 70.9, SE 3.7	0.002**
FEV_1 FVC	n = 271	n = 15	- -
# Patients <70	124 (45.8%)	12 (80.0%)	0.0097**
Numerical value	Mean 68.8, SE 0.58	Mean 61.1, SE 2.9	0.021*
% DLCO	n = 255	n = 15	- -
# Patients <60	39 (15.3%)	3 (20.0%)	0.6251
Numerical value	Mean 79.3, SE 1.2	Mean 75.0, SE 5.3	0.441
Wedge Resection	171/216 (79.2%)	7/13 (53.8%)	0.0331*
Intraoperative Complications	13/287 (4.5%)	3/17 (17.6%)	0.0186*

The results from univariate analysis on risk factors for severe PAAL. Results are expressed as numeric value of the count, followed by percentage of total of the group in brackets except in rows with continuous variable statistics. The total number of collected cases for each category is expressed as n = number. PAAL, Prolonged alveolar air leak; CT, computed tomography; FEV_1, forced expiratory volume in 1 second; %FEV_1, percentage of predicted value for FEV_1; FVC, forced vital capacity; %DLCO, percentage of predicted values for DLCO for age, gender, and height; SE, standard error. *p<0.05. **p<0.01.

Post-operative management of prolonged alveolar air leak, length of stay, readmission, and additional complications

LOS was significantly longer in the PAAL group (86.2% stayed for >5 days) compared with the control group (29.2%) (p=0.001). Likewise, the rate of readmission within 30 days was 24.6% in patients with PAAL compared with 4.2% in patients without PAAL. Patients with PAAL maintained an indwelling chest drain for an average of 18 ± 16.4 days, and 46.2% of those with

PAAL were discharged with a chest drain. The duration of PAAL is shown in **Table 1.4**. The time course of severe PAAL is significantly longer than that of non-severe PAAL.

Table 1.4 Differences in the duration of non-severe PAAL and severe PAAL.

	"Non-severe PAAL" (%)	"Severe PAAL" (%)	*p* Value
Duration of PAAL	n = 48	n = 17	- -
≤5 days	3 (6.2%)	1 (5.9%)	0.5941
6-10 days	18 (37.5%)	2 (11.7%)	0.0482*
11-30 days	24 (50.0%)	7 (41.2%)	0.5312
30 days	3 (6.2%)	7 (41.2%)	0.0024**

The proportion of patients with PAAL of different duration, broken down by non-severe and severe PAAL groups. PAAL, Prolonged alveolar air leak. *$p<0.05$. **$p<0.01$.

Patients with PAAL also were found to have a higher average number of complications in comparison with the control group (1.26 per patient vs 0.42 per patient, respectively; $p<0.05$). The breakdown of grades of complications is shown in **Table 1.5**, and the pie charts for grades and types of complications are shown in **Figure 1.1**. PAAL developed in 65 patients (18.1% of total); 48 patients had non-severe PAAL (grade II, 73.8% of patients with PAAL and 13.2% of all patients), and 17 patients had severe PAAL (grade ≥ III, 26.1% of patients with PAAL and 4.6% of all patients). With grade II PAAL, the patients required discharge with a chest tube or experienced a prolonged LOS and were managed with the chest drains placed at the time of pulmonary resection. However, a grade IV complication developed in 1 patient in the non-severe PAAL group, not as a result of PAAL, but as a result of pulmonary embolism, atelectasis, and pneumonia. Grade IIIA PAAL occurred in 12 patients (18.5% of PAAL, 3.4% of all), who required interventions such as bronchoscopy or insertion of additional chest drains. Grade IIIB occurred in 3 patients (4.6% of PAAL, 0.85% of all), who required a reoperation to control the air leak. Grade IV PAAL occurred in 2 patients (3.1% of PAAL, 0.57% of all), who were

admitted to the intensive care unit as a result of air leak. There was no mortality within 30 days of surgery in the PAAL group. The control group had significantly fewer grade IIIA complications and pulmonary and pleural complications than the PAAL group (p<0.05). The control group had more cardiac complications (41.3% control vs. 7.3% PAAL, p<0.05), and 40% of these were atrial fibrillation. Of the 82 complications in the PAAL group, most (90%) were pulmonary and pleural in nature. The associated complications of severe PAAL (17 patients) included empyema (n=2), pneumonia (n=2), hemothorax (n=1), and pulmonary embolism (n=1).

Table 1.5 Difference in complication rates between groups by severity.

Complication Grade	Control	PAAL	*p* Value
Grade I	10 (3.5)	0 (0)	0.1268
Grade II	86 (30.0)	58 (89.2)	<0.0001*
Grade III			
Grade IIIA	6 (2.1)	15 (23.1)	<0.0001*
Grade IIIB	5 (1.7)	6 (9.2)	0.0007*
Grade IV			
Grade IVA	9 (3.1)	3 (4.6)	0.5528
Grade IVB	2 (0.7)	0 (0)	0.4997
Grade V	3 (1.0)	0 (0)	0.4078
Total # complications	121 (42.2)	82 (126.2)	0.0013*
Total # patients	287	65	- -

Numeric counts of complications broken down by the TM&M grades in control and PAAL groups. PAAL, Prolonged alveolar air leak. *p<0.001.

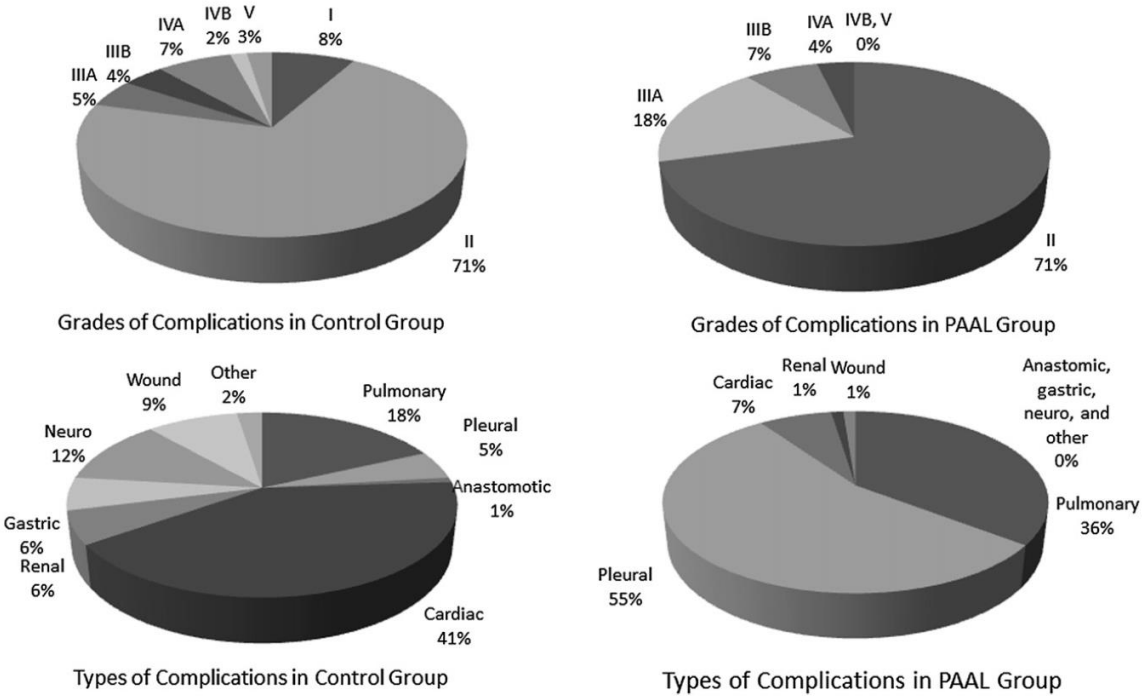


Figure 1.1 Types and grades of complications in control and PAAL groups.

DISCUSSION

The objectives of this study were 2-fold: to systematically quantify the burden of illness from PAAL using the TM&M system and to stratify risk factors for non-severe PAAL (grades I and II) and severe PAAL (grades III, IV, and V). We showed that as a whole, PAAL poses a burden for patients and hospital resources by prolonging the median LOS by 4 days and increasing the rate of readmission within 30 days by 20.4%. The TM&M system showed that the majority of PAAL cases are managed by chest drains inserted at the time of surgery, and 90% of associated complications with PAAL are limited to pleural and pulmonary complications, such as empyema and pneumonia. There is less association with complications in other organ systems, such as cardiac, anastomotic, and wound, which is similar to past studies.[15,16]

In the current study, we defined PAAL as the presence of forced expiratory air leak on POD 5. However, it is clear that there are other possibilities of defining PAAL. For example, an air leak on POD 3 may be considered PAAL if the patient underwent a single wedge resection only, because these patients are usually expected to be discharged sooner than those who underwent lobectomy or extended lobectomy. Nonetheless, given feedback from peers, in keeping with the majority of prior reports, and in compliance with Society of Thoracic Surgeons database, we elected to standardize the definition as air leak lasting greater than 5 days from the operation.

Identifying patients at risk for PAAL and severe PAAL helps design strategies to prevent these conditions. Of clinical importance, we found that emphysema seen on preoperative CT scan and obstructive pattern on PFT are significant risk factors for severe PAAL. However, emphysema seen on histopathology is not a significant risk factor. The CT and histopathology findings of emphysema agree approximately 40% of the time, and more patients were found to have emphysematous changes evident on histopathology. It is plausible that mild emphysematous changes that could be seen only under the microscope may not have as much impact on the post-operative course as emphysematous changes visible on CT scan. Therefore, CT findings of emphysema and obstructive pattern on PFT together can be used preoperatively to identify patients at risk for severe PAAL. The current study was not able to use multivariate analysis to identify predictors of severe PAAL because of the small sample size (4.6% of all patients). We suggest future studies to pool larger patient populations to properly perform such an analysis. In addition, validation from independent cohorts is necessary to evaluate reliability.

By delineating the risk factors and burden of illness, the cost versus benefit of various intraoperative prevention and post-operative treatment options can be weighed more accurately. Because the majority of PAAL cases are non-severe in nature and self-limited in time course, **Table 1.4**, these expensive measures may not be necessary for every patient. However, they could benefit patients at risk for severe PAAL, who are at risk for associated complications. Examples of intraoperative measures include the use of buttressed stapled lines with bovine pericardium (Bio-Vascular Dry Peri-Strips, Minneapolis, Minn), pleural tents for upper lobectomy, pneumoperitoneum after lower lobectomy, focal seal (Genzyme, Biosurgery, Cambridge, Mass), BioGlue (CryoLife, Europa Ltd, Surrey, UK), collagen patch, and so forth.[19-21] These methods also have their drawbacks other than cost and potential prolonged operative time. Pleural tent can cause bleeding, and synthetic materials may cause irritation and hypersensitivity. Although some randomized trials showed a reduction of post-operative arrhythmias, the meta-analysis by Malapert and colleagues [19] did not find any reduction in atelectasis, hemothorax, pneumonia, pneumothorax, and death by using glue, patch, or buttress. Staple-line buttress, fibrin glue, synthetic sealant, and collagen patch were not used in the current study. We are engaged in further studies to determine whether these intraoperative preventative measures could decrease the incidence of severe PAAL.

In addition to intraoperative prevention methods, patients who are at risk for severe PAAL should be closely monitored and managed aggressively to prevent further complications. **Table 1.4** shows the importance of properly managing severe PAAL because a significant portion of these patients had PAAL for extended periods of time compared with those with non-severe PAAL. The post-operative management of PAAL varies widely among institutions, and even

between surgeons of the same institution. One of the limitations of this study is that the amount of air leak in the evacuation chambers was not recorded. We hope that future studies will correlate the quantity of air leak with the severity of PAAL by using the scale by Cerfolio and colleagues.[17] The optimal algorithm of suction and water seal for the management of PAAL with analogue chest drainage systems remains highly controversial. Several randomized studies suggest the superiority of early water seal instead of ongoing suction, with respect to decreased days to resolution of air leak, LOS, and duration of indwelling chest drain.[17] Reduced suction can be achieved by continuous low suction setting or alternating "on" at night and "off" during the day. However, larger air leaks defined as bubbling greater than 4/7 by the scale used by Cerfolio and colleagues[17] may be better managed by suction to treat pneumothorax and subcutaneous emphysema. Past studies have shown that some form of reduced suction was safe with close monitoring. Outpatient 1-way valve chest drainage systems, such as the Heimlich or Pneumostat valve (Atrium Medical Corp), also have been shown to be safe and effective in asymptomatic patients with a small stable pneumothorax.[7] An idea for future work is to randomize post-operative suction versus underwater seal algorithms to determine whether either algorithm can reduce the incidence of severe PAAL.

Our data showed that the most notable complication in the control group was cardiac complication in the form of atrial fibrillation (rate of 40%). Despite the high prevalence of atrial fibrillation in this group, the median LOS was relatively short (median, 4 days, compared with 8 days in PAAL), and fewer complications developed in patients (0.42 per patient) compared with the PAAL group (1.26 per patient). It is unclear exactly why patients with PAAL had a lower incidence of atrial fibrillation compared with patients without PAAL. Past studies found that the

risk factors for post-operative arrhythmias are mostly related to the extent of pulmonary resection (especially pneumonectomy), hilar manipulation, and pre-existing heart disease.[18] In the current study, there was no significant difference in cardiac comorbidities between the groups, and the rates of atrial fibrillation may be artificially low because all cases of pneumonectomy were excluded.

Conclusion

Our study stratified the risk factors predicting non-severe and severe PAAL after pulmonary resection by using the TM&M classification system. Findings of emphysema on CT of the thorax along with an obstructive pattern on PFT were significant predictors of severe PAAL. We showed that the majority of PAAL cases in this series were nonsevere in degree and managed with a single chest tube inserted at the time of operation. In addition, patients with PAAL were more likely to have additional post-operative adverse events, and the majority of these were pleural or pulmonary in nature (empyema and pneumonia). Future research should focus on facilitating outpatient management and hospital resource-savings for non-severe PAAL cases. With respect to severe PAAL, further investigation is needed to examine the use of intraoperative preventative measures in those at risk and the resultant post-operative rates.

ACKNOWLEDGEMENTS

The authors thank Sarah Ramzan, Jennifer Threader, Dr. Ahmed Al-Hussaini, Dr. P. James Villeneuve, Dr. Anna McGuire, and Dr. Tim Ramsay for contributing to this article.

DISCLOSURES

No relevant financial relationships with commercial interests to disclose.

AUTHOR CONTRIBUTIONS

Study Conception and Design: Seely, Andrew; Liang, Shuyin; Ivanovic, Jelena

Acquisition of Data: Liang, Shuyin; Ivanovic, Jelena

Analysis and Interpretation of Data: Ivanovic, Jelena; Liang, Shuyin; Seely, Andrew

Drafting of Manuscript: Shuyin, Liang; Ivanovic, Jelena

Critical Revision: Shuyin, Liang; Ivanovic, Jelena; Seely, Andrew; Gilbert, Sebastian; Maziak, Donna; Shamji, Farid; Sundaresan, Sudhir

REFERENCES

1. Bardell T, Petsikas D. What keeps postpulmonary resection patients in hospital? Can Respir J 2003;10:86−9.
2. Irshad K, Feldman LS, Chu VF, et al. Causes of increased length of hospitalization on a general thoracic surgery service: A prospective observational study. Can J Surg 2002;45:264−68.
3. Brunelli A, Monteverde M, Borri A, et al. Predictors of prolonged air leak after pulmonary lobectomy. Ann Thorac Surg 2004;77:1205−10.
4. Abolhoda A, Liu D, Brooks A, Burt M. Prolonged air leak following radical upper lobectomy. Chest 1998;113:1507−10.
5. Stolz AJ, Schutzner J, Lischke R, et al. Predictors of prolonged air leak following pulmonary lobectomy. Eur J Cardiothorac Surg 2005;27:334−36.
6. Brunelli A, Varela G, Refai M, et al. A scoring system to predict the risk of prolonged air leak after lobectomy. Ann Thorac Surg 2010;90:204−09.
7. Cerfolio RJ, Bass CS, Pask AH, Katholi CR. Predictors and treatment of persistent air leaks. Ann Thorac Surg 2002;73:1727−31.
8. Linden PA, Bueno R, Colson YL, et al. Lung resection in patients with preoperative FEV1 < 35% predicted. Chest 2005;127:1984−90.
9. Singhal S, Ferraris VA, Bridges CR, et al. Management of alveolar air leaks after pulmonary resection. Ann Thorac Surg 2010;89:1327−35.
10. Varela G, Jimqnez MF, Novoa N, Aranda JL. Estimating hospital costs attributable to prolonged air leak in pulmonary lobectomy. Eur J Cardiothorac Surg 2005;27:329−33.
11. Massone P, Magnani B, Conti B, et al. Cauterization versus fibrin glue for aerostasis in precision resections for secondary lung tumors. Ann Surg Oncol 2003;10:441−46.
12. Seely AJ, Ivanovic J, Threader J, et al. Systematic classification of morbidity and mortality after thoracic surgery. Ann Thorac Surg 2010;90:936−42.
13. Ivanovic J, Al-Hussaini A, Al-Shehab D, et al. Evaluating the reliability and reproducibility of the Ottawa Thoracic Morbidity and Mortality classification system. Ann Thorac Surg 2011;91:387−93.
14. Clavien PA, Sanabria JR, Strasberg SM. Proposed classification of complications of surgery with examples of utility in cholecystectomy. Surg 1992;111:518−26.
15. Okereke I, Murthy SC, Alster JM, et al. Characterization and importance of air leak after lobectomy. Ann Thorac Surg 2005;79:1167−73.
16. Brunelli A, Xiume F, Al Refai M, et al. Air leaks after lobectomy increase the risk of empyema but not of cardiopulmonary complications. Chest 2006;130:1150−56.
17. Cerfolio RJ, Bass C, Katholi CR. Prospective randomized trial compares suction versus water seal for air leaks. Ann Thorac Surg 2001;71:1613−17.
18. Vaporciyan AA, Correa AM, Rice DC, et al. Risk factors associated with atrial fibrillation after noncardiac thoracic surgery: Analysis of 2588 patients. J Thorac Cardiovasc Surg 2004;127:779−86.
19. Malapert G, Hanna HA, Pages PB, Bernard A. Surgical sealant for the prevention of prolonged air leak after lung resection: Meta-analysis. Ann Thorac Surg 2010;90:1779−85.
20. Fabian T, Federico JA, Ponn RB. Fibrin glue in pulmonary resection: A prospective, randomized, blinded study. Ann Thorac Surg 2003;74:1587−92.

21. Belcher E, Dusmet M, Jordan S, et al. A prospective, randomized trial comparing BioGlue and Vivostat for the control of alveolar air leak. J Thorac Cardiovasc Surg 2010;140:32−38.

Commentary

Post-operative atrial fibrillation (PAF) is the most common sustained arrhythmia, the second most common complication after pulmonary resection, and a major, potentially preventable, adverse outcome.[1] PAF has a significant burden on the patient, and is often associated with longer intensive care unit and hospital stays, increased morbidity and mortality, as well as, higher resource utilization. The identification of risk factors for PAF is paramount in order to direct prophylactic measures to the appropriate patients. In the following study, we have attempted to quantify the incidence and severity of PAF using the Thoracic Morbidity & Mortality (TM&M) classification system, and to identify risk factors for PAF.

The following novel information has contributed to the literature: standardized definitions for uncomplicated/transient (≤7 days) or complicated/persistent (>7 days) electrocardiographically documented PAF requiring initiation of pharmacological therapy (Grade II). Likewise, since their publication, the results have been applied to the 2014 American Association for Thoracic Surgery guidelines for the prevention and management of PAF and flutter for thoracic surgical procedures.[2]

The Ottawa Hospital's Division of Thoracic Surgery has also embarked upon different measures to prevent PAF in patients undergoing pulmonary resection for lung cancer through detailed discussions of best practice measures at divisional positive deviance seminars. This has allowed

for the targeted management of patients with increased risk of developing this common complication.

REFERENCES

1. Seely AJ, Ivanovic J, Threader J, et al. Systematic classification of morbidity and mortality after thoracic surgery. Ann Thorac Surg 2010;90:936−42.
2. Frendl G, Sodickson AC, Chung MK, et al. 2014 AATS guidelines for the prevention and management of perioperative atrial fibrillation and flutter for thoracic surgical procedures. J Thorac Cardiovasc Surg 2014;148:153−93.

Chapter II

Incidence, severity, and perioperative risk factors for atrial fibrillation following pulmonary resection

Interact Cardiovasc Thorac Surg 2014;18:340–346

Presented at the 2013 Canadian Society for Epidemiology and Biostatistics Conference, St. John's, Newfoundland, Canada, June 2013

Jelena Ivanovic, MSc;[1,3,4]
Donna E. Maziak, MDCM, MSc, FRCSC, FACS;[1,3,4]
Sarah Ramzan, MD;[2]
Anna L. McGuire, MD, FRCSC;[3]
P. James Villeneuve, MDCM, PhD, FRCSC;[3]
Sebastien Gilbert, MD, FRCSC;[3]
R. Sudhir Sundaresan, MD, FRCSC, FACS;[3,4]
Farid M. Shamji, MBBS, FRCSC;[3]
and Andrew J.E. Seely, MD, PhD, FRCSC.[1,3,4]

Department of Epidemiology and Community Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Canada[1]
Faculty of Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Canada[2]
Division of Thoracic Surgery, Department of Surgery, The Ottawa Hospital, 501 Smyth Road, Ottawa, Canada[3]
Clinical Epidemiology Program, Ottawa Hospital Research Institute, 501 Smyth Road, Ottawa, Canada[4]

ABSTRACT

OBJECTIVES: Post-operative atrial fibrillation (PAF) occurs commonly following pulmonary resection. Our aims were to quantify the incidence and severity of PAF using the Thoracic Morbidity & Mortality classification system, and identify risk factors for PAF.

METHODS: All consecutive patients undergoing pulmonary resection at a single centre (January 2008 - April 2010) were enrolled. PAF was defined as post-operative, electrocardiographically documented, and requiring initiation of pharmacological therapy. Univariate and multivariate analyses of risk factors associated with the development of PAF were conducted.

RESULTS: The incidence of PAF was 11.8% (n=43) of 363 pulmonary resections (open: n=173; 47.7%; video-assisted: n=177; 48.8%; converted: n=13; 3.6%): sublobar (n=93; 25.6%), lobectomy (n=237; 65.3%), bilobectomy (n=7; 1.9%) and pneumonectomy (n=24; 6.6%). Twenty-eight cases (65.1%) were uncomplicated/transient, and 15 cases (34.9%) were complicated/persistent PAF, defined as lasting for >7 days (40.0%), requiring cardioversion (13.3%), vasopressors (33.3%), in-hospital use of anticoagulants (46.7%) and/or anticoagulants on discharge (26.7%). Patients with PAF had increased mean lengths of hospital stay (10.5 days vs 6.9 days; p=0.04). Peak onset of PAF occurred 2.5 (standard deviation (SD) ± 1.3) days after pulmonary resection, lasting for 1.8 ± 2.8 (mean, ±SD) days. Multivariate analysis identified (relative risk; 95% confidence interval): age ≥70 years (2.3; 1.1–5.1), history of angioplasty/stents/angina (4.0; 1.4–11.3), thoracotomy (3.6; 1.4–9.3), conversion to open thoracotomy (16.5; 2.2–124.0) and extent of surgery/stage (7.1; 1.0–49.4) as predictors of PAF.

CONCLUSIONS: While the majority of PAF is uncomplicated and transient, one-third of cases lead to persistence or major intervention. Age, coronary artery disease, and extent of surgery/stage increase the risk of PAF following pulmonary resection. Identifying patients with elevated risk may lead to targeted prophylaxis to reduce the incidence of PAF.

INTRODUCTION

Post-operative atrial fibrillation (PAF) has remained one of the most common complications that occur following non-cardiac thoracic surgery. Although it is difficult to determine the true incidence of PAF due to various methodologies used to identify its occurrence, reported rates have varied between 4% and 37%.[1-3] The occurrence of PAF is associated with significant morbidity, such as increased risk of stroke, atrial thrombosis and systemic embolization, post-operative mortality, and significant increases in hospital length of stay (LOS) and costs.[1,4] Accordingly, several prophylactic antiarrhythmic drug treatments have been proposed in an attempt to reduce the incidence of PAF.[5-7]

In addition, a number of efforts have been made at identifying risk factors for PAF.[1,8-13] Risk factors of PAF have included increasing age, male sex, pre-existing history of ischaemic heart disease, valvular heart disease, congestive heart failure, peripheral vascular disease, intrapericardial dissection, increasing extent of pulmonary resection, intraoperative blood transfusions, post-operative electrolyte imbalance and hypoxia.[1,8-12] A definite causative relationship between PAF development and preoperative and intraoperative risk factors has not yet been firmly established. Moreover, past studies have not used a standardized and validated definition of PAF; nor have previous studies attempted to stratify the severity of disease burden associated with PAF.

In accordance with the Clavien–Dindo classification system of adverse events,[14] we developed a standardized system to identify both the presence and severity of Thoracic Morbidity & Mortality (TM&M).[14,15] The TM&M classification system was implemented at The Ottawa

Hospital in January 2008, and is a prospective database documenting all complications and their severity for all thoracic surgical procedures. The TM&M database is created to prospectively record post-operative adverse event information, and it provided an essential platform for the current study, as the onset of PAF most commonly occurs in the first 7 days following pulmonary resection. Thus, using the TM&M classification system, our objectives were to quantify the incidence and severity of PAF, and identify preoperative and intraoperative risk factors for PAF following pulmonary resection.

MATERIALS & METHODS

Study population

A retrospective review of the prospectively collected TM&M database was conducted. All consecutive patients who underwent pulmonary resection for benign or malignant disease by The Ottawa Hospital's Division of Thoracic Surgery between January 2008 and April 2010 were considered for inclusion into this study. Among the 371 patients initially identified, 4 patients were excluded because they did not undergo formal preoperative assessment. Of the remaining 367 patients, 13 more were excluded due to a history of AF. If a patient underwent multiple separate pulmonary resections during the study period, data were collected and analysed for each surgery separately. Data were collected on patient demographics, medical history, surgical history, active medications, neoadjuvant chemotherapy and/or radiotherapy, preoperative cardiopulmonary functional testing, operative details and post-operative adverse events. Data on PAF characteristics and hospital LOS were also recorded. This study was approved by the Ottawa Hospital Research Ethics Board.

Classification of post-operative adverse events, and definition and monitoring of PAF

The incidence and severity of PAF were classified prospectively during the study period using the TM&M classification system, which grades adverse events on a severity scale from Grade I to V based on the effort required to treat the event.[14]

PAF was defined as uncomplicated/transient (≤7 days) or complicated/persistent (>7 days) electrocardiographically documented PAF requiring initiation of pharmacological therapy (Grade II). PAF was considered complicated/persistent if it lasted for >7 days and/or if the patient required: (i) cardioversion or vasopressors (Grade III); (ii) intensive care unit (ICU) admission (Grade IV); (iii) in-hospital use of anticoagulants and/or (iv) the patient was discharged on amiodarone, diltiazem, digoxin or anticoagulants. During the post-operative period, clinical examination of patients was performed daily until hospital discharge and electrocardiogram (ECG) monitoring was performed daily or when required by clinical examination.

Statistical analysis

Summary statistics for continuous variables were recorded as medians and means, and analysed using Student's *t*-test; categorical data were summarized as frequencies and percentages, and comparisons between the two groups were performed with the Pearson $\chi 2$ test or Fisher exact test. A *p*-value of <0.05 was considered statistically significant. Stepwise logistic regression analysis was used to determine independent correlates for both PAF and complicated/persistent PAF development. A two-sided *p*-value of <0.05 was considered statistically significant for inclusion into the multivariate model. Discrimination and calibration of the model were assessed using the C-statistic and the Hosmer–Lemeshow goodness of fit test. Data collection was

performed using paper case-report forms, data entry was performed using Excel (Microsoft, Redmond, WA, USA) and data analyses were conducted using IBM SPSS Statistics Version 20® software (SPSS, Chicago, IL, USA).

RESULTS

Baseline features and outcomes

During the study period, a total of 354 patients underwent 363 pulmonary resections. A total of 195 (53.7%) patients were female and 139 (38.3%) patients were over the age of 70. Disease diagnoses included primary lung cancer (n=279; 76.9%), pulmonary metastasis (n=71; 19.6%) and benign lung disease (n=13; 3.6%). A comparison of preoperative risk factors in patients with and those without PAF is presented in **Table 2.1** and **Table 2.2**.

Table 2.1 Patient demographics, medical history, and pre-operative evaluation.

	– PAF (n = 320), %	+ PAF (n = 43), %	*p* Value
Patient characteristics			
Age			0.01
≤69	205 (64)	19 (44)	
≥70	115 (36)	24 (56)	
Sex			0.45
Male	149 (47)	19 (44)	
Female	171 (53)	24 (56)	
BMI, median	26.2	26.3	0.85
Past medical history			
Cardiac risk factors			
Angioplasty/Stents/Angina	22 (7)	9 (21)	0.01*
Myocardial infarction	21 (7)	4 (9)	0.51
Cerebrovascular disease	16 (5)	2 (5)	0.64
Peripheral vascular disease	24 (8)	2 (5)	0.50
Valvular disease	15 (5)	3 (7)	0.52
Other risk factors			
Diabetes mellitus	35 (11)	7 (16)	0.30
Previous cancer	125 (39)	14 (33)	0.41
Previous cardiac surgery	18 (6)	4 (9)	0.34
Previous thoracic surgery	17 (5)	1 (2)	0.40
Preoperative chemotherapy	7 (2)	3 (7)	0.07
Preoperative radiation therapy	3 (1)	1 (2)	0.41
Lung function tests			
$FEV_{1\ predicted,\ median,\ \%}$	82.0	82.5	0.70
$FVC_{\ predicted,\ median,\ \%}$	89.0	83.0	0.21
$FEV_1/FVC\%$	70.0	69.0	0.51
$DLCO_{\ predicted,\ median,\ \%}$	78.0	83.0	0.06
Echocardiographic variables	(n = 258)	(n = 33)	
Left systolic dysfunction	13 (5)	6 (18)	<0.05*
Left diastolic dysfunction	23 (9)	6 (18)	0.09
Left atrial enlargement	12 (5)	3 (9)	0.28
Left ventricular hypertrophy	6 (2)	2 (6)	0.2
Pulmonary hypertension	11 (4)	1 (3)	0.60
Atrial septal defect	3 (1)	2 (6)	0.04*
Valvular pathology	23 (9)	5 (15)	0.25
Diagnosis			0.74
Benign	12 (8)	1 (2)	
Malignant NSCLC	244 (76)	35 (81)	
Other Malignant	64 (20)	7 (16)	

Overall incidence of PAF was 11.8% (n=43), including 28 cases (65.1%) of uncomplicated/transient PAF and 15 cases (34.9%) of complicated/persistent PAF. Approximately one-quarter (23.3%) of patients with PAF experienced PAF during the first 24 h. The peak onset of PAF occurred 2.5 (standard deviation (SD) ± 1.3) days after pulmonary resection, lasting for a mean duration of 1.8 (SD ± 2.8) days. Sixteen patients (37.2%) developed PAF after Day 3. Five patients (11.6%) had PAF after Day 5.

Table 2.2 Patient demographics, medical history, and pre-operative evaluation.

	– PAF (n = 320), %	Uncomplicated/ Transient PAF (n = 28), %	*p* Value	Complicated/ Persistent PAF (n = 15), %	*p* Value
Patient characteristics					
Age			0.01*		0.40
≤69	205 (64)	11 (39)		8 (53)	
≥70	115 (36)	17 (61)		7 (47)	
Sex					0.31
Male	149 (47)	18	0.18	6 (40)	
Female	171 (53)	10		9 (60)	
BMI, median	26.2	25.8	0.49	28.2	0.38
Past medical history					
Cardiac risk factors					
Angioplasty/Stents/Angina	22 (7)	5 (18)	0.24	11(73)	0.02*
Myocardial infarction	21 (7)	1 (4)	0.46	3 (20)	0.08
Cerebrovascular disease	16 (5)	1 (4)	0.45	0 (0)	0.77
Peripheral vascular disease	24 (8)	0 (0)	0.12	2 (13)	0.33
Valvular disease	15 (5)	3 (11)	0.17	0 (0)	0.50
Other risk factors					
Diabetes mellitus	35 (11)	4 (14)	0.39	3 (20)	0.23
Previous cancer	125 (39)	9 (32)	0.31	5 (33)	0.66
Previous cardiac surgery	18 (6)	2 (7)	0.49	2 (13)	0.22
Previous thoracic surgery	17 (5)	0 (0)	0.23	1 (7)	0.57
Preoperative chemotherapy	7 (2)	0 (0)	0.55	3 (20)	0.01*
Preoperative radiation therapy	3 (1)	0 (0)	0.77	1 (7)	0.17
Lung function tests					
$FEV_{1\ predicted,\ median,\ \%}$	82.0	82.5	0.39	82.0	0.74
$FVC_{\ predicted,\ median,\ \%}$	89.0	83.5	0.47	91.0	0.84
$FEV_1/FVC\%$	70.0	69.0	0.64	70.0	0.52
$DLCO_{\ predicted,\ median,\ \%}$	78.0	83.5	0.05*	83.0	0.18
Echocardiographic variables	(n = 258)	(n = 21)		n = 12	
Left systolic dysfunction	13 (5)	3 (14)	0.11	3 (25)	0.03*
Left diastolic dysfunction	23 (9)	4 (19)	0.13	2 (17)	0.31
Left atrial enlargement	12 (5)	1 (5)	0.65	2 (17)	0.12
Left ventricular hypertrophy	6 (2)	2 (10)	0.12	0 (0)	0.76
Pulmonary hypertension	11 (4)	1 (5)	0.62	0 (0)	0.60
Atrial septal defect	3 (1)	1 (5)	0.27	1 (8)	0.17
Valvular pathology	23 (9)	1 (5)	0.44	4 (33)	0.02*
Diagnosis			0.52		0.92
Benign	12 (8)	1 (4)		0 (0)	
Malignant NSCLC	244 (76)	23 (82)		12 (80)	
Other Malignant	64 (20)	4 (14)		3 (20)	

Operative characteristics

The types of procedures performed were sublobar resection (n=93; 25.6%), lobectomy (n=237; 65.3%), bilobectomy/extended lobectomy (n=7; 1.9%) and pneumonectomy (n=24; 6.6%). The majority of the operations were approached through open thoracotomy (n=173; 47.7%). A comparison of operative characteristics, intraoperative complications, and pathological staging in patients with and without PAF is presented in **Table 2.3** and **Table 2.4**.

Table 2.3 Operative characteristics, intraoperative complications, pathologic staging, hospital length of stay.

	– PAF (n = 320), %	+ PAF (n = 43), %	*p* Value
Approach			<0.05
Thoracoscopy	167 (52)	9 (21)	
Thoracotomy	143 (45)	30 (70)	
Converted	10 (3)	4 (9)	
Procedure Performed			
Sub-lobar resection	89 (28)	4 (9)	<0.05
Lobectomy	213 (67)	30 (70)	0.41
Right upper lobectomy	80 (38)	8 (27)	0.25
Right middle lobectomy	16 (8)	0 (0)	0.12
Right lower lobectomy	32 (15)	7 (23)	0.18
Left upper lobectomy	54 (25)	10 (33)	0.35
Left lower lobectomy	31 (15)	5 (17)	0.76
Bilobectomy/Extended lobectomy	7 (2)	0 (0)	0.33
Pneumonectomy	15 (5)	9 (21)	<0.05
Left	11 (73)	4 (44)	0.16
Right	4 (27)	5 (56)	0.16
Unresectable	2 (1)	0 (0)	0.60
Intraoperative Complications	18 (6)	7 (16)	0.02
Bleeding	7 (2)	4 (9)	0.01*
Hypotension	3 (1)	2 (5)	0.05*
Hypoxemia	2 (1)	0 (0)	0.60
Open/Close	4 (1)	0 (0)	0.46
AJCC 7th edition 2009 stage	n = 239	n = 35	0.01
IA/IB	151 (63)	16 (46)	0.05
IIA/IIB	56 (23)	11 (31)	0.30
IIIA/IIIB	28 (12)	4 (11)	0.96
IV	4 (2)	4 (11)	<0.05
Hospital length of stay, mean, days	6.9	10.5	0.04

The comparison of PAF occurring in total number of patients operated on for malignant non-small-cell lung cancer (NSCLC) or other lung diseases revealed a trend towards more frequent PAF in patients with malignant NSCLC (data not shown). Similarly, the comparison of PAF occurring in total number of patients operated on for malignant NSCLC or other lung diseases revealed a trend towards more frequent PAF in patients who underwent pneumonectomy.

Table 2.4 Operative characteristics, intraoperative complications, pathologic staging, hospital length of stay.

	– PAF (n = 320), %	Uncomplicated/ Transient PAF (n = 28), %	*p* Value	Complicated/ Persistent PAF (n = 15), %	*p* Value
Approach			0.02*		<0.05
Thoracoscopy	167 (52)	7 (25)		2 (13)	
Thoracotomy	143 (45)	19 (68)		11 (73)	
Converted	10 (3)	2 (7)		2 (13)	
Procedure Performed					
Sub-lobar resection	89 (28)	3 (11)	0.05	1 (7)	0.07
Lobectomy	213 (67)	20 (40)	0.38	10 (67)	0.99
Right upper lobectomy	80 (38)	6 (30)	0.50	3 (30)	0.63
Right middle lobectomy	16 (8)	0 (0)	0.20	0 (0)	0.37
Right lower lobectomy	32 (15)	6 (30)	0.08	1 (10)	0.66
Left upper lobectomy	54 (25)	5 (25)	0.92	4 (40)	0.30
Left lower lobectomy	31 (15)	3 (15)	0.95	2 (20)	0.63
Bilobectomy/Extended lobectomy	7 (2)	0 (0)	0.63	0 (0)	0.33
Pneumonectomy	15 (5)	5 (18)	0.02	4 (27)	0.01
Left	11 (73)	3 (60)	0.32	1 (25)	0.08
Right	4 (27)	2 (40)	0.57	3 (75)	0.08
Unresectable	2 (1)	0 (0)	0.67	0 (0)	0.76
Intraoperative Complications	18 (6)	3 (11)	0.23	4 (27)	0.01
Bleeding	7 (2)	1 (4)	0.50	3 (20)	0.01
Hypotension	3 (1)	1 (4)	0.29	1 (7)	0.17
Hypoxemia	2 (1)	0 (0)	0.85	0 (0)	0.91
Open/Close	4 (1)	0 (0)	0.71	0 (0)	0.83
AJCC 7th edition 2009 stage	n = 239	n = 23	0.28	n = 12	<0.05
IA/IB	151 (63)	11 (48)	0.11	5 (42)	0.13
IIA/IIB	56 (23)	9 (39)	0.08	2 (17)	0.59
IIIA/IIIB	28 (12)	2 (9)	0.49	2 (17)	0.61
IV	4 (2)	1 (4)	0.37	3 (25)	<0.05
Hospital length of stay, mean, days	6.9	6.7	0.92	17.7	<0.05

Univariate analysis to identify risk factors associated with PAF following pulmonary resection

Univariate predictors of PAF are outlined in **Table 2.5**. Predictors that are associated with PAF include (odds ratio; 95% confidence interval): age ≥70 (2.3; 1.2–4.3), cardiac history positive for: angioplasty/stents/angina (3.6; 1.5–8.4), left systolic dysfunction (4.2; 1.5–11.9) and atrial septal defect (5.5; 0.9–34.1); surgical approach: open thoracotomy (3.9; 1.8–8.5) and thoracoscopic surgery converted to open thoracotomy (7.4; 1.9–28.3); increasing extent of pulmonary resection: pneumonectomy (13.4; 3.6–48.9); presence of intraoperative complications, such as bleeding (4.6; 1.3–16.3) and hypotension (5.2; 0.8–31.8), and higher American Joint Committee on Cancer (AJCC) 7th edition 2009 stage: IV (7.6; 1.8–31.9).

Risk factors identified by univariate subanalysis for complicated/persistent PAF were similar to those described above. Age ≥70 years, however, did not prove to be a significant risk factor for complicated/persistent PAF upon univariate analysis (data not shown).

Table 2.5 Results of univariate analysis to identify pre-operative and intra-operative risk factors associated with post-operative atrial fibrillation after pulmonary resection.

Variable	Odds Ratio	95% CI Lower	95% CI Upper	*p* Value
Demographics				
Age				0.01
≤69	1.0			
≥70	2.3	1.2	4.3	0.01
Medical History				
Angioplasty/Stents/Angina	3.6	1.5	8.4	0.01
Echocardiographic Variables				
Left systolic dysfunction	4.2	1.5	11.9	<0.05
Atrial septal defect	5.5	0.9	34.1	0.04
Surgical Approach				<0.05
Thoracoscopy	1.0			
Thoracotomy	3.9	1.8	8.5	<0.05
Converted	7.4	1.9	28.3	0.05
Procedure Performed				<0.05
Sub-lobar resection	1.0			
Lobectomy	3.2	1.1	9.4	0.68
Pneumonectomy	13.4	3.6	48.9	<0.05
Intraoperative complications	3.3	1.3	8.3	0.02
Bleeding	4.6	1.3	16.3	0.01
Hypotension	5.2	0.8	31.8	0.05
AJCC 7th edition 2009 stage				<0.05
IA/IB	1.0			
IIA/IIB	1,5	0.7	3.2	0.30
IIIA/IIIB	1.0	0.3	3.0	0.96
IV	7.6	1.8	31.9	<0.05

Multivariate analysis to identify risk factors associated with PAF following pulmonary resection

Stepwise logistic regression analysis was used to determine independent correlates of PAF using the variables listed in **Table 2.6**. Significant predictors of PAF in multivariate analysis included (relative risk; 95% confidence interval): age ≥70 years (2.3; 1.1–5.1), history of angioplasty/stents/angina (4.0; 1.4–11.3), thoracotomy (3.6; 1.4–9.3), conversion to open thoracotomy (16.5; 2.2–124.0) and higher AJCC 7th edition 2009 stage (7.1; 1.0–49.4). The presence of intraoperative complications was not a significant predictor for PAF development

upon multivariate analysis. The final model had a C-statistic of 0.81, Hosmer–Lemeshow $\chi 2$ value of 2.9 and P-value of 0.89 (**Table 2.6**).

Table 2.6 Results of multivariate analysis to identify pre-operative and intra-operative risk factors associated with post-operative atrial fibrillation after pulmonary resection.

Variable	Relative Risk	95% CI Lower	95% CI Upper	*p* Value
Demographics				
Age				
≤69	1.0			
≥70	2.3	1.1	5.1	0.04
Medical History				
Angioplasty/Stents/Angina	4.0	1.4	11.3	0.01
Surgical Approach				<0.05
Thoracoscopy	1.0			
Thoracotomy	3.7	1.5	9.3	0.01
Converted	16.5	2.2	124.0	0.01
Intraoperative Complications	1.1	0.3	4.6	0.90
AJCC 7th edition 2009 stage				0.12
IA/IB	1.0			
IIA/IIB	2.0	0.8	4.9	0.12
IIIA/IIIB	1.0	0.3	3.7	0.95
IV	7.1	1.0	49.4	0.05

C Statistic, 0.81; Hosmer and Lemeshow Goodness of Fit Test, p=0.89

Multivariate subanalysis of risk factors identified for complicated/persistent PAF were similar to those described above (data not shown). Increasing age, thoracotomy and the presence of intraoperative complications were not significant predictors for complicated/persistent PAF development upon multivariate analysis.

Clinical course, management, and sequela **of PAF**

The majority of PAF in this series were uncomplicated/transient in nature and were managed with pharmacological therapy alone (Grade II) (**Table 2.7**). Uncomplicated/transient PAF was

successfully managed with pharmacological therapy: 96.4% (n=27) cases were treated with a beta-blocking agent such as metoprolol and 21.4% (n=4) were treated with digoxin. Two cases (7.1%) required in-hospital treatment with amiodarone. A total of 19 cases (67.9%) with uncomplicated/transient PAF were discharged from hospital on metoprolol.

Table 2.7 Post-operative atrial fibrillation characteristics.

Characteristic	+ PAF (n = 43), %	Uncomplicated/Transient +PAF (n = 28), %	Complicated/Persistent + PAF (n = 15), %
Paroxysmal	37 (86.0)	28 (100)	9 (60.0)
Persistent	6 (14.0)	0 (0)	6 (40.0)
Cardioversion	2 (4.7)	0 (0)	2 (13.3)
Vasopressors	5 (11.6)	0 (0)	5 (33.3)
Anticoagulants in hospital	7 (16.3)	0 (0)	7 (46.7)
Metoprolol in hospital	41 (95.3)	27 (96.4)	14 (93.3)
Amiodarone in hospital	8 (18.6)	2 (7.1)	6 (40.0)
Diltiazem in hospital	2 (4.7)	0 (0)	2 (13.3)
Digoxin in hospital	19 (44.2)	6 (21.4)	13 (86.7)
Anticoagulants on discharge	4 (9.3)	0 (0)	4 (26.7)
Metoprolol on discharge	32 (74.4)	19 (67.9)	13 (86.7)
Amiodarone on discharge	2 (4.7)	0 (0)	2 (13.3)
Diltiazem on discharge	0 (0)	0 (0)	0 (0)
Digoxin on discharge	10 (23.3)	0 (0)	10 (66.7)
Discharge with PAF	2 (4.7)	0 (0)	2 (13.3)
LOS > 5 days	29 (67.4)	16 (57.1)	13 (86.7)

Complicated PAF was persistent in 6 cases (40.0%), required cardioversion (Grade III) in 2 cases (13.3%) and vasopressors haemodynamic support with ICU admission (Grade IV) in 5 cases (33.3%). With respect to anticoagulation for complicated/persistent PAF, 7 patients (46.7%) were started on anticoagulants in hospital due to increased risk of stroke following cardiology

assessment, and 4 patients (26.7%) continued anticoagulation on discharge from hospital. Two cases (4.7%) displayed persistent PAF on discharge from hospital.

Patients with PAF had increased mean hospital LOS (10.5 days vs. 6.9 days; p=0.04) (**Table 2.2**). Approximately 11.6% (n=5) of patients with PAF required hospital readmission following initial discharge in comparison with 7.2% (n=23) in the control group (p=0.31); an additional 11.6% (n=5) required admission to the ICU for management; no patients in the control group were admitted to the ICU for management (p<0.05).

Approximately 20% (n=3) of patients with complicated/persistent PAF required hospital readmission following initial discharge; an additional 33.3% (n=5) required admission to the ICU for management (data not shown).

PAF was associated with a greater incidence of additional post-operative complications in comparison with the control group: 1.58 per patient vs. 0.48 per patient (p<0.05), respectively (data not shown). There was 1 death (1.5%) reported in the PAF group, and 4 deaths (2.6%) in the control group (p=0.47) (data not shown).

DISCUSSION

Despite ongoing efforts to decrease its occurrence, PAF remains the most common cardiac complication in patients following pulmonary resection.[7] The incidence of PAF in our study was 11.8% and was lower in comparison with other published reports.[8,13,16] This difference in incidence rates can be explained by the absence of a uniform definition of PAF, the oversight of

its sometimes transient nature, and the various methodologies used to record its occurrence. In the present study, a standardized definition of PAF was applied requiring electrocardiographic evidence, initiation of pharmacological therapy, and prospective monitoring and documentation of PAF using the TM&M classification system. The TM&M classification system facilitates monitoring, reporting and evaluation of surgical adverse events, and has recently been evaluated for its inter-rater agreement.[17]

Several risk factors have been identified for the development of PAF.[1,8-13] In the current study, age ≥70 years, history of angioplasty/stents/angina, open thoracotomy surgical approach, conversion and higher AJCC 7th edition 2009 stage/extent of surgery were significantly associated with PAF development on multivariate analysis. First, when preoperative risk factors are taken into account, increasing age has been the most consistent predictor of PAF due to age-related structural changes in atrial connective tissue, dilatation and irregular anisotropic conduction.[12,18]

Secondly, higher AJCC 7th edition 2009 stage/increasing extent of pulmonary resection was associated with a significantly higher incidence of PAF compared with lesser resections. It has been suggested that the removal of one lung can increase the cardiac function with ventricular dilatation, increased right-heart pressure and transient pulmonary hypertension.[19] Laterality of pneumonectomy has also been linked with higher incidence of PAF development.[16] Laterality of pneumonectomy was not a significant predictor of PAF development in our data. Of note, however, patients who underwent pneumonectomy represented a small proportion of the total patient population in the current study, limiting the comparison.

Thirdly, in contrast to results of previously reported studies,[20] this study revealed a significantly higher incidence of PAF following open thoracotomy compared with the minimally invasive surgical approach; suggesting that incision-related effects may be responsible for the pathogenesis of PAF in high-risk patients.

Fourthly, the presence, extent and severity of coronary artery disease have not been a consistent predictor of PAF in the literature. The current study did, however, identify previous percutaneous transluminal coronary angioplasty and stent placement and pre-existing history of angina to be associated with an increased risk of PAF development.

Significant risk factors for complicated/persistent PAF identified on multivariate analysis were similar to those found in the overall PAF population, including: conversion to open thoracotomy, cardiac history positive for: angina, angioplasty and coronary artery stenting, and higher AJCC 7th edition 2009 stage/extent of pulmonary resection. Increasing age did not prove to be a significant risk factor for complicated/persistent PAF development. Due to the small sample size of the dependent variable, however, these results should be interpreted with caution.

The most common time for onset of PAF is during the first 24 h following major thoracic surgery.[21] Approximately one-quarter (23.3%) of our patients with PAF experienced PAF during the first 24 h with a peak incidence on post-operative Day 2. The majority of our patients (37.2%), however, developed PAF after Day 3. These data suggest that extended ECG monitoring should be performed in higher risk patients for timely diagnosis and prevention of additional complications of PAF.

In addition to increased morbidity and mortality, previous studies have found that PAF is associated with increased hospital LOS.[1,4,10,13] Consistent with these findings, the current study found PAF to be associated with prolonged hospital LOS in the majority of patients (67%). We were also able to document the rate of PAF readmission and its significant association with a greater complication rate.

Professional opinions remain at odds whether PAF should be treated with beta-blockers that reduce the adrenergic response or with calcium-channel blockers that decrease the pulmonary hypertensive response.[22] Medications most commonly used to treat PAF in our series included beta-blockers such as metoprolol (95.4%) and digoxin (46.5%). The safety and efficacy of other pharmacological therapies for PAF prophylaxis have also been studied. In high-risk patients, only amiodarone and diltiazem have been shown to provide effective prophylaxis against AF.[5-7,18,23] Amiodarone and diltiazem, however, were not commonly used in this study for management of PAF as beta-blockers were found to successfully control the rhythm disturbance in the majority of cases. Moreover, it is not our standard of care to use prophylactic antiarrhythmic treatment.

This study has several limitations. Perioperative imbalances in serum electrolytes have been associated with the development of PAF.[24] Data on perioperative potassium and magnesium levels were not collected in this study, nor were data on volemia, hypoxia, anemia and hypothermia, all of which have been associated with increased sympathetic activity and PAF.[24] Although our patients were selected from a prospective database, it remained a retrospective review, and is subject to limitations attributed to retrospective studies, including difficulties in

controlling bias and confounders and establishing cause and effect relationships. Small sample size of the dependent variable and the use of a single institution may limit the generalizability of our results to other institutions as well.

There are several strengths to this study. First, a frequent post-operative complication after pulmonary resection was explored and our findings are similar to those of previously published studies. Secondly, we believe that the reported incidence of PAF was accurate, as routine ECG monitoring was employed in all patients, untreated transient PAF was not included in the analysis, and a standardized and validated definition of PAF was used. Thirdly, while the majority of PAF cases were classified as uncomplicated/transient and were well-controlled with minimal intervention in this study, complicated/persistent PAF accounted for 34.9% of cases warranting further discussion. Preoperative risk stratification may be used (i) to better inform the surgical patient of the risk of PAF development; (ii) for patient selection for whom prophylactic drug therapy might be most safe and beneficial; (iii) to improve standard therapeutic regimens and ultimately (iv) to develop criteria for patient selection for future randomized controlled trials.

Conclusion

In summary, patient- and procedure-related risk factors of PAF development following pulmonary resection were identified. The TM&M classification system further quantified the incidence and severity of disease burden from PAF. Severity stratification and accurate prediction of PAF following pulmonary resection may lead to more aggressive prophylaxis of specific populations that may reduce the incidence of PAF.

DISCLOSURES

No relevant financial relationships with commercial interests to disclose.

AUTHOR CONTRIBUTIONS

Study Conception and Design: Seely, Andrew; Ivanovic, Jelena; Ramzan, Sarah

Acquisition of Data: Ramzan, Sarah; Ivanovic, Jelena

Analysis and Interpretation of Data: Ivanovic, Jelena; Seely, Andrew; Maziak, Donna

Drafting of Manuscript: Ivanovic, Jelena; Ramzan, Sarah

Critical Revision: Ivanovic, Jelena; Ramzan, Sarah; McGuire, Anna; Seely, Andrew; Gilbert, Sebastian; Maziak, Donna; Shamji, Farid; Sundaresan, Sudhir

REFERENCES

1. Vaporciyan A, Correa AM, Rice DC, et al. Risk factors associated with atrialfibrillation after noncardiac thoracic surgery: analysis of 2588 patients. J Thorac Cardiovasc Surg 2004;127:779–86.
2. Dyszkiewicz WF, Skrzypczak M. Atrial fibrillation after surgery of the lung: Clinical analysis of risk factors. Eur J Cardiothorac Surg 1998;13:625–28.
3. Gomez-Caro A, Moradiellos MF, Ausin PF, et al. Risk factors for atrial fibrillation after thoracic surgery. Arch Bronconeumol 2006;42:9–13.
4. Amar DF, Zhang HF, Roistacher N. The incidence and outcome of ventricular arrhythmias after noncardiac thoracic surgery. Anesth Analg 2002;95:537–43.
5. Amar D, Roistacher N, Rusch VW, et al. Effects of diltiazem prophylaxis on the incidence and clinical outcome of atrial arrhythmias after thoracic surgery. J Thorac Cardiovasc Surg 2000;120:790–98.
6. Lanza LA, Visbal AI, DeValeria PA, et al. Low-dose oral amiodarone prophylaxis reduces atrialfibrillation after pulmonary resection. Ann Thorac Surg 2003;75:223–30.
7. Tisdale JE, Wroblewski HA, Kesler KA. Prophylaxis of atrial fibrillation after noncardiac thoracic surgery. Semin Thoracic Surg 2010;22:310–20.
8. Onaitis MF, D'Amico TF, Zhao YF, et al. Risk factors for atrialfibrillation after lung cancer surgery: analysis of the society of thoracic surgeons general thoracic surgery database. Ann Thorac Surg 2010;90:368–74.
9. Amar DF, Zhang H, Leung DHY, et al. Older age is the strongest predictor of post-operative atrial fibrillation. Anesthesiology 2002;96:352–56.
10. Polanczyk CA, Goldman L, Marcantonio ER, et al. Supraventricular arrhythmia in patients having noncardiac surgery: Clinical correlates and effect on length of stay. Ann Intern Med 1998;129:279–82.
11. Harpole DH, Liptay MJ, DeCamp MM Jr, et al. Prospective analysis of pneumonectomy: Risk factors for major morbidity and cardiac dysrhythmias. Ann Thorac Surg 1996;61:977–82.
12. Hollings DD, Higgins RSD, Faber LP, et al. Age is a strong risk factor for atrial fibrillation after pulmonary lobectomy. Am J Sur 2010;199:558–61.
13. Passman RS, Gingold DS, Amar D, Lloyd-Jones D, et al. Prediction rule for atrial fibrillation after major noncardiac thoracic surgery. Ann Thorac Surg 2005;79:1698–703.
14. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: A new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205–13.
15. Seely AJ, Ivanovic J, Threader J, et al. Systematic classification of morbidity and mortality after thoracic surgery. Ann Thorac Surg 2010;90:936–42.
16. Roselli EE, Murthy SC, Rice TW, et al. Atrial fibrillation complicating lung cancer resection. J Thorac Cardiovasc Surg 2005;130:438–44.
17. Ivanovic J, Al-Hussaini A, Al-Shehab D, et al. Evaluating the reliability and reproducibility of the Ottawa thoracic morbidity and mortality classification system. Ann Thorac Surg 2011;91:387–93.

18. Ciriaco PF, Mazzone PF, Canneto BF, Zannini P. Supraventricular arrhythmia following lung resection for non-small cell lung cancer and its treatment with amiodarone. Eur J Cardiothorac Surg 2000;18:12–16.
19. Mansour Z, Kochetkova EA, Santelmo N, et al. Risk factors for early mortality and morbidity after pneumonectomy: A reappraisal. Ann Thorac Surg 2009;88:1737–43.
20. Park BJ, Zhang H, Rusch VW, Amar D. Video-assisted thoracic surgery does not reduce the incidence of post-operative atrial fibrillation after pulmonary lobectomy. J Thorac Cardiovasc Surg 2007;133:775–79.
21. Ritchie AJ, Bowe P, Gibbons JR. Prophylactic digitalization for thoracotomy: A reassessment. Ann Thorac Surg 1990;50:86.
22. Amar D, Roistacher N, Burt B, et al. Effects of diltiazem versus digoxin on dysrhythmias and cardiac function after pneumonectomy. Ann Thorac Surg 1997;63:1374–82.
23. Bobbio AF, Caporale DF, Internullo EF, et al. Post-operative outcome of patients undergoing lung resection presenting with new-onset atrialfibrillation managed by amiodarone or diltiazem. Eur J Cardiothorac Surg 2007;31:70–74.
24. Chelazzi C, Villa G, De Gaudio AR. Post-operative atrial fibrillation. ISRN Cardiol 2011; 2011:203179.

Commentary

Accurate and unbiased data collection is vital for clinical research and in studies of surgical outcomes. The American College of Surgeons-National Surgical Quality Improvement Program (ACS-NSQIP) has been considered as the most prominent surgical quality improvement initiative for standardizing and comparing outcomes after surgery. Significant improvements in the delivery of care across different institutions and for different types of operations have been consistently demonstrated after analyzing post-implementation ACS-NSQIP data. The following study explores the inter-system reliability of reported outcomes following thoracic surgery from the ACS-NSQIP database and the TM&M classification system in order to better understand to what extent the methods used to collect data may be impacting study results. To our knowledge, this is the first study to examine the accuracy of adverse event reporting specific to thoracic surgery in comparison to the most prominent surgical quality improvement effort.

Upon careful analysis of these two different systems, our results demonstrate that the method of data collection can significantly alter the reported rate of post-operative adverse events. A thorough evaluation of the methods of data collection, and the goals of each system in the corresponding discussion make it evident as to why these differences in reported rates exist.

Despite specific and unique criteria, and definitions for reporting, the respective objectives of the two systems remain similar: to report and evaluate post-operative adverse events so as to improve the quality of surgical care.

Chapter III

Measuring surgical quality: A comparison of post-operative adverse events with the American College of Surgeons-National Surgical Quality Improvement Program and the Thoracic Morbidity & Mortality Classification System

J Am Coll Surg 2014;218:1024–1031

Presented at the 2013 American College of Surgeons-National Surgical Quality Improvement Program (ACS-NSQIP) Conference, San Diego, CA, July 2013

Jelena Ivanovic, MSc;[1,2,3]
Andrew J.E. Seely, MD, PhD, FRCSC;[1,2,3]
Caitlin Anstee, BA;[2]
P. James Villeneuve, MDCM, PhD, FRCSC;[2]
Sebastien Gilbert, MD, FRCSC;[2]
Donna E. Maziak, MDCM, MSc, FRCSC, FACS;[1,2,3]
Farid M. Shamji, MBBS, FRCSC;[2]
Alan J. Forster, MD, MSc, FRCPC;[1,3]
and R. Sudhir Sundaresan, MD, FRCSC, FACS.[2,3]

Department of Epidemiology and Community Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Canada[1]
Division of Thoracic Surgery, Department of Surgery, The Ottawa Hospital, 501 Smyth Road, Ottawa, Canada[2]
Clinical Epidemiology Program, Ottawa Hospital Research Institute, 501 Smyth Road, Ottawa, Canada[3]

ABSTRACT

BACKGROUND: Monitoring surgical outcomes is critical to quality improvement; however, different data-collection methodologies can provide divergent evaluations of surgical outcomes. We compared post-operative adverse event reporting on the same patients using two classification systems: the retrospectively recorded American College of Surgeons (ACS) NSQIP and the prospectively collected Thoracic Morbidity and Mortality (TM&M) system.

METHODS: Utilizing the TM&M system, complications and deaths were documented daily by fellows and reviewed weekly by staff for all thoracic surgical cases conducted at our institution (4/1/2010-12/31/2011). ACS-NSQIP recording was performed 30-120 days after index surgery by trained surgical clinical reviewers on a systemic sampling of major cases during the same time period. Univariate analyses of the data were performed.

RESULTS: During the study period, 1788 thoracic procedures were performed (1091 were designated "major" as per ACS-NSQIP inclusion criteria). ACS-NSQIP evaluated 182 of these procedures, representing 21.1% and 16.7% of patients and procedures. Mortality rates were: TM&M 1.4% versus ACS-NSQIP 2.2% (p=0.42). Total patients and procedures with complications reported were 24.4% and 31.1% by TM&M versus 20.2% and 39.0% by ACS-NSQIP (p=0.23 and 0.03), respectively. Rates of reported cardiac complications were higher in TM&M vs. ACS-NSQIP (5.8% versus 1.1%; p=0.01); while wound complications were lower (2.5% versus 6.0%; p=0.01).

CONCLUSIONS: Although overall rates were similar, significant differences in collection, definitions, and classification of post-operative adverse events were observed when comparing TM&M and ACS-NSQIP. While both systems offer complementary value, harmonization of definitions and severity classification would further enhance quality improvement programs.

INTRODUCTION

It is increasingly important to measure and evaluate the quality of surgical care, as surgery has become more and more complex. Surgical care is technologically advanced, highly specialized, and involves invasive procedures performed frequently on high-risk and complex patients.[1] Surgical outcomes, particularly post-operative complications, are the most commonly used indicator for surgical quality assessment.[2] As such, data on post-operative complications are often used as a means of comparing surgical techniques, individual surgeon outcomes, and institutional performance.[3] However, different data-collection methodologies can provide divergent evaluations of surgical outcomes.

In 2010, the American College of Surgeons-National Surgical Quality Improvement Program (ACS-NSQIP) was implemented at The Ottawa Hospital (TOH) as a means to evaluate surgical quality, and facilitate quality improvement initiatives through rigorously collected risk-adjusted outcomes. The ACS-NSQIP traditionally has assessed cases from the fields of general and vascular surgery;[1] however, there is now a model that allows inclusion of cases from multiple specialties, including thoracic surgery. The ACS-NSQIP methodology provides estimates of both unadjusted and risk-adjusted post-operative morbidity and mortality (M&M) rates, has been demonstrated to considerably improve surgical outcomes,[4] and is widely considered the most prominent surgical quality improvement effort.

Similarly, Clavien and colleagues were the first to introduce an innovative system to grade post-operative complications by severity proportional to the effort required to treat the complication as a means to facilitate surgical quality improvement.[5] This system, now known as the Clavien-

Dindo classification system, was validated in 2004 in a large cohort of patients who underwent a number of general surgical procedures, and it has since been used in several surgical subspecialties,[6] including thoracic surgery.[7]

The Thoracic Morbidity and Mortality (TM&M) classification system was developed in 2008 by TOH's Division of Thoracic Surgery in accordance to the Clavien-Dindo classification of surgical adverse events (AEs). The TM&M classification system is a prospective system that documents all post-operative AEs and their severity for all thoracic surgical procedures. The TM&M classification system has recently been evaluated for its reliability, reproducibility, and inter-rater agreement.[8] The TM&M classification system facilitates reporting, monitoring, and evaluation of post-operative AEs.[7] However, despite its proven feasibility as an effective method for continuous surgical quality assessment, no studies have been done to compare the TM&M classification system with the most prominent surgical quality improvement effort.

The objectives of this study were two-fold. First, we sought to compare outcomes and the relative effectiveness of post-operative AE reporting on the same patients using the two classification systems: the retrospectively recorded ACS-NSQIP and the prospectively collected TM&M classification system. Second, we performed a descriptive analysis of the context and processes of data collection for the two systems to yield insights into the strengths and weaknesses of each system.

MATERIALS & METHODS

This study was approved by TOH Research Ethics Board. Data used in this study originated from thoracic surgical patients operated at TOH between April 2010 to December 2011, spanning one 18-month period. TOH is a high-volume, single academic thoracic surgery center serving a population of 1.35 million people; thoracic surgical care is consolidated at one campus by six thoracic surgeons. The annual operative volume for thoracic surgery averages approximately 1,200 patients.

The analytic approach and methods of ACS-NSQIP have been described previously.[1,9,10] Briefly, ACS-NSQIP is based on manual and retrospective review of medical records using strict AE definitions. The ACS-NSQIP collects preoperative patient demographics, risk factors, procedure, and 30-day complications relating to an index surgical procedure using a systematic and temporal approach for a typical institution.[1] Specifically, the first 40 successive surgical cases meeting the inclusion criteria are collected during an 8-day cycle.[1] This sampling may or may not result in a 20% sample of eligible cases.

In comparison, the TM&M classification system is a prospective database that provides an accurate summary of the absolute rate of complications and quantifies their severity for all cases. The TM&M system was developed according to the Clavien-Dindo classification schema[6] of surgical AEs. Definitions of surgical AEs were modified according to complications in patients after noncardiac thoracic surgery through peer-review and questionnaire, and adjusted based on surgeons' experience. TM&M data collection and reporting is a continuous, collegial, and divisional activity that is composed of daily reporting (by thoracic surgical fellows), weekly

review (by staff surgeons), monthly discussion at M&M rounds (by the entire Division of Thoracic Surgery), and quarterly analysis (by the entire Division of Thoracic Surgery).

To improve the value and overall quality of TM&M data reporting, a web-based reporting system has recently been developed. Patients with complications are recorded in real-time on a daily basis by thoracic surgical fellows. Post-operative AEs are chosen from a series of standardized definitions (a complete list of definitions is available at: https://ottawatmm.org/). Information on surgical volume, priority of surgery, disease diagnosis, procedure class, and surgical approach/incision is also collected and stored in the TM&M database. The result is a powerful source of information for all thoracic surgical patients that is available for monthly presentation and discussion at M&M rounds, and quality assurance and scientific analyses by a multidisciplinary team of surgeons, clinical epidemiologists, and clinical managers. With respect to the current study, all major TM&M surgical cases (i.e., the same population being selected for ACS-NSQIP review) were selected for the comparison.

Statistical analysis

Univariate analyses of the data were performed using chi-square analysis of contingency tables, with a p value <0.05 considered statistically significant. Reported rates of complications common to both systems were analyzed. A descriptive comparison of the two systems was conducted to identify their strengths and weaknesses. Data were analyzed using SAS software (version 9.2, SAS Institute).

RESULTS

Quantitative comparison

During the study period, 1,788 thoracic procedures were performed, 1,091 of which met ACS-NSQIP inclusion criteria. The ACS-NSQIP evaluated 182 (16.7%) of these procedures and 178 (21.1%) patients.

Table 3.1 illustrates the types of operative procedures performed during the study period. No significant differences were noted in the sampling of major operative procedures performed, including mediastinoscopy/mediastinotomy, bullectomy/pleurectomy, sublobar resection, bilobectomy, pneumonectomy, esophagectomy, gastrectomy, and laparoscopic surgery of the gastroesophageal junction. Significant differences in the sampling scheme were noted in the rate of lobectomy procedures captured by ACS-NSQIP (26.4% vs 19.2%; $p = 0.0247$), and the rate of other major types of operative procedures captured by ACS-NSQIP (16.5% vs 29.7%; $p = 0.002$). Other major surgical procedures included excision/resection of mediastinal tumors, empyema/decortication, and other explorative procedures.

Table 3.1 Number and types of major operative procedures performed, April 2010 – December 2011.

Procedure	TM&M (n = 1091)		ACS-NSQIP (n = 182)		*p* Value
	n	%	n	%	
Mediastinoscopy/mediastinotomy	239	21.9	37	20.3	0.6327
Bullectomy/pleurectomy	38	3.5	10	5.5	0.1872
Sub-lobar resection	127	11.6	30	16.5	0.0659
Lobectomy	209	19.2	48	26.4	0.0247*
Bilobectomy	6	0.5	3	1.7	0.1016
Pneumonectomy	12	1.1	1	0.5	0.4941
Esophagectomy	31	2.8	5	2.7	0.9434
Gastrectomy	29	2.7	4	2.2	0.7175
Laparoscopic surgery of the gastroesophageal junction	85	7.8	14	7.7	0.9633
Other	324	29.7	30	16.5	0.0002*

Total patients and procedures with complications reported were 24.4% and 31.1%, respectively, by TM&M vs. 20.2% and 39.0%, respectively, by ACS-NSQIP (p=0.2299 and p=0.03) (**Table 3.2**). Mortality rate was 1.4% as reported by TM&M vs. 2.2% as reported by ACS-NSQIP (p=0.4214) (**Table 3.2**). Rate of readmission was 1.5% as reported by TM&M vs. 1.7% as reported by ACS-NSQIP (p=0.88) (**Table 3.2**).

Table 3.2 Post-operative morbidity, mortality rates, and re-admission rates.

	TM&M		ACS-NSQIP		*p* Value
	n	%	n	%	
Total patients	843	100	178	21.1	n/a
Total number of operative procedures	1091	100	182	16.7	n/a
Total patients with complications	206	24.4	36	20.2	0.2299
Total number of complications	339	31.1	71	39.0	0.0339*
30-day mortality	12	1.4	4	2.2	0.4214
30-day readmission	13	1.5	3	1.7	0.8888

Table 3.3 illustrates the absolute rates of post-operative occurrences as detected by both systems. Rates of cardiac complications were significantly higher as reported by TM&M vs. ACS-NSQIP (specifically, 5.8% vs. 1.1%; p=0.0080); and rates of wound (6.0% vs. 2.5%; p=0.0088) and other types of complications (12.6% vs. 6.0%; p=0.0013) were significantly higher as reported by ACS-NSQIP. Of the 63 cardiac events captured by the TM&M system, 43 were atrial fibrillation, representing a total of 68% of all cardiac events and a total of 3.9% of all AEs. No significant differences were noted in the rates of reported complications in the remaining groupings of complications.

Table 3.3 Rates of total post-operative occurrences detected by both systems.

Post-operative occurrences by system	TM&M (n =1091)		ACS-NSQIP (n = 182)		*p* Value
	n	%	n	%	
Cardiac^	63	5.8	2	1.1	0.0080*
Central Nervous System	22	2.0	2	1.1	0.3994
Respiratory	134	12.3	24	13.2	0.7319
Urinary	27	2.5	9	4.9	0.0627
Wound	27	2.5	11	6.0	0.0088*
Other	66	6.0	23	12.6	0.0013*
Total	339	31.1	71	39.0	0.0339*

^Of the 63 cardiac events captured by the TM&M system, 43 were atrial fibrillation representing a total of 68% of all cardiac events, and a total of 3.9% of all adverse events.

Table 3.4 illustrates specific post-operative complications detected by both ACS-NSQIP and the TM&M system. Rates of congestive heart failure (2.7% vs. 0.1%; p<0.0001), pulmonary embolism (1.6% vs. 0.3%; p=0.0123), wound dehiscence (1.6% vs. 0.1%; p=0.0005), transfusion (9.3% vs. 0.4%; p<0.0001), and sepsis (1.1% vs. 0.2%; p=0.0410) were significantly higher as reported by ACS-NSQIP. A trend toward significance was observed in rates of surgical site

infections (SSIs) between the two systems, with higher reported rates of SSIs captured by ACS-NSQIP (3.8% vs 2.2%; p=0.1822). No significant differences in reporting rates were detected among the remaining complications.

Table 3.4 Rates of specific post-operative occurrences recorded by both the TM&M system and the ACS-NSQIP.

System	Post-operative occurrence	TM&M (n =1091)	ACS-NSQIP (n = 182)	*p* Value
Cardiac	Congestive heart failure	1 (0.1)	5 (2.7)	<.0001*
	Deep venous thrombosis	1 (0.1)	1 (0.5)	0.1489
	Ischemia (myocardial infarction)	9 (0.8)	2 (1.1)	0.7116
Central Nervous System	Cerebrovascular accident/stroke	4 (0.4)	1 (0.5)	0.7151
Respiratory	Pneumonia	26 (2.4)	8 (4.4)	0.1190
	Pulmonary embolism	3 (0.3)	3 (1.6)	0.0123*
Urinary	Renal insufficiency	8 (0.7)	1 (0.5)	0.7841
	Urinary tract infection	15 (1.4)	6 (3.3)	0.0595
Wound	Wound dehiscence	1 (0.1)	3 (1.6)	0.0005*
	Wound (surgical site) infection	24 (2.2)	7 (3.8)	0.1822
Other	Transfusion intraoperative/post-operative	4 (0.4)	17 (9.3)	<.0001*
	Sepsis	2 (0.2)	2 (1.1)	0.0410*

Descriptive comparison

Table 3.5 illustrates the similarities and differences in definitions of post-operative AEs between the two systems for the seven types of events with statistically significant differences in reported rates, including atrial fibrillation, congestive heart failure, pulmonary embolism, wound dehiscence, SSIs, transfusion, and sepsis.

Table 3.5 Selected post-operative occurrences as defined by the TM&M classification system and the ACS-NSQIP.

System	Post-operative occurrence	TM&M*	ACS-NSQIP**
Cardiac	Congestive heart failure	Cardiac output is insufficient to meet the body's normal requirements for oxygen and nutrients, and pulmonary edema develops	Congestive heart failure is the inability of the heart to pump a sufficient quantity of blood to meet the metabolic needs of the body or can do so only at increased ventricular filling pressure; or ff indication on chest x-ray of pulmonary edema
Respiratory	Pulmonary embolism	Occlusion of one or more pulmonary arteries by thrombi that originate elsewhere	Lodging of a blood clot in a pulmonary artery with subsequent obstruction of blood supply to the lung parenchyma
Wound	Wound dehiscence	Previously closed wound reopening	Post-operative incision dehiscence (superficial or dehiscence to fascia; fascia remains intact)
	Surgical site infection	Purulent wound discharge and/or local host response	Deep incisional surgical site infection, organ space surgical site infection, superficial incisional surgical site infection
Other	Transfusion intraoperative/post-operative	Post-operative hemorrhage or hematoma	Bleeding transfusions ≥ 4 units
	Sepsis	Confirmed or suspected infection in the presence of the systemic inflammatory response syndrome	Sepsis (two clinical signs and symptoms of systemic response to infection) or septic shock (associated with organ and/or circulatory dysfunction)

*TM&M complications are further sub-classified according to the severity and effort required to treat the complication. Specifically, grades I and II complications require no therapy or pharmacologic intervention only. Grades III and IV require surgical intervention or life support. Grade V complications result in patient death.
** The above noted definitions do not provide the detailed specifications of each event that are applied within ACS-NSQIP.

Table 3.6 provides a descriptive comparison of the TM&M classification system and the ACS-NSQIP. Significant differences in data collection and reporting of post-operative AEs were observed when comparing an in-hospital prospectively collected methodology on all patients vs. a retrospective methodology to measure AEs in systematically selected patients.

Table 3.6 Descriptive comparison of the TM&M classification system and the ACS-NSQIP with respect to key quality indicators.

Quality indicator	TM&M	ACS-NSQIP
Data collection	Prospective, in-hospital	Retrospective, 30-day
Data entry	Thoracic surgical fellows (with weekly affirmation by staff surgeons)	Trained & audited surgical case reviewers
Data reporting	Continuous	Bi-annual
Patient population	All thoracic surgical patients	20% sample
Burden of complications	Incidence & severity	Incidence only
Application & relevance	Divisional monitoring	Institutional benchmarking

DISCUSSION

The reporting and evaluation of surgical outcomes is imperative to improving surgical quality.[11] Post-operative complications are important surgical outcomes that impact the health of the patient, as well as, increase hospital costs and length of stay.[12] The ACS-NSQIP is considered the most prominent surgical quality improvement effort, and has been demonstrated to considerably improve surgical M&M.[4] We performed a comparative analysis on the relative effectiveness of post-operative AE reporting using the ACS-NSQIP thoracic surgical patient database and the TM&M classification system from a single institution, TOH, in one 18-month period.

Both ACS-NSQIP and TM&M are robust and designed for their own individual purposes. The ACS-NSQIP identifies predefined post-operative AEs based on documented data in the clinical medical record and patient-reported events during the follow-up period by trained and audited surgical clinical reviewers. The ACS-NSQIP is a validated methodology and is useful for comparing risk-adjusted perioperative surgical outcomes across participating institutions and represents a systems-based approach to surgical quality improvement.[1] Retrospective reviews of patients' medical records have been the foundation of research into errors and AEs,[13] and studies have shown that medical record review is more detailed, robust, and informative than are administrative claims, and has greater validity than voluntary reporting.[14] However, the method is retrospective and can be limited by poor documentation in clinical records.[15] For this purpose, the ACS-NSQIP has implemented training and audit procedures for its hospital participants that are highly effective in collecting robust data. An analysis of inter-rater reliability of variables in the ACS-NSQIP found that the reliability of the data was high from the inception and has improved over time (3.2% disagreement in 2005 vs. 1.6% disagreement in 2008).[16,17] In addition, disagreement levels for individual variables have continually improved, with 26 individual variables demonstrating >5% disagreement in 2005, to only two such variables in 2008.[17]

The TM&M classification system is an in-hospital prospectively collected monitoring system. Our staff and residents are trained to proactively monitor patients for post-operative AEs. Post-operative AEs are chosen from a series of standardized definitions and complications are recorded in real-time on a daily basis by thoracic surgical residents using a web-based AE reporting system. Weekly review by staff surgeons allows for affirmation of complications. Ongoing feedback in the process of quality reporting plays an essential role in maintaining the

accuracy and completeness of data. The TM&M classification system allows users to analyze outcomes in the full census of thoracic surgical patients and data can be subdivided by priority of surgery, disease diagnosis, procedure class, and surgical approach/incision. The system can be used to evaluate severity and burden of post-operative AEs, and represents a continuous and divisional approach to surgical quality assessment. Prospective clinical surveillance has been cited as the most precise and accurate method of reporting AEs and is ideally suited for assessing the effectiveness of specific interventions to decrease explicitly defined AEs. However, prospective clinical surveillance is limited by practical and methodological issues, including the requirement for an observer who clearly understands clinical processes to ensure reliability.[18,19] A previous study done by our group has demonstrated that the TM&M classification systems offers high inter-rater reliability: 87% of kappa statistics were >0.81, a range that is interpreted as "almost perfect agreement;" and the remaining 13% ranged between 0.61 and 0.8, interpreted as "substantial agreement."[8]

Our results show that overall rates of reported M&M were similar between ACS-NSQIP and the TM&M classification system. However, significant differences were observed in the raw incidence of specific post-operative AEs. The differences in incidence reflect different definitions within each system, the difference between prospective and retrospective data collection, and the differing time horizons of the two programs.

First, altered definitions might seem trivial, but impact the data and yield differing results. For example, rates of wound dehiscence were significantly higher as reported by ACS-NSQIP. Wound dehiscence as defined by ACS-NSQIP refers to post-operative incision dehiscence in

which the fascia remains intact; TM&M defines it as a previously closed wound reopening; the term *dehiscence* is reserved for fascial dehiscence. Rate of intraoperative/post-operative transfusion was significantly higher as reported by ACS-NSQIP, again reflecting the differential definitions. The precise definition of post-operative bleeding remains controversial in the surgical community. The ACS-NSQIP defines post-operative bleeding as requiring a transfusion of ≥4 U, which might be due to a variety of causes (e.g., gastrointestinal bleed), or occur over several days.

Despite its many positives, our data also highlighted several important limitations to the ACS-NSQIP. One important drawback is that ACS-NSQIP is not yet comprehensive – some thoracic surgical-specific definitions simply do not appear within ACS-NSQIP, such as post-operative atrial fibrillation. Our results showed that atrial fibrillation presents a considerable post-operative burden in our patient population. Of the 63 cardiac events captured by the TM&M system, 43 were atrial fibrillation, representing a total of 68% of all cardiac events, and a total of 3.9% of all AEs.

Post-operative atrial fibrillation has remained one of the most frequent complications that occur after noncardiac thoracic surgery. Although it is difficult to determine the true incidence of post-operative atrial fibrillation due to various methodologies used to identify its occurrence, reported rates have varied between 4% and 37%.[20-22] The occurrence of post-operative atrial fibrillation is associated with significant morbidity, such as increased risk of stroke, atrial thrombosis and systemic embolization, post-operative mortality, and substantial increases in hospital length of stay and costs.[20,23] We would suggest that additional procedure-specific standards of reporting

within ACS-NSQIP would aid quality improvement programs to alleviate the burden of this costly complication.

Second, the data collected for M&M for ACS-NSQIP extend to 30-days post-operatively, and TM&M captures inpatients during their hospital stay, reflecting the differing time horizons of the two programs. Complications that become evident after patients leave the hospital can be particularity difficult to track and are not recorded by the TM&M system, and this might have resulted in under-reporting of specific complications.

Third, one advantage of a retrospective approach, such as the ACS-NSQIP, is its ability to capture events post-discharge. For example, a trend toward significance was observed in differing rates of SSIs between the two systems, with higher reported rates of SSIs captured by ACS-NSQIP, reflecting the longer follow-up periods. Surgical site infections can be acquired after hospital discharge (e.g., in follow-up clinic visits, with visits to the general practitioner, or emergency room visits) and recorded in ACS-NSQIP, but will not be picked up by TM&M. Weigelt and colleagues have demonstrated a 53% increase in infection rates reported by complete 30-day inpatient and outpatient reporting.[24]

Fourth, the impact of serial or cascading complications was addressed differently by the two systems. The ACS-NSQIP assesses and records all post-operative AEs in a patient, even if they are serial. As suggested by Clavien and colleagues, our goal is to record only the most severe complication pertaining to the affected system when those complications of a lower grade are a step in the process leading to the more serious outcomes.[7,25] For example, aspiration leading to

respiratory failure is recorded as a single grade IVa using the TM&M system; ACS-NSQIP will record aspiration and respiratory failure as two separate complications. However, no significant differences were noted in the rates of the total number of patients with complications between the two systems.

There is complementary value to the institutionally-focused ACS-NSQIP compared with the divisionally-focused TM&M. Importantly, ACS-NSQIP uses state-of-the-art validated risk-adjustment methodology to address the confounding effect of case mix on the frequency and nature of surgical complications,[26-28] which provides useful information for benchmarking and comparisons across institutions.[29,30] However, it is less applicable as a continuous quality improvement measure for an individual thoracic surgical program, as understanding and improving the delivery of a particular operation might require measures tailored to that operation.[2] As ACS-NSQIP assesses a systematic sample of cases, the sample size is not large enough for subgroup analyses, such as results from specific procedures or individual surgeon performance. Similarly, semi-annual reports allow institution-specific comparisons that form the basis for development of institution-specific quality improvement action plans; however, at the expense of timely identification of problems within a single surgical service. Ultimately, prompt identification and recognition of a problem would mean a more rapid response to rectify it.[31]

Conversely, the TM&M classification system does not yet have a model for risk-adjustment of outcomes. Yet, the absence of adjustment for illness should not limit the use of TM&M data for quality assessment because we are not attempting to measure differences between individual hospitals, but rather are monitoring outcomes within one surgical service. A recent study by

Salatia and colleagues[32] has demonstrated the usefulness of the TM&M classification system in auditing the quality of care within a single surgical unit. The authors concluded that the TM&M classification system revealed a decline in quality of care within their unit otherwise undetected by applying traditional outcomes measures, and that the system can be used as an additional graded outcomes end point to refine internal audit of performance.[32] Taken together, risk-adjusted performance feedback enabled by ACS-NSQIP, coupled with a continuous and prospective data-collection methodology such as the TM&M system, is fundamental for monitoring surgical outcomes, and for tailored quality-improvement efforts. Our institution and division have invested in both systems.

Our results have demonstrated that neither system is more or less effective than the other at driving quality improvement; however, efforts to harmonize definitions of AEs need to be undertaken. Based on our experience with the Clavien-Dindo−inspired TM&M system, we suggest that ACS-NSQIP consider objectively characterizing the severity of complications in addition to documenting incidence. Quantification of severity of post-operative AEs is possible using ACS-NSQIP, and can be useful in assessing surgical outcomes.[33,34] In our division, severity grading has helped to assess overall complication burden, in comparison with considering only events. This approach underscores the substantial impact of higher-grade or major complications. Although grade I and II complications represent the majority of complications in our patient population, they contribute the least burden to hospital resources. On the contrary, grade III to V complications comprise a minor portion of complications, but the majority of the burden.

There are a number of limitations in this study. First, the data are derived from a single institution, limiting generalization of our results. Second, a direct comparison of the two systems is difficult due to differences in recording methodologies and differences in definitions of post-operative AEs between the two systems. As discussed here, these inherent differences might have contributed to our findings. Third, although there is a difference in target patient populations between the two systems (i.e., TM&M contains data on all thoracic surgical patients, and ACS-NSQIP targets a specific sample of the inpatient and outpatient settings), we strived to ensure the two patient populations were analogous.

Conclusion

Both ACS-NSQIP and TM&M systems have strengths and limitations, and offer complementary value. Harmonization of definitions, including the addition of definitions relevant to the thoracic surgical subspecialty, along with a severity classification of post-operative complications, would enhance quality improvement programs.

DISCLOSURES

No relevant financial relationships with commercial interests to disclose.

AUTHOR CONTRIBUTIONS

Study Conception and Design: Seely, Andrew; Ivanovic, Jelena; Sundaresan, Sudhir

Acquisition of Data: Ivanovic, Jelena; Anstee, Caitlin; Forster, Alan

Analysis and Interpretation of Data: Ivanovic, Jelena; Seely, Andrew

Drafting of Manuscript: Ivanovic, Jelena

Critical Revision: Ivanovic, Jelena; Seely, Andrew; Gilbert, Sebastian; Maziak, Donna; Shamji, Farid; Sundaresan, Sudhir

REFERENCES

1. Khuri SF, Daley J, Henderson W, et al. The Department of Veterans Affairs' NSQIP: The first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. Ann Surg 1998;228:491−507.
2. Birkmeyer JD, Dimick JB, Birkmeyer NJO. Measuring the quality of surgical care: Structure, process, or outcomes? J Am Coll Surg 2004;198:626−32.
3. Veen EJ, Steenbruggen J, Roukema JA. Classifying surgical complications: A critical appraisal. Arch Surg 2005;140:1078−83.
4. Hall BL, Hamilton BH, Richards K, et al. Does surgical quality improve in the American College of Surgeons National Surgical Quality Improvement Program: An evaluation of all participating hospitals. Ann Surg 2009;250:363−76.
5. Clavien PA, Sanabria JR, Strasberg SM. Proposed classification of complications of surgery with examples of utility in cholecystectomy. Surgery 1992;111:518−26.
6. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: A new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205−13.
7. Seely AJE, Ivanovic J, Threader J, et al. Systematic classification of morbidity and mortality after thoracic surgery. Ann Thor Surg 2010;90:936−42.
8. Ivanovic J, Threader J, Alhussaini A, et al. Testing the reliability and reproducibility of the Ottawa Hospital Thoracic Morbidity and Mortality classification system. Ann Thor Surg 2011;91:387−93.
9. Khuri SF, Daley J, Henderson WG. The comparative assessment and improvement of quality surgical care in the Department of Veterans Affairs. Arch Surg 2002;137:20−27.
10. Fink AS, Campbell DA Jr, Mentzer RM Jr, et al. The National Surgical Quality Improvement Program in non-veterans administration hospitals: Initial demonstration of feasibility. Ann Surg 2002;236:344−53.
11. Davenport DL, Henderson WG, Khuri SF, Mentzer RM Jr. Preoperative risk factors and surgical complexity are more predictive of costs than post-operative complications: A case study using the National Surgical Quality Improvement Program (NSQIP) database. Ann Surg 2005;242:463−68.
12. Dimick JB, Chen SL, Taheri PA, et al. Hospital costs associated with surgical complications: A report from the private sector national surgical quality improvement program. J Am Coll Surg 2004;199:531−37.
13. Luck J, Peabody JW, Dresselhaus TR, et al. How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. Am J Med 2000;108:642−49.
14. Madsen JJ. Comparison of concurrent and retrospective methods of detecting adverse drug reactions. Am J Hosp Pharm 1993;50:2556−57.
15. Forster AJ, Dervin G, Martin C, et al. Improving patient safety through the systematic evaluation of patient outcomes. Can J Surg 2012;55:418−25.
16. Ingraham AM, Richards KE, Hall BL, Ko CY. Quality improvement in surgery: The American College of Surgeons National Surgical Quality Improvement Program approach. Adv Surg 2010;44:251−67.

17. Shiloach M, Frencher SK, Steeger JE, et al. toward robust information: data quality and inter-rater reliability in the American College of Surgeons National Surgical Quality Improvement Program. J Am Coll Surg 2010;210:6−16.
18. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. J Gen Intern Med 2003;18:61−67.
19. Etchells E, O'Neill C, Bernstein M. Patient safety in surgery: Error detection and prevention. World J Surg 2003;27:936−42.
20. Vaporciyan A, Correa AM, Rice DC, et al. Risk factors associated with atrial fibrillation after noncardiac thoracic surgery: analysis of 2588 patients. J Thorac Cardiovasc Surg 2004;127: 779−86.
21. Dyszkiewicz WF, Skrzypczak M. Atrial fibrillation after surgery of the lung: Clinical analysis of risk factors. Eur J Cardiothorac Surg 1998;13:625−28.
22. Gomez-Caro A, Moradiellos MF, Ausin PF, et al. Risk factors for atrial fibrillation after thoracic surgery. Arch Bronconeumol 2006;42:9−13.
23. Amar DF, Zhang HF, Roistacher N. The incidence and outcome of ventricular arrhythmias after noncardiac thoracic surgery. Anesth Analg 2002;95:537−43.
24. Weigelt JA, Dryer D, Haley RW. The necessity and efficiency of wound surveillance after discharge. Arch Surg 1992;127:77−81.
25. Clavien PA, Barkun J, de Oliveira ML, et al. The Clavien-Dindo classification of surgical complications: Five-year experience. Ann Surg 2009;250:187−96.
26. Daley J, Khuri SF, Henderson W, et al. Risk adjustment of the post-operative morbidity rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:328−40.
27. Henderson WG, Daley J. Design and statistical methodology of the National Surgical Quality Improvement Program: why is it what it is? Am J Surg 2009;198[Suppl]:S19−S27.
28. Davis CL, Pierce JR, Henderson W, et al. Assessment of the reliability of data collected for the Department of Veterans Affairs national surgical quality improvement program. J Am Coll Surg 2007;204:550−60.
29. David RS, Worthington JR, Kitts JB. Implementation of an integrated peri-operative quality management program at The Ottawa Hospital. Healthc Manage Forum 2011;24:S34−S40.
30. Velanovich V, Rubinfeld I, Patton JH, et al. Implementation of the National Surgical Quality Improvement Program: Critical steps to success for surgeons and hospitals. Am J Med Qual 2009;24:474−79.
31. Forster AJ, Worthington JR, Hawken S, et al. Using prospective clinical surveillance to identify adverse events in hospital. BMJ Qual Saf 2011;20:756−63.
32. Salatia M, Pompili C, Refai M, et al. The use of the Thoracic Morbidity and Mortality system for the internal analysis of performance: a case-matched temporal audit. Eur J Cardiothorac Surg 2014;45:859−63.
33. Strasberg SM, Hall BL. Post-operative Morbidity Index: a quantitative measure of severity of post-operative complications. J Am Coll Surg 2011;213:616−26.
34. Porembka MR, Hall BL, Hirbe M, Strasberg SM. Quantitative weighting of post-operative complications based on the accordion severity grading system: Demonstration of potential impact using the American College of Surgeons National Surgical Quality Improvement Program. J Am Coll Surg 2010;210:286−98.

Commentary

Surgeon-specific reporting of outcomes of care is gaining increasing attention, as hospitals strive to increase the transparency and accountability of surgical care. There are significant benefits to this approach for both hospitals and surgeons, including the opportunity to drive quality improvement initiatives. Several concerns do persist, however, regarding the use of surgeon-specific outcome reports, including resistance to change, and the many ways information may be misinterpreted and misused. In the following manuscript, we review the benefits and limitations surrounding surgeon-specific outcomes reporting, and then suggest some possible solutions to these concerns in a stride toward better implementation and utilization of surgeon-specific reporting to improve surgical outcomes.

However, surgeon-specific outcome reporting is only half the battle. Data collection needs to be complemented with innovative means to deliver data back to surgeons, such that it improves care. Positive deviance is one such approach; and its flexibility, relevance, and use in identifying practical and sustainable solutions in improving outcomes following surgical care is further explored.

Chapter IV

Using surgeon-specific outcome reports and positive deviance for continuous quality improvement in thoracic surgery

Ann Thorac Surg 2015; In press

Presented at the 51st Society of Thoracic Surgeons Annual Meeting, San Diego, California, January 2015

Jelena Ivanovic, MSc;[1,2,3]
Caitlin Anstee, BA;[2]
Tim Ramsay, PhD;[1,2]
Sebastien Gilbert, MD, FRCSC;[2]
Donna E. Maziak, MDCM, MSc, FRCSC, FACS;[1,2,3]
Farid M. Shamji, MBBS, FRCSC, FACS;[2]
R. Sudhir Sundaresan, MD, FRCSC, FACS;[2,3]
P. James Villeneuve, MDCM, PhD, FRCSC;[2]
and Andrew J.E. Seely, MD, PhD, FRCSC.[1,2,3]

Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada[1]
Division of Thoracic Surgery, Department of Surgery, The Ottawa Hospital, Ottawa, Canada[2]
Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada[3]

ABSTRACT

BACKGROUND: Using the Thoracic Morbidity & Mortality classification to document all post-operative adverse events between 10/12-02/14, we created surgeon-specific outcome reports (SSORs) to promote self-assessment and to implement a divisional continuous quality improvement (CQI) program to improve individual surgeon's clinical performance, on the construct of positive deviance (PD).

METHODS: Mixed-methods study within a division of six thoracic surgeons, involving: i) iterative development of real-time/web-based/risk-adjusted SSORs; ii) implementation of CQI seminars (n=6; 09/13-06/14) for evaluation of results, collegial discussion on quality improvement based on identification of positive outliers, and selection of quality indicators for future discussion; and iii) in-person interviews to identify facilitators/barriers to using SSORs and CQI. Interview transcripts were analyzed using thematic analysis.

RESULTS: Interviews revealed enthusiastic support for SSORs as a means to improve patient care through awareness of personal outcomes with blinded divisional comparison for similar operations and diseases, and apply the learning objectives to continuous professional development and maintenance of certification. Perceived limitations of SSORs included: difficulty measuring surgeon expertise, limited understanding of risk-adjustment, resistance to change, and belief that knowledge of sensitive data could lead to punitive actions. All surgeons believed CQI/PD seminars led to collegial interactions and discussions, while perceived limitations included quorum participation, and failing to circle back on actionable items.

CONCLUSIONS: Real-time performance feedback using SSORs can motivate surgeons to improve their practice, while CQI/PD seminars offer the opportunity to review and interpret results, and address issues in a supportive environment. Whether SSORs and CQI/PD can lead to improvements in rates of post-operative adverse events is a matter of ongoing research.

INTRODUCTION

The rate of post-operative adverse events (AEs) is often used as a parameter for evaluating both the effectiveness of treatment and the quality and safety of surgical care. The selection of patients for operation, patient factors, disease factors, and surgical expertise are all important variable considerations in AEs. These are serious considerations now and in the future as the occurrence of post-operative AEs have been directly linked to mortality,[1] hospital length of stay,[2] and post-operative quality of life.[3] There is also increasing evidence demonstrating that post-operative AEs affect the overall costs and resource utilization in major surgery.[4,5] Post-operative AEs are a major influence on both clinical and economic outcomes of surgical care, and methodologies to better categorize, report, and monitor their incidence are essential for ongoing efforts to minimize their occurrence and impact.[6] It is the responsibility of the surgeon to be diligent in assessing, reporting, and improving the quality of surgical care delivered at all times.[7]

A number of strategies have been advocated to promote improvement in the quality of surgical care, including: performance measurement and feedback, positive deviance (PD), and dissemination of best practice measures. Specifically, performance measurement and feedback are increasingly being used as a strategy to provide surgeons with benchmarking information to use for individual and institutional quality improvement.[8] Performance measurement and feedback is formally defined as the summation of clinical performance of care over a specified time period aimed at providing information to health professionals to allow them to assess and adjust their performance.[9] Performance measurement and feedback are intended to increase the accountability of healthcare professionals, enhance clinical performance, and thereby improve the quality of care and patient safety.[10] However, lack of transparent, explicit, systematic, data-

driven performance measurement and feedback mechanisms for surgeons has been considered to be a major impediment in fully adopting this type of quality improvement strategy.

The concept of PD originated in international public health initiatives and was based on the observation that in most communities, including quality of healthcare research, there were individuals and groups whose uncommon practices produce better outcomes than their peers.[11] Specifically, the PD approach identifies innovative strategies from "positive deviants" in healthcare, those individuals that consistently demonstrate exceptionally high performance in an area of interest, and designs and implements activities enabling others to exercise these new practices.[12] The PD approach has recently been used to improve quality and safety of healthcare delivery in a number of settings.[13-15] To date, no studies have used the approach of PD as means to promote improvement in the quality of surgical care.

Using the validated Thoracic Morbidity & Mortality (TM&M) classification of AEs,[16-17] the objectives of this study were three-fold. First, to create risk-adjusted surgeon-specific outcome reports (SSORs) to enable individualized performance measurement and feedback. Second, to implement a divisionally-focused, surgeon-led, continuous quality improvement (CQI) program to review results, select quality indicators in need of improvement, discuss quality improvement strategies based on identification of positive deviants along with best practice measures, and select topics for subsequent discussion. Third, to understand surgeons' perceptions, including the benefits and limitations, on the use of SSORs and a CQI/PD program as a means to enable surgeons to actively participate in assessment of their clinical performance.

MATERIALS & METHODS

Study Design

We performed a mixed-methods study within a division made up of six thoracic surgeons (Division of Thoracic Surgery, The Ottawa Hospital, Ottawa, Canada) involving: i) development of real-time, web-based, risk-adjusted SSORs; ii) implementation of CQI seminars (n=6; 09/13-06/14) to review results, select quality indicators, discuss quality improvement strategies based on identification of positive outliers along with best practice measures, and select topics for future discussion; and iii) in-person confidential interviews to identify facilitators and barriers of using SSORs and CQI/PD. The study was approved by The Ottawa Hospital Research Ethics Board.

Thoracic Morbidity & Mortality classification system of post-operative adverse events

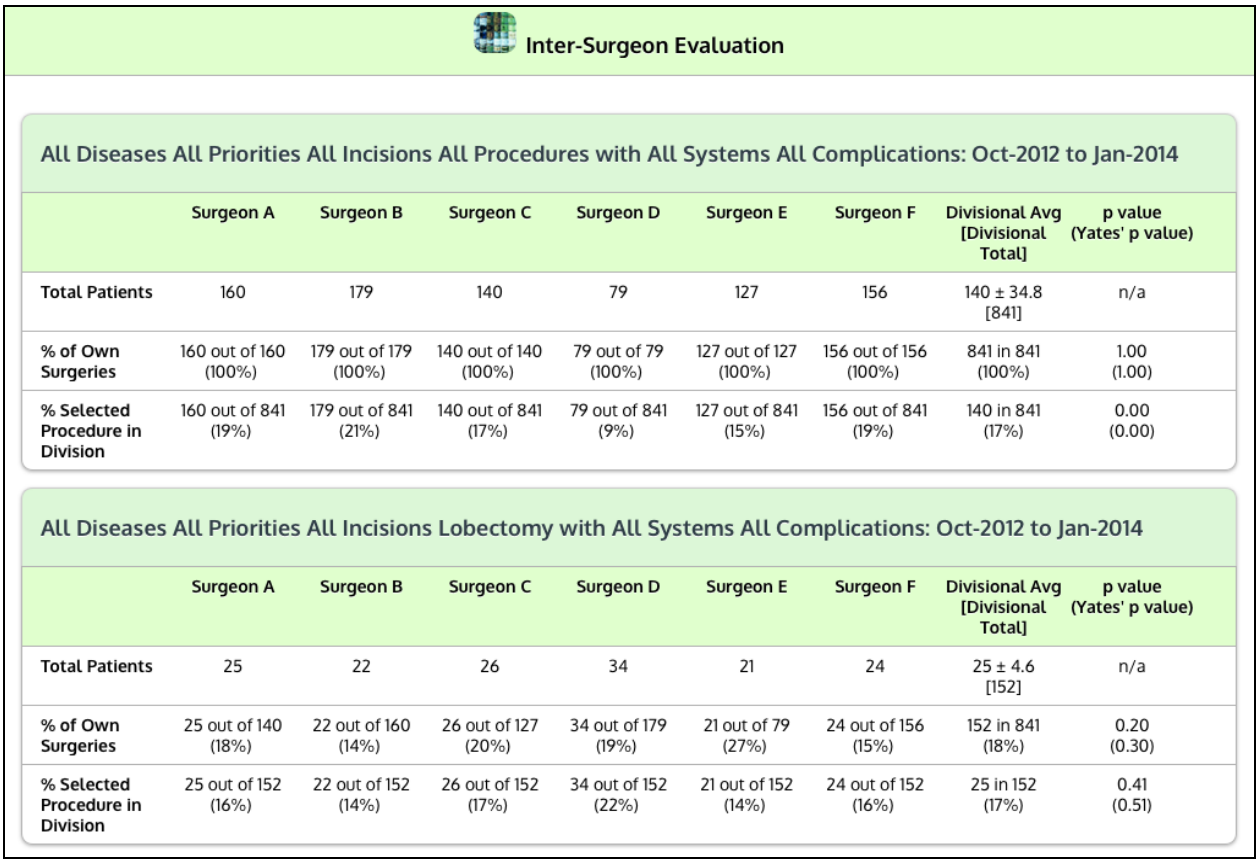
The TM&M classification system is an in-hospital prospective database that provides a reliable summary of the rate of post-operative AEs and quantifies their severity. The TM&M system was developed according to the Clavien-Dindo classification schema of surgical AEs.[18] TM&M data collection and reporting is a continuous, collegial, and divisional activity that is composed of daily reporting (by thoracic surgical fellows), weekly review (by staff surgeons), and monthly discussion and analysis at morbidity & mortality rounds (by the entire Division of Thoracic Surgery). The process of TM&M data collection and reporting has been facilitated by a web-based, menu-driven, iPad-optimized software application. Post-operative AEs are chosen from a series of standardized definitions (a complete list of definitions is available at: http://ottawatmm.org). Information on surgical volume, priority of surgery, disease diagnosis, procedure class, and surgical approach/incision are also collected and stored in the TM&M

database. Additional quality metrics including if a complication resulted in: prolonged length of hospital stay, hospital readmission, and return to emergency are also reported and monitored.

Creation of web-based volume, complication, and surgeon-specific outcomes reports (SSORs)

The web-based software application was derived from the TM&M classification of AEs based on data from thoracic surgical patients operated at The Ottawa Hospital between October 1, 2012 to February 28, 2014, spanning a 16-month time period. The software application is comprised of three reports, including: a divisional volume report, a divisional outcomes report, and a SSOR, all created to be dynamic, interactive, and anonymous (visit http://tsqic.org).

i. Volume report: This report shows the number of surgeries performed by each surgeon since October 2012. Surgeons can refine or filter the results by selecting a specific procedure, the surgical approach/incision, and/or a specific time period (**Figure 4.1**).



Inter-Surgeon Evaluation

All Diseases All Priorities All Incisions All Procedures with All Systems All Complications: Oct-2012 to Jan-2014

	Surgeon A	Surgeon B	Surgeon C	Surgeon D	Surgeon E	Surgeon F	Divisional Avg [Divisional Total]	p value (Yates' p value)
Total Patients	160	179	140	79	127	156	140 ± 34.8 [841]	n/a
% of Own Surgeries	160 out of 160 (100%)	179 out of 179 (100%)	140 out of 140 (100%)	79 out of 79 (100%)	127 out of 127 (100%)	156 out of 156 (100%)	841 in 841 (100%)	1.00 (1.00)
% Selected Procedure in Division	160 out of 841 (19%)	179 out of 841 (21%)	140 out of 841 (17%)	79 out of 841 (9%)	127 out of 841 (15%)	156 out of 841 (19%)	140 in 841 (17%)	0.00 (0.00)

All Diseases All Priorities All Incisions Lobectomy with All Systems All Complications: Oct-2012 to Jan-2014

	Surgeon A	Surgeon B	Surgeon C	Surgeon D	Surgeon E	Surgeon F	Divisional Avg [Divisional Total]	p value (Yates' p value)
Total Patients	25	22	26	34	21	24	25 ± 4.6 [152]	n/a
% of Own Surgeries	25 out of 140 (18%)	22 out of 160 (14%)	26 out of 127 (20%)	34 out of 179 (19%)	21 out of 79 (27%)	24 out of 156 (15%)	152 in 841 (18%)	0.20 (0.30)
% Selected Procedure in Division	25 out of 152 (16%)	22 out of 152 (14%)	26 out of 152 (17%)	34 out of 152 (22%)	21 out of 152 (14%)	24 out of 152 (16%)	25 in 152 (17%)	0.41 (0.51)

Figure 4. 1 Surgeon-specific volumes of all procedures (upper panel) and lobectomies (lower panel) performed since October 2012.

ii. Complication report: This report shows the number of surgeries performed by each surgeon since October 2012, in addition to the number of reported post-operative complications. Similar to the volume report, surgeons can refine or filter the results by selecting a specific procedure, a specific time period, the organ system affected, or the post-operative complication(s) of interest (**Figure 4.2**).

All Diseases All Priorities All Incisions Lobectomy with All Systems Prolonged Air Leak: Oct-2012 to Jan-2014

	Surgeon A	Surgeon B	Surgeon C	Surgeon D	Surgeon E	Surgeon F	Divisional Avg [Divisional Total]	p value (Yates' p value)
Pts w Complications	23%	27%	25%	12%	43%	12%	6 in 25 (23%) [34 in 152 (22%)]	0.10 (0.24)
Pts w Minor Complications	23%	27%	25%	12%	38%	12%	6 in 25 (24%) [33 in 152 (22%)]	0.20 (0.43)
Pts w Major Complications	0%	0%	4%	0%	5%	0%	0 in 25 (0%) [2 in 152 (1%)]	0.43 (0.99)
Pts w Grade V Complications	0%	0%	0%	0%	0%	0%	0 in 25 (0%) [0 in 152 (0%)]	1.00 (1.00)

All Diseases All Priorities All Incisions Lobectomy with All Systems Prolonged Air Leak: Oct-2012 to Jan-2014

NSQIP	Surgeon A	Surgeon B	Surgeon C	Surgeon D	Surgeon E	Surgeon F	Divisional Avg	
Maker-Wasserman (Occurence) 95% CI	0.06 0.04 - 0.08	0.07 0.04 - 0.10	0.06 0.04 - 0.08	0.03 0.02 - 0.04	0.11 0.06 - 0.16	0.03 0.02 - 0.04	0.06	
Maker-Wasserman (Severity) 95% CI	0.06 0.04 - 0.08	0.07 0.04 - 0.10	0.1 0.06 - 0.14	0.03 0.02 - 0.04	0.13 0.07 - 0.19	0.04 0.03 - 0.05	0.07	
EVAD (Age, DLCO, FEV1)	Surgeon A	Surgeon B	Surgeon C	Surgeon D	Surgeon E	Surgeon F	Divisional Avg	p value (Yates' p value)
O/E 95% CI	0.41 0.38 - 1.07	0.45 0.37 - 1.10	0.32 0.94 - 1.39	0.22 0.29 - 1.01	0.69 0.64 - 1.33	0.19 0.43 - 1.01	0.38	0.00 (0.00)
Dynamic Risk Picker (Age,Sex,ASAClass) Age, Sex, A... 3	Surgeon A	Surgeon B	Surgeon C	Surgeon D	Surgeon E	Surgeon F	Divisional Avg	p value (Yates' p value)
Complications - O/E 95% CI	0.5 0.49 - 1.35	0.55 0.46 - 1.36	0.44 1.32 - 1.94	0.26 0.35 - 1.21	0.9 0.84 - 1.76	0.24 0.53 - 1.23	0.48	0.00 (0.00)

Figure 4.2 Surgeon-specific rates of raw and risk-adjusted air leaks are shown for the period between October 2012 and January 2014.

Throughout both the volume report and the complication report, a Chi-square test was added with the purpose of identifying significant inter-surgeon differences in volume of procedures performed, where a *p* value of <0.05 is considered statistically significant. The Yates' correction, a conservative estimate, was also added in conjunction with the Chi-square test to prevent overestimation of statistical significance for small data. The Yates' correction is used when at least one cell of the table has a count < 5 (**Figure 4.1 and Figure 4.2**).

Within the complication report, three different risk-adjustment scores were applied to account for differences in the case-mix for each surgeon for those patients undergoing lung resection, including: limited resection (wedge resection, segmentectomy) standard lobectomy, extended pulmonary resection (sleeve lobectomy, bilobectomy, enbloc chest wall resection), standard pneumonectomy, and sleeve pneumonectomy (**Figure 4.2**).

The first score, the Maker-Wasserman score, is a validated score used by American College of Surgeons-National Surgical Quality Improvement Program (ACS-NSQIP) to compare surgeon-specific data and identify positive outliers, where the lowest score indicates better outcome based on the calculation of the risk index/outcome index.[19] The selection of risk factors for inclusion in the score was based on validated factors used in the universal ACS-NSQIP risk calculator.[20]

The second score, the expiratory volume, age, and diffusing capacity (EVAD score), was selected based on a systematic literature review of thoracic-surgery specific risk scores. The EVAD score was selected because it was demonstrated as easier to use and at least as accurate as other scoring systems used in the thoracic surgical setting, including the Physiological and

Operative Severity Score for Enumeration of Mortality and Morbidity and the Cardiopulmonary Risk Index.[21] The EVAD score uses three variables to predict the risk of post-operative complications following major lung resection, including a patient's age, %FEV_1 and %DLCO.

The third score, the Dynamic Risk Picker, was internally-developed and based on a study by Dimick and colleagues where the authors demonstrated that **procedure-specific** quality measures can be adequately risk-adjusted with a limited number of variables.[22] Within this score, the surgeon can select up to five risk factors appropriate for the procedure to be performed. As with the Maker-Wasserman score, the surgeon selects risk factors for inclusion in the model based on those factors used in the universal ACS-NSQIP risk calculator. The use of the previously identified risk factors for post-operative prolonged alveolar air leak and atrial fibrillation by our group were considered for risk adjustment in this score as well.

For each of the scores, missing data were imputed by averaging the non-missing values. For the EVAD score and the Dynamic Risk Picker, established statistical methods were used to develop a logistic regression model. Once the logistic regression equation is computed, the equation is then used to calculate a probability of the AE for each patient. These probabilities are then summed for each individual surgeon to obtain the expected (E) number of AEs for the patient sample for that individual surgeon. An O/E ratio and confidence interval (CI) is then calculated for the patient sample for the individual surgeon, where O is the number of patients observed to have the AE. If the surgeon's O/E ratio is below 1 and the upper limit of the confidence interval is <1, then the surgeon has a statistically significant smaller number of AEs than would be expected on the basis of his/her patient characteristics = positive outlier (or positive deviant).

iii. SSORs: These reports are available for each surgeon to view their own surgeon-specific data and benchmark themselves to the divisional average. These reports allow surgeons to perform additional in-depth analysis of their own performance data.

Positive deviance (PD) in healthcare and the implementation of a surgeon-led, continuous quality improvement (CQI) program

A divisional, surgeon-led, continuous quality improvement (CQI) program (n=6 seminars; 09/13-06/14) was implemented to serve as a complement to the SSORs and to give surgeons an opportunity to first review group results in a blinded-fashion, select outcomes in need of improvement, discuss best practice strategies based on the consented de-identification of positive deviants/outliers, and finally, select topics for subsequent discussion. Seminars were scheduled on OR-free days to ensure that all or the majority of surgeons could attend. Seminars were generally 1.5 to 2 hours in length.

An expert in the field of PD, was invited to lead the introductory seminar and highlighted the approaches' flexibility, its relevance to a range of quality improvement issues, and its use in identifying practical and sustainable solutions in healthcare.[23] Subsequent seminars were limited to staff surgeons, a database manager (CA), and a research associate (JI).

Semi-structured in-person interviews

Following completion of the 6th CQI/PD seminar, private, semi-structured in-person interviews of a target length of 30 minutes or less were conducted by two of the studies investigators (JI and CA). These investigators were in the best position to conduct the interviews to increase response rates, maintain motivation with longer questions, probe for responses, clarify ambiguous

questions, and aid in the recall of events. Interviews were audio-recorded and then transcribed, with identifying information removed and random numbers assigned to serve as unique identifiers of transcripts. Interview questions were open-ended and designed to assess adoption, usefulness of SSORs, suggested improvements, and the surgeons' overall impression of the CQI/PD program.[24] Interview transcripts were systematically analysed into codes, and clustered to form data themes. The transcripts were treated as narratives and as perceptions of the interviewees.[25]

RESULTS

Patient demographics and volume of procedures performed

During the study period, 258 patients underwent 279 major lung operations by six thoracic surgeons. The number of cases ranged from 79 to 179 per surgeon for all operative procedures performed, and between 32 (11%) to 61 (22%) cases of major lung resections, highlighting significant inter-surgeon differences in the volume of major lung resections (p=0.03) (data not shown). Demographics and risk factors used for inclusion in the risk models for the 258 patients are shown in **Table 4.1**.

Table 4.1 Demographics and pre-operative risk factors for major lung resection surgeries.

Risk Factor	Procedure					
	Wedge Resection (n=80), %	**Segmentectomy (n=15), %**	**Lobectomy (n=144), %**	**Extended Lobectomy (n=7), %**	**Pneumonectomy (n=12), %**	**Total (n=258), %**
Gender						
Female	37 (46.3)	10 (66.7)	86 (59.7)	3 (42.9)	3 (25)	139 (53.9)
Male	43 (53.7)	5 (33.3)	58 (40.3)	4 (57.1)	9 (75)	119 (46.1)
Age						
<49	7 (8.8)	0 (0)	5 (3.5)	1 (14.3)	1 (8.3)	14 (5.4)
50-59	15 (18.8)	2 (13.3)	31 (21.5)	0 (0)	3 (25)	51 (19.8)
60-69	36 (45)	8 (53.3)	46 (31.9)	4 (57.1)	5 (41.7)	99 (38.4)
70-79	18 (22.5)	5 (33.3)	48 (33.3)	2 (28.6)	3 (25)	76 (29.5)
>80	4 (5)	0 (0)	14 (9.7)	0 (0)	0 (0)	18 (7.0)
BMI						
<18.5	3 (3.8)	0 (0)	5 (3.5)	0 (0)	0 (0)	8 (3.1)
18.5-25	22 (27.5)	5 (33.3)	47 (32.6)	6 (14.3)	5 (41.7)	85 (32.9)
25-30	30 (37.5)	7 (46.7)	52 (36.1)	1 (85.7)	6 (50)	96 (37.2)
>30	22 (27.5)	3 (20)	39 (27.1)	0 (0)	1 (8.3)	65 (25.2)
ASA Class						
II	3 (3.7)	0 (0)	15 (10.4)	1 (14.3)	0 (0)	19 (7.4)
III	63 (78.8)	15 (100)	122 (84.7)	6 (85.7)	12 (100)	218 (84.5)
IV	14 (17.5)	0 (0)	7 (4.9)	0 (0)	0 (0)	21 (8.1)
Smoking Status						
Never	28 (35)	6 (40)	36 (25)	3 (42.9)	2 (16.7)	75 (29.1)
Past	38 (47.5)	9 (60)	82 (56.9)	4 (57.1)	7 (58.3)	140 (54.3)
Current	14 (17.5)	0 (0)	26 (18.1)	0 (0)	3 (25)	43(16.7)
Other Pre-Operative Risk Factors						
Steroid Use	13 (16.3)	0 (0)	28 (19.4)	0 (0)	1 (8.3)	42 (16.3)
Disseminated Cancer	31 (38.8)	4 (26.7)	8 (5.6)	1 (14.3)	5 (41.7)	49 (19.0)
Diabetic	15 (18.8)	6 (40)	30 (20.8)	0 (0)	0 (0)	51 (19.8)
Hypertension	34 (42.5)	9 (60)	80 (55.6)	4 (57.1)	2 (16.7)	129 (50)
Previous Cardiac Event	9 (11.3)	3 (20)	27 (18.8)	2 (28.6)	2 (16.7)	43 (16.7)
Dyspnea	22 (27.5)	1 (6.7)	37 (25.7)	3 (42.3)	1 (8.3)	64 (24.8)
COPD	21 (26.3)	5 (33.3)	45 (31.3)	1 (14.3)	4 (33.3)	76 (29.5)
Pulmonary Function Test Results						
Average DLCO (%)	70.4	75.2	73.3	65.3	72.1	72.3
Average FEV1 (%)	78.3	87.1	78.6	76.3	84.4	792
Average FEV1/FVC (%/%)	70.7	70.6	67.5	67	69.9	68.8

Attendance rate for CQI seminars and topics of discussion

Six seminars were held over a 10 month period, with attendance rate varying from 50-100%. Topics of discussion per seminar were largely based on primary outcomes of interest, which included management options for reducing rates of post-operative atrial fibrillation, prolonged air leak, and anastomotic leak.

Semi-structured, in-person interviews

All surgeons agreed to participate in semi-structured, in-person confidential interviews, and consented to having their responses audio-recorded and transcribed for research purposes only. It was mandated that confidentiality was to be observed by the interviewers. Experience level of the surgeons ranged from 2 to 32 years.

SSORs – benefits, limitations, and suggestions for improvement (Table 4.2)

Confidential individual interviews revealed that the majority of surgeons believed SSORs can lead to improvements in care through knowledge gained of personal outcomes with comparison within the division for similar operations, and eventual balancing of outcomes. One surgeon described:

> *"It's a spur to do better, and if you see that you are not performing up to the same average, that's something to say, well, we can do better, we can strive for improvement. The other thing is – it can give you a sense of – well, we're actually on the right track, we're doing very well, our morbidity is low, our mortality is low. So, it can be, sort of an affirmation that practice is being administered in the right fashion."* (**Theme**: Self-assessment; self-improvement)

Other reported benefits included the format through which the data were presented, real-time data access and performance monitoring, and team-building or capacity-building through collegial discussions, as described below:

> *"I like the dynamic way to evaluate the data, i.e., the ability to change procedures, dates, types of complications, all of that. It's like a microscope that I can use to evaluate my complications, and I can just adjust the filters."* (**Theme**: Format of data presentation)
>
> *"I think it is very positive for surgeons to have a constant, nice up to date tracking of their outcomes. You can go over your results at home, go over the results in the office, at any point, day or night, whenever you feel like it - you can go over things, see what your outcomes are, your complications, and get a nice snapshot of how your performance is - realizing that a lot of complications arise as a result of patients' health status, co-morbidities, things like that."* (**Theme**: Real-time data access and performance monitoring)
>
> *"First of all, it's an excellent system, excellent initiative, what I like about it, is that we control the initiative, we decide what is appropriate reporting for us, what is appropriate risk-adjustment. That's all getting ahead of the curve, rather than having someone else from the outside telling us, this is what you're going to do, and that's the way it's going to be."* (**Theme**: Team-building; capacity-building)

One surgeon believed that SSORs can be used for continuous professional development and maintenance of certification:

> *"We can provide letters back to clinicians based on the amount of time that they have been using the software saying, ok, we note that you have used this software for two hours in the last month, please note that that is eligible for Royal College CPD, which is continuous professional development credits."* (**Theme**: Maintenance of certification; continuous professional development)

Perceived *limitations* of SSORs included: potential for inaccurate data, false representation of performance, limited understanding of risk-adjustment metrics, increase risk for liability,

discouragement of surgeons, failing to circle back on actionable items, avoidance of high-risk patients, and resource requirements for local database maintenance, as outlined below:

> *"I was actually looking at reports around esophagectomy patients last week and to me it seemed that the rate of readmission was a little bit high and some of the complications were probably, I don't know, if some of the complications were like actual true complications reported in the system." (**Theme**: Potential for inaccurate data)*
>
> *"We do about 30 [esophagectomies] a year on our service and if there's six surgeons, and if they were equally distributed, we each do about 5 a year. And the problem with that is, that's not statistically [significant]. All one person needs is one leak or one chylothorax to all of a sudden jump to a 20% risk of that complication over the year?"* (**Theme**: False representation of performance)
>
> *"When we looked at the risk-adjustment scores - I forgot how they work. So, I looked at the data, I said, ok, does this mean that it's good or bad? Does that mean I have sicker patients or more healthy patients? I wasn't entirely certain…"* (**Theme**: Limited understanding of risk-adjustment metrics)
>
> *"One thing that is vital to these open discussions, is that there has to be medico-legal protection and the proper mechanisms put in place. I might see myself having to sort of deal with that kind of information and that's not the purpose of these surgeon-specific reports."* (**Theme**: Increase risk for liability)
>
> *"It's sensitive data, and so if it is used in the wrong way, if you will, if it is used for punitive type of actions, then, I fear, that could jeopardize the whole culture of quality improvement and safety." "If you are looking at assessing outcomes, certainly, if the news is not great, one may think about, perhaps self-doubt, with regards to outcomes."* (**Theme**: Discouragement of surgeons)
>
> *"Collecting data without an action to follow is useless. Ok? You are collecting this thing for a reason, for a purpose, for improvement, if those are the three things that you want then you better act on it for the reason that you collected the data in the first place."* (**Theme**: Failing to circle back on actionable items)
>
> *"If surgeons know their own results are going to be scrutinized in detail and commented upon at this level, why do they want to appear as outliers? So, they'll just stop taking those patients."*

"People will just actually gravitate towards taking on more straightforward cases and avoid the tougher ones." (**Theme**: Avoidance of high-risk patients)

"Could this system live on, let's say [the database manager] moves on to somewhere else…are we going to be able to keep running this or not? What is currently being done to make that happen?" (**Theme**: Resource requirements for local database maintenance)

Most surgeons suggested that descriptions of risk-adjustment scores would be helpful to improve their understanding of the reports. Surgeons expressed that this additional information would be valuable if available as an optional tab within the report. Most of the surgeons also reported a need to have access to the entire medical record of patients reported in their data.

Table 4.2 Perceived benefits and limitations of surgeon-specific outcome reports and continuous quality improvement seminars, and suggestions for improvement.

SSORs	Benefits	Self-assessment; self-improvement Real-time data access and performance monitoring Team-building; capacity building Collegial discussion Balancing of outcomes Format Maintenance of certification; continuous professional development Closing the loop
	Limitations	Potential for inaccurate data False representation of performance Limited understanding of risk-adjustment scores Increase risk for liability Discouragement of surgeons Failing to circle back on actionable items Avoidance of high-risk patients Resource requirements for local database maintenance
	Suggestions for improvement	Supplementary information
CQI Seminars	Benefits	Team-building; capacity building Change management Stimulates further personal study Collegial discussion
	Limitations	Quorum participation Scheduling; accessibility Resistance to change Failing to circle back on actionable items
	Suggestions for improvement	Quorum participation Scheduling; accessibility Supplementary information Closing the loop

CQI/PD Seminars – Benefits, Limitations, and Suggestions for Improvement (Table 4.2)

Most surgeons believed a program of CQI/PD has stimulated further personal study on topics and encouraged change management based on best practice measures:

> *"It made me think a little bit and be more thoughtful about what I'm planning to do, what I am doing. Well, most people are doing this [procedure] this way, and we'll try it this way. So that's where the thoughtfulness comes in to say, well, there's a different way of tackling this problem, maybe I'll do it this way and we'll see what happens. So it's kind of, still I guess in evolution. But that's kind of the thoughtfulness that it gives you anyways."* (**Theme**: Stimulates further personal study)
>
> *"They have altered the way that I perform the anastomosis during esophageal reconstruction. They have altered the way I divide the fissure during lobectomies and they have changed the way I think about electrolyte replacement after thoracic surgery."* (**Theme**: Change management)

Similar to the reported benefits of SSORs, all surgeons believed that a program of CQI/PD has led to team-building or capacity-building through collegial interactions and discussions, as elucidated below:

> *"I think that it's been an important force for bringing our division together during the last year. Its been a source of collegial interaction that reinforces the positive relationships that are inherent within our division." "The fact that we could persist, knowing the disparate nature of the people in the group, and a lot of the trepidations expressed at the beginning, the fact that we can pull people together and not only that, but persist with these sessions and make progress with it and come out with even recommendations regarding the topics - I think that's a huge success."* (**Theme**: Team-building; capacity building, and collegial discussion)

Perceived *limitations* of a CQI/PD program included resistance to change, where several surgeons suggested that adopting new approaches and management techniques can pose a serious

challenge. Other limitations included scheduling and accessibility of CQI/PD seminars, and quorum participation.

> *"Everyone is tied to the way they practice in a certain way, to a certain degree. Everyone has their own, for the lack of a better word, dogma of practice, things that they have done, things that they taught, because they think it's the optimal approach, and it's hard to deviate from that." "Not everyone practices surgery the same way. So, as an example, most of us do thoracoscopy, and one of us does not. And so many of the solutions that are potentially implementable, if they are specific to one type of operation, technique, versus another, then it's going to be limited in its applicability." "You see five fingers are different. Five surgeons are going to be different and that's the thing. People are individuals, and they have their prejudices, people think only they know the best, that's a human characteristic." "I think, just resistance to change, surgeons are a bit notorious for latching onto something in the course of their training, or early on in their career, that works for them. Getting them to adopt new approaches and things can be a serious challenge."* (**Theme**: Resistance to change)
>
> *"The challenge I think will remain scheduling and getting people to be there because there's always something going on." "This only works when six surgeons are talking around the table. And scheduling has been the Achilles heel of implementing this in a meaningful way." "The only one [challenge] for me is that you keep holding them at a time when it is really difficult for me to attend personally." "The biggest one is getting all six thoracic surgeons together for the meetings. It has been profoundly valuable to have all surgeons present. It's something that I have really found - its been a learning point - is how important to have 100% participation is to the process." "I think they're good. As long as you get a quorum." (**Theme**: Scheduling; accessibility, quorum participation)*

To improve the conduct of CQI/PD seminars, all surgeons suggested that seminars should be held quarterly to facilitate participation by all staff. Several surgeons suggested seminars should be supplemented with supporting literature of evidence-based best measures. Closing the loop was reported as another suggestion for improving the conduct of CQI/PD seminars, where future

discussions would circle back to previous recommendations as a means to evaluate what impact changes have had on patient outcomes, as described below:

> *"The only suggestion I have is that, you keep repeating it, do you know what I mean? Like, if you take for example, air leak, we discussed it, we did the cautery, so we need to make sure like in 6 months or 8 months to revisit to see if there's been a change and a change not only as a whole, but within each surgeon, and to see if they have applied that."* (**Theme**: Closing the loop)

DISCUSSION

Various methodologies are becoming increasingly utilized to enhance the quality of surgical care. Both SSORs and confidential, peer discussions represent distinct, yet, complementary approaches to quality assessment and improvement. An absence of literature regarding the reliability of SSORs, together with on-going peer discussions, on their impact on the quality of surgical care prompted us create SSORs to enable self-assessment, and implement a complementary CQI program, on the construct of PD, as a means to assess and improve clinical performance. Structured interviews were then conducted to assess limitations and benefits of SSORs, and the surgeons' overall impression of the CQI/PD program.

Perceived limitations of SSORs and complementary CQI/PD seminars

The potential for inaccurate data and the small sample size of AEs leading to false representation of performance were identified as impediments to SSOR use and adoption. Small numbers mean that a large amount of cases need to be aggregated to reach a reliable number of events to obtain meaningful statistical comparisons. This may be especially applicable to surgeons at the beginning and end of their career where case numbers may be lower, and particularly, in the case of junior surgeons, whose techniques may be changing rapidly. As SSORs are continued over

time, case numbers for surgeons with initially small sample sizes will eventually accumulate and provide stable estimates of their performance, and for a single moderate size division as a whole.

Limited understanding of risk-adjustment metrics was identified as another barrier to SSOR use and adoption. Ultimately, the entire division has to agree upon which risk adjustment score best fits the purpose for driving improvement.[26] The ideal risk adjustment score should be simple, easily reproducible, objective, applicable to all patients and operations, both sensitive and specific,[27] and be used in conjunction with the surgeon's own intuition to answer the following questions: *Does the patient need the proposed operation? Is the patient fit for the planned operation? What is the margin between life and death for the patient if an operation is to be undertaken? Are we taking into consideration intraoperative mishaps?* Studies demonstrate, and support interview findings, that if individual reports are not adequately adjusted to account for variations in the risk of a poor outcome, surgeons may avoid caring for high-risk patients.[28]

Surgeons also feared that sensitive data contained within SSORs could be used for punitive actions, in turn, jeopardizing the culture of patient safety and quality improvement. Surgeons thus believed that proper mechanisms should be in place in this environment of open discussions about AEs. Surgeons further believed that comparison of outcomes should not be made in a coercive or punitive fashion, but in a cooperative and constructive spirit, with the ultimate aim to generalize successful best practices across the division. As long as respective medical protective associations and colleges, give support to the concept, and the hospital administration is at arm's length without risk of withdrawal of privileges to work, the increase risk for liability,

discouragement of surgeons, and avoidance of high-risk patients should not be barriers to adopting SSORs.

Scheduling and accessibility, quorum participation, and resistance to change were identified as limitations of complementary CQI/PD seminars. First, studies have shown that educational-type seminars alone or combined with other interventions, can improve professional practice and patient outcomes. Suggested strategies to increase attendance at educational seminars include using mixed interactive and didactic formats, and focusing on outcomes that are likely to be perceived as serious may increase the effectiveness of educational seminars.[29] Topics of discussion at CQI/PD seminars were thus largely based on management options for reducing rates of post-operative atrial fibrillation, prolonged air leak, and anastomotic leak, which pose the highest burden on our patient population. Second, resistance to change by surgeons is an established finding in the published literature. This finding relates to the introduction of evaluation and healthcare IT,[30] and also evidence-based changes to clinical practice, for example, management of hospital-acquired infection.[31] It is contended that resistance to change is a wider issue within healthcare that requires communication, participation in decision making, support and negotiation.[32]

Perceived benefits of SSORs and complementary CQI/PD seminars

Interview transcripts revealed that the majority of surgeons believed SSORs can lead to improvements in care through self-assessment and knowledge of personal outcomes with divisional comparison. A study of primary care physicians showed that they considered report cards either an important or very important part of their practice and supported its continuation;

over half of the participants said the reports had influenced them to make positive changes in their practice.[33] It is also commonly believed that by simply inspecting their own results, surgeons would make their practices better, a concept commonly known as the Hawthorne effect.[34] Similarly, medical specialties employ numerous methods to examine physician performance, including clinical dashboards, evaluation of encounters with simulated patients, observation of patient care, medical record audits, and peer assessment.[35] SSORs, which provide feedback that is both continuous and transparent, can serve as an additional strategy for a comprehensive clinical performance evaluation.

Other reported benefits of SSORs included the format through which the data were presented allowing for real-time data access and performance monitoring. A key attribute of any quality reporting system is in its ability to collect and analyze routinely collected clinical data in order to quickly generate quality reports, and to monitor the quality of the care provided.[36] When initially creating the SSORs, a major objective of the division was to ensure that metrics were consistent with thoracic surgical practice, centered on patient care, and easy to extract and interpret, so as to help guide current and future goals in quality improvement. To minimize the effect of bias in our measurements, the division ultimately decided to focus on the measurement and extraction of readily available electronic TM&M data, thereby enabling the creation of an automatic, reliable, cost-effective process applied on a continuous basis and in real-time. To maintain the benefits of on-going automated data gathering, however, involves substantial information technology expertise and local resource requirements, and ultimately depends on the buy-in, size, and intricacy of the surgical division.

Interview transcripts revealed that the majority of surgeons believed a program of CQI/PD has stimulated further personal study or continuing education on best practice measures. Continuing education has been cited as one way to respond to the results of performance evaluation if an area of practice is found not to meet expectations.[37] A recent study suggests that physicians and surgeons benefit from reflection on their progress and development.[38] Given its inherent self-assessment nature, SSORs are also eligible for maintenance of certification credits in surgery in support of lifelong learning needs.

Interview transcripts also revealed that all surgeons believed that a program of CQI/PD, in conjunction with the use of SSORs, has led to team-building or capacity-building through collegial interactions and discussions. Discussions were marked by camaraderie and respect, and were conducted in a manner that maintained confidentiality. We believe that discussing performance data in a non-threatening way, colleague to colleague, is an effective method of bringing about change. Supporting evidence suggests that many surgeons seek information from colleagues over other sources, highlighting the important role of interaction, collaboration, and professionalism for quality improvement purposes.[39] Gagliardi and colleagues,[40] found that sharing of clinical experience made possible collective decision making for complex cancer cases. Surgeons thought that collegial interaction improved awareness of current evidence appropriate care delivery, and continuity. In addition, by comparing proposed treatment with that of the group and gaining exposure to decision making for more cases than they would see in their own practices, surgeons developed clinical expertise that could be applied to future cases.[40]

Study limitations

A number of limitations are noted in this work. First, this is a single center study, and is susceptible to both sampling bias and response bias due to the small sample size of our thoracic surgery division. As such, it is possible that theoretical saturation[41] (the same themes repeatedly seen) was not reached from the six interview transcripts that were analyzed.

Another concern is that of collation of individual surgeon data to create SSORs. Data from a 16-month period do not provide statistically meaningful results because of small sample sizes and low adverse event rates; risk-adjusted surgical outcomes, thus, often do not meet reliability benchmarks for distinguishing individual surgeon performance. The PD approach, which served as a backbone to CQI seminars, was intended to be a springboard for collegial discussion and interaction, and not a true depiction of exemplary individual outcomes.

Other quality indicators, including patient experience, which may include patient feedback on communication skills and professionalism, long-term outcome measures, and wait time measurements, are potentially valuable but are subject to limitations. All, however, are needed for improving the quality of care, and will be included in future iterations of SSORs.

Last, recording of TM&M relies on consistent vigilance and team-involvement and may at times be susceptive to reporting bias; this vigilance in reporting also applies for risk factors (the preoperative variables associated with post-operative outcomes which were used in the model building for risk-adjustment). On-going feedback in the process of quality reporting, thus, plays an essential role to maintaining the accuracy and completeness of data.

Conclusion

We created innovative and dynamic thoracic SSORs for reporting information on multiple quality indicators. By monitoring individual outcomes and providing feedback, SSORs allow surgeons to evaluate their performance over time and in comparison to colleagues, provide automated and real-time data monitoring, and are a fundamental component of contemporary efforts to improve the quality of surgical care. Unique to our study, is an effort to link performance results with on-going CQI seminars based on the concept of PD in an attempt to provide an additional forum for discussion and reinforcement of trust in ones' colleagues. Whether SSORs and a complementary program of CQI/PD can lead to improvements in rates of post-operative AEs is a matter of ongoing research.

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the mentorship of Curt Lindberg, a leading expert in the field of PD, for his encouragement of this project, and his input and assistance in leading the introductory CQI seminar.

DISCLOSURES

No relevant financial relationships with commercial interests to disclose.

AUTHOR CONTRIBUTIONS

Study Conception and Design: Seely, Andrew; Ivanovic, Jelena

Acquisition of Data: Ivanovic, Jelena; Anstee, Caitlin

Analysis and Interpretation of Data: Ivanovic, Jelena

Drafting of Manuscript: Ivanovic, Jelena

Critical Revision: Ivanovic, Jelena; Seely, Andrew; Ramsay, Tim; Gilbert, Sebastian; Maziak, Donna; Shamji, Farid; Sundaresan, Sudhir

REFERENCES

1. Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. JAMA 2003;290:1868−74.
2. Khan NA, Quan H, Bugar JM, et al. Association of post-operative complications with hospital costs and length of stay in a tertiary care center. J Gen Intern Med 2006;21:177−180.
3. Derogar M, Orsini N, Sadr-Azodi O, et al. Influence of major post-operative complications on health related quality of life among long-term survivors of esophageal cancer surgery. J Clin Oncol 2012; 30:1615−19.
4. Vonlanthen R, Slankamenac K, Breitenstein S, et al. The impact of complications on costs of major surgical procedures: A cost analysis of 1200 patients. Ann Surg 2011;254:907−13.
5. Dimick JB, Pronovost PJ, Cowan JA, et al. Complications and costs after high-risk surgery: Where should we focus quality improvement initiatives? J Am Coll Surg 2003;196:671−78.
6. Martin RC, Brennan MF II, Jaques DP. Quality of complication reporting in the surgical literature. Ann Surg 2002;235:803−13.
7. Hall BL, Campbell DA Jr, Phillips LRS, Hamilton BH. Evaluating individual surgeons based on total hospital costs: Evidence for variation in both total costs and volatility of costs. J Am Coll Surg 2006;202:565−76.
8. Dimick JB, Hendren SK. Hospital report cards: Necessary but not sufficient? JAMA Surg 2014;149:143−44.
9. O'Brien MA, Oxman AD, Davis DA, et al. Audit and feedback versus alternative strategies: Effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2000;2:CD000260.
10. Shekelle PG, Lim YW, Mattke S, et al. Does public release of performance results improve quality of care? A systematic review. Health Foundation, 2008;1−45.
11. Wishik SM, Van der Vynckt S. The use of nutritional "positive deviants" to identify approaches for modification of dietary practices. Am J Public Health 1976;66:38−42.
12. Bradely EH, Curry LA, Ramanadhan S, et al. Research in action: Using positive deviance to improve quality of health care. Implement Sci 2009;4:25.
13. Awad SS, Palacio CH, Subramanian A, et al. Implementation of a methicillin-resistant Staphylococcus aureus (MRSA) prevention bundle results in decreased MRSA surgical site infections. Am J Surg 2009;198:607−10.
14. Marra AR, Guastelli LR, Pereira de Araújo CM, et al. Positive deviance: A new strategy for improving hand hygiene compliance. Infect Control Hosp Epidemiol 2010;31:12−20.
15. Jain R, Kralovic SM, Evans ME, et al. Veterans Affairs initiative to prevent methicillin-resistant Staphylococcus aureus infections. N Engl J Med 2011;364:1419−30.
16. Seely AJE, Ivanovic J, Threader J, et al. Systematic classification of morbidity and mortality after thoracic surgery. Ann Thor Surg 2010;90:936−42.
17. Ivanovic J, Threader J, Alhussaini A, et al. Testing the reliability and reproducibility of the Ottawa Hospital Thoracic Morbidity and Mortality classification system. Ann Thor Surg 2011;91:387−93.
18. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: A new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205–13.

19. Maker V. Single-score ACS-NSQIP report cards: Key to improving your institution. Presented at The Ottawa Hospital, Ottawa, Canada. November 21, 2013.
20. Bilimoria KY, Liu Y, Paruch JL, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: A decision aid and informed consent tool for patients and surgeons. J Am Coll Surg 2013;217:833−42.
21. Ferguson MK, Durkin AE. A comparison of three scoring systems for predicting complications after major lung resection. Euro J Cardiothorac Surg 2003;23:35−42.
22. Dimick JB, Osborne NH, Hall BL, et al. Risk adjustment for comparing hospital quality with surgery: How many variables are needed? J Am Coll Surg 2010;210:503−08.
23. Lindberg C, Clancy TR. Positive deviance: An elegant solution to a complex problem. J Nurs Admin 2010;40:150−53.
24. Yi SG, Wray NP, Jones SL et al. Surgeon-specific performance reports in general surgery: An observational study of initial implementation and adoption. J Am Coll Surg 2013;217:636–647 .
25. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psy 2006;3:77−101.
26. Opelka FG. Performance measurement and optimal care for surgical patients. Ann Surg 2014;259:850−51.
27. Shah N, Hamilton M. Clinical review: Can we predict which patients are at risk of complications following surgery? Crit Care 2013;17:226.
28. Schneider EC, Epstein AM. Influence of cardiac-surgery performance reports on referral practices and access to care: A survey of cardiovascular specialists. N Engl J Med 1996;335:251−56.
29. Forsetlund L, Bjorndal A, Rashidian A, et al. Continuing education meetings and workshops: Effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2009;2:CD003030.
30. Boonstra A, Broekhuis M. Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. BMC Health Serv Res 2010;10:231.
31. Saint S, Kowalski CP, Banaszak-Holl J, et al. How active resisters and organizational constipators affect health care-acquired infection prevention efforts. Jt Comm J Qual Patient Saf 2009;35:239−46.
32. Lamb MC. Implementing change in the National Health Service. J Manag Med 1999;13:288–97.
33. Custers T, Hurley J, Klazinga NS, et al. Selecting effective incentive structures in health care: A decision framework to support health care purchasers in finding the right incentives to drive performance. BMC Health Serv Res 2008;8:66.
34. Traverso LW. Surgical outcomes: What are they and why should they be measured? Surg Endosc 2000;14:103−05.
35. Ehrenfeld JM, Henneman JP, Peterfreund RA, et al. Ongoing professional performance evaluation (OPPE) using automatically captured electronic anesthesia data. Jt Comm J Qual Patient Saf 2012;38:73−80.
36. Capurro D, Yetisgen M, van Eaton E, et al. Availability of structured and unstructured clinical data for comparative effectiveness research and quality improvement: A multi-site assessment. eGEMS 2014;2:1−7.

37. Mazmanian PE & Davis DA. Continuing medical education and the physician as a learner: Guide to the evidence. JAMA 2002;288:1057−60.
38. Campbell CM, Parboosingh J, Gondocz T, et al. Study of the factors influencing the stimulus to learning recorded by physicians keeping a learning portfolio. J Cont Educ Health Prof 1999;19:16−24.
39. Violato C, Lockyer J, Fidler H. Multisource feedback: A method of assessing surgical practice. BMJ 2003;326:546−48.
40. Gagliardi AR, Wright FC, Anderson MA, et al. The role of collegial interaction in continuing professional development. J Cont Edu Health Prof 2007;27:214−19.
41. Bowen GA. Naturalistic inquiry and the saturation concept: A research note. Qual Res 2008;8:137−52.

Commentary

Quality of care and patient safety have been receiving sustained attention since the publication of 1999 Institute of Medicine Report, *To Err is Human*,[1] and interest in these areas are likely to increase even further. An evolutionary understanding of the heterogeneity of Thoracic Morbidity & Mortality (TM&M) is a critical first step to following improvements in care, and to developing novel and successful approaches to quality and patient safety research.

However, measurement of evolutionary TM&M data alone, although necessary, it is not sufficient for quality improvement. The following study summarises our division's progress to date, including both successes and challenges with the use of the TM&M classification system, and lists our recommendations for the next steps that will shape the future of patient safety and quality improvement in surgery. We suggest that in addition to implementing a complementary point-of-care, interactive, web-based quality monitoring system, key factors for improving quality and patient safety include a combination of temporal analyses of adverse events, effective surgeon-specific feedback mechanisms, actionable information based on best practice measures, and standardization of case reviews.

Our vision is that the rigorous analytical components of standardized adverse event monitoring and reporting will be transposed into other surgical disciplines allowing more accurate tracking, and evaluation of post-operative adverse events.

REFERENCES

1. Kohn LT, Corrigan J, Donaldson MS. To Err is Human: Building a Safer Health System. Washington, DC: National Academy Press 2000.

Chapter V

Analysing impact of a prospective, standardized, and comprehensive post-operative adverse event monitoring and reporting system as a means for quality improvement

To be submitted for publication consideration

Presented at the 19th International Forum for Quality and Safety, Paris, France, April 2014

Jelena Ivanovic, MSc;[1,2,3]
Caitlin Anstee, BA;[3]
Christian Finley, MD, MPH, FRCSC;[4]
Sebastien Gilbert, MD, FRCSC;[2,3]
Donna E. Maziak, MDCM, MSc, FRCSC, FACS;[1,2,3]
Farid M. Shamji, MBBS, FRCSC;[3]
R. Sudhir Sundaresan, MD, FRCSC, FACS;[1,3]
P. James Villeneuve, MDCM, PhD, FRCSC;[3]
Tim Ramsay, PhD;[1,2]
and Andrew J.E. Seely, MD, PhD, FRCSC.[1,2,3]

Clinical Epidemiology Program, Ottawa Hospital Research Institute, 501 Smyth Road, Ottawa, Canada[1]
Department of Epidemiology and Community Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Canada[2]
Division of Thoracic Surgery, Department of Surgery, The Ottawa Hospital, 501 Smyth Road, Ottawa, Canada[3]
Division of Thoracic Surgery, Department of Surgery, St. Joseph's Healthcare, 2757 King Street East, Hamilton, Canada[4]

ABSTRACT

BACKGROUND: Objective reporting of post-operative adverse events (AEs) is the basis of surgical quality assurance. In 2008, we developed a standardized classification to identify presence and severity of Thoracic Morbidity & Mortality (TM&M). However, paper-based forms that were originally used to support AE reporting were inefficient, with potential for data entry error. Our objectives were to create an interactive, portable method to record and report AEs, integrated with process-of-care delivery, the Thoracic Surgery Quality Monitoring, Information Management, and Clinical Documentation (TSQIC) system; and to use the system as a foundation for evaluating temporal trends in volume and TM&M rates based on evolutionary data.

METHODS: The TM&M classification was developed in accordance to the Clavien-Dindo classification of AEs. Daily recording of all thoracic operative procedures and all AEs is performed by residents and staff surgeons, along with weekly review by all staff. Monthly data review, discussions, and iterative software refinements led to an evolution of the web-based TSQIC system allowing for real-time reporting, monitoring, and analyses of volume and quality.

RESULTS: From 01/2008-10/2014, 3384 patients (mean age 61, range 14-97) underwent 4293 thoracic surgical procedures, of which 1317 (30.9%) patients had at least one complication. Major complication rates (grades III-IV) for lobectomy, pneunomenctomy, and esophagectomy were 11.6% (n=104/898), 30.1% (n=22/73), and 48.1% (n=78/162); mortality rates were 0.7% (n=6/898), 4.1% (n=3/73), and 4.3% (n=7/162). 17.0% (n=728/4293) of AEs led to prolonged length of stay, 7.1% (n=306/4293) led to hospital readmission/return to emergency, and 2.5% (n=108/4293) of AEs were unrelated to index procedure. We have incorporated automated reporting of AEs into M&M rounds, procedure-specific and risk-adjusted surgeon-specific outcome reports (SSORs), and positive deviance (PD) seminars. Analysis of post-implementation of SSORs and CQI/PD seminars (01/2012 to 11/2013 vs. 12/2013 to 02/2015) suggests a non-significant decrease in overall rates of atrial fibrillation following pulmonary resection (8.7% vs 7.0%; $p=0.42$), a significant decrease in overall rates of prolonged alveolar air leak following pulmonary resection (21.6% vs. 7.8%; $p<0.001$), and a significant decrease in grade II rates of anastomotic leak following esophagectomy (14.3% vs. 0%; $p<0.04$).

CONCLUSIONS: Prospective and standardized definitions facilitate reporting of all post-operative AEs, while a web-based, point-of-care system provides an effective method for AE data entry and review. Together, they have enabled powerful team-oriented quality improvement opportunities resulting in a positive cultural change.

INTRODUCTION

Quality and patient safety are of foremost importance to surgeons. Assessing post-operative adverse events (AEs) is essential for evaluating both the effectiveness of treatment and the quality and safety of surgical care. A systematic review of eight chart review studies (from United States, Canada, United Kingdom, Australia, and New Zealand) found a median overall incidence of AEs of 9.2% (of which approximately 40% were surgical in nature).[1] Post-operative AEs are costly,[2] cause pain and suffering,[3] extend hospital length of stay,[4] and are the most important determinant of decreased post-operative survival.[5]

The importance of rigorous recording of clearly defined AEs is widely recognized, yet is rarely and poorly performed in practice. Martin and colleagues[6] conducted a systematic review on the quality of AE reporting in the surgical literature, and found that the lack of uniformity in documenting and reporting AEs results in incomplete capture of data, making comparisons among different surgical approaches and institutional series difficult. The authors strongly argue for the creation and generalized use of standards for reporting this information in order to assess the impact of therapeutic changes on patient outcomes, and to follow improvements in quality and safety over time.[6]

In 1992, Clavien and colleagues were the first to introduce the concept of severity grading of post-operative AEs. The principle underlying this classification is that the severity of an AE is proportional to its impact on a patient and the degree of effort to rectify it.[7] The rationale for this approach was to eliminate subjective interpretation of serious AEs and any tendency to down-grade complications, as it is based on data that are well-documented and easily verified. The

system, now known as the Clavien-Dindo classification of AEs, was validated in 2004 in a large cohort of general surgical patients, and has since been used in a number of different surgical sub-specialities.[8]

In 2008, modeled on the Clavien-Dindo system, The Ottawa Hospital's Division of Thoracic Surgery developed a standardized approach to identify both presence and severity of Thoracic Morbidity & Mortality (TM&M), with the aim to capture all AEs after all surgeries.[9] The rigorous implementation in clinical practice and evaluation of such a system requires time. Since its inception, supporting studies have shown that the TM&M classification is feasible,[9] reliable and reproducible,[10] facilitates objective comparison, identifies burden of illness of individual complications,[11,12] provides an effective method for continuous surgical quality assessment, and serves as a complement to the most prominent surgical quality improvement effort, the National Surgical Quality Improvement Program (NSQIP).[13] More recently, the TM&M classification has been adopted by numerous surgical groups internationally.[14,15] We are also aware of randomized trials, which have included this classification in their end point evaluation.[16]

Every post-operative AE has been recorded in a prospective quality database since the beginning of its application; however, paper-based forms that were originally used to support TM&M reporting were cumbersome and inefficient, and limited in their use for benchmarking purposes. Modern information technology (IT) represents an especially valuable tool to overcome the barriers of traditional AE reporting methods, most notably, by facilitating a more rapid response after an AE has occurred, and by monitoring and providing feedback about AEs.[17] These

potential benefits of IT are particularly crucial for improving the quality and safety of surgical care given its complexity.

Realizing the full potential of IT, our primary objective was to develop an interactive, dynamic, and portable method to record and report AEs, integrated with process of care delivery, namely the Thoracic Surgery Quality Monitoring, Information Management, and Clinical Documentation (TSQIC) system. Second, to use the TSQIC system as a foundation for evaluating temporal trends in operative volume and TM&M rates following major thoracic surgery based on our institution's evolutionary data. Our tertiary objective was to describe the evolution of both the TM&M classification and the TSQIC system, and explore how together they have afforded powerful quality assessment and improvement opportunities.

MATERIALS & METHODS

Study design, setting, and population

Ottawa Hospital Research Ethics Board (OHREB) approval was obtained for a prospective cohort study of post-operative AEs that occurred in patients undergoing operations for thoracic conditions between January 1, 2008 and October 31, 2014, spanning a 7-year period. The Ottawa Hospital is a 1,195-bed academic health sciences center serving a population of 1.35 million people where thoracic surgical care is consolidated at one campus by six thoracic surgeons and two thoracic surgical fellows.

Separate OHREB approval was obtained for the development and implementation of TSQIC. Informed consent was waived for both ethics applications owning to implied consent on the account of the study objectives.

TSQIC development

Paper-based forms that were originally used to support AE reporting involved cumbersome event reporting steps and resulted in inefficient organizational processes when attempting to use the information for overall improvement. Recognizing that software was necessary to efficiently record and review TM&M data, as well as to report the data back to surgeons, iterative development led to an evolution of a secure, real-time, web-based, point-of-care, iPad-optimized TSQIC software system.

TSQIC is divided into three components: i) Mandatory OR, for prospective collection of all thoracic surgical cases within the practice, and TM&M, for tracking all complications associated with the operative procedures performed; ii) automated and dynamic reporting of statistics for surgical volume and quality; and iii) clinical documentation for preoperative clinic consults, daily progress notes, and follow-up notes.

The TSQIC system was iteratively pilot tested and implemented within The Ottawa Hospital's Division of Thoracic Surgery in December 2012. Database support and maintenance requires the equivalent of one full-time employee to be successful. Maintenance of data integrity is crucial and considerable time is spent ensuring that data are complete and accurate.

i) Mandatory Operative Report (OR) and Thoracic Morbidity & Mortality (TM&M) Modules

A recently validated synoptic OR for lung cancer surgery was embedded within the TSQIC system.[18] The creation of this base dataset of the synoptic OR for lung cancer surgery utilized an anonymous modified Delphi method[19,20] using three iterative rounds.[18] Within the synoptic OR module, patients' demographics, medical history, pre- and post-operative diagnoses are also recorded. The synoptic OR was further iteratively refined to include **mandatory** elements relating to all thoracic procedures performed, including lung, pleura, foregut, tracheobronchial, mediastinum and neck, chest and abdominal wall, diaphragm and phrenic nerve, and heart, pericardium, and vascular procedures. Thus, mandatory ORs are completed for all procedures, including those that are done in endoscopy. The surgical procedures may be further risk stratified according to surgical complexity using either the universal ACS-NSQIP risk calculator[21] or the forced expiratory volume (1 sec), age, and diffusing capacity (EVAD) score for patients undergoing major lung resection.[22]

The prospective nature of the mandatory OR module allows a surgeon to record data directly into the database immediately following case completion. Each case takes 5-10 minutes to enter into the database depending on the complexity of the case and patient. At the surgeon's discretion, completion of the mandatory OR can be delegated to a responsible second party (e.g. thoracic surgery fellow).

The TM&M module of TSQIC is designed to systematically track complications associated with each surgical case. Definitions of what constitutes a complication and exactly what should be

reported in the TSQIC system were standardized.[9] Complications that do not fit into the standardized definition list are included under the heading "other complications." All in-hospital AEs (morbidities, deaths, prolonged length of stay, hospital readmissions, emergency room visits, and/or post-operative clinic visits beyond the usual expected visits) must be reported. All AEs are classified as to whether they are related to the index surgical procedure, as rarely they are not. Multiple complications can be entered for each case.

TM&M data collection is a continuous, collegial and divisional activity that is comprised of daily reporting (by thoracic surgical fellows), and weekly review (by nurses, thoracic surgical fellows, and staff surgeons) to facilitate uniform classification of complications. The process of mandatory OR and TM&M recording and reporting using TSQIC is demonstrated in **Figure 5.1**.

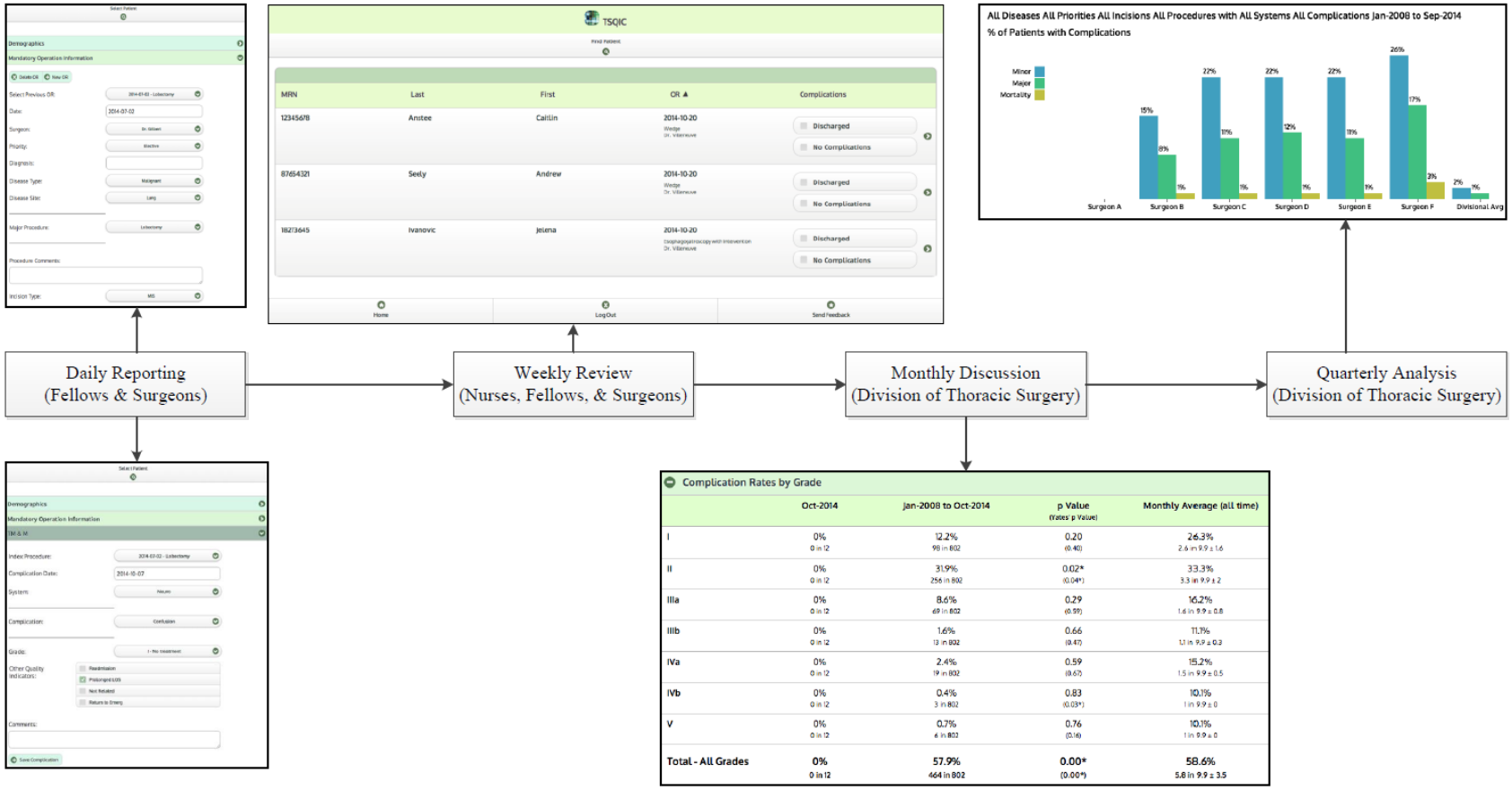


Figure 5.1 TM&M data collection is a continuous, collegial and divisional activity that is comprised of daily reporting (by thoracic surgical fellows and staff surgeons), weekly review and affirmation of complications (by nurses, thoracic surgical fellows, and staff surgeons), monthly discussion at M&M conferences (by entire division), and quarterly analysis of individual surgeon performance. The process of TM&M reporting, review, and analysis has been enhanced by point-of-care TSQIC software.

ii) Automated reporting of surgical volume and quality, and statistical analysis

Data from the combined mandatory OR and TM&M modules are used to generate real-time, dynamic, and interactive volume, complication, and procedure-specific AE reports that are in turn used to support divisional M&M conferences, to inform individual surgeon performance measurement and feedback, to identify areas for local quality improvement efforts, and to serve as a platform for clinical research.

Raw frequency data are queried for each operative procedure and each complication to develop graphic tools, including tables, temporal line graphs, and pie charts to allow for monitoring of surgical volume and quality for the entire patient cohort. Where possible, all data are presented as proportions, means and standard deviations. Chi-square tests are calculated to compare M&M rates in a given time period to our divisional historical AE rates. Yates' correction, a conservative estimate, is added in conjunction with each Chi-square test to prevent overestimation of statistical significance for small data. Yates' correction is used when at least one cell of the table has a count <5. ANOVA is used to analyze historical data for temporal trends. Statistical significance is defined as $p<0.05$. All analyses are custom built and programmed using the PHP 5 scripting language (The PHP Group).

iii) Clinical documentation modules: the future of TSQIC

Future refinements to the TSQIC system are ongoing and entail multidisciplinary point-of-care data entry and clinical documentation of all elements of the preoperative clinic consult, including surgical history and physical examination, daily progress notes, and patient follow-up notes.

Standardization of these processes will facilitate data generation for future research and clinical practice improvement.

RESULTS

Patient demographics

From January 1, 2008 to October 31, 2014, 3384 patients (median age, 61 years; range, 14 to 97 years) underwent 4293 procedures (**Table 5.1**). Of the 3384 patients, 1798 were males (53.1%) and 1586 were females (46.9%). Primary malignancies of the lung remained the most common indication for surgery over time (**Table 5.1**).

Table 5.1 Patient demographics and surgical volume, January 1, 2008 – October 31, 2014.

	n = 4293 cases	%
Patient Demographics		
Sex		
Males	2314	53.9
Females	1979	46.1
Age (mean, range)	61	14 – 97
<64	2290	53.3
65-74	1249	29.1
75-84	692	16.1
>85	82	1.9
Priority of Surgery		
Elective	3934	91.6
Emergent	358	8.3
Surgical Approach/Incision		
Minimally invasive surgery	1846	43.0
Open	1219	28.4
Converted	121	2.8
Disease Type		
Malignant	3285	76.5
Benign	986	23.0
Unknown	21	0.5
Disease Site		
Lung	1659	38.6
Foregut	780	18.2
Mediastinum, diaphragm, trachea, and chest wall	251	5.8
Pleura	113	2.6
Minor	1469	34.2
Other	21	0.5

Temporal trends in surgical volume

Figure 5.2 shows temporal trends in volume of procedures performed. Overall, 1659 lung procedures were performed over 7 years. The proportion of patients who had undergone wedge resections (n=460; 32.0% of total number of procedures; *p*=0.24), segmentectomies (n=73; 4.4% of total number of lung procedures; *p*=0.09), lobectomies (n=898; 54.1% of total number of lung procedures; *p*=0.61), and pneumonectomies (n=73; 4.4% of total number of lung procedures; *p*=0.17) remained stable over time; while the proportion of extended lobectomies performed increased over time (n=28; 1.7% of total number of procedures; *p*=0.02).

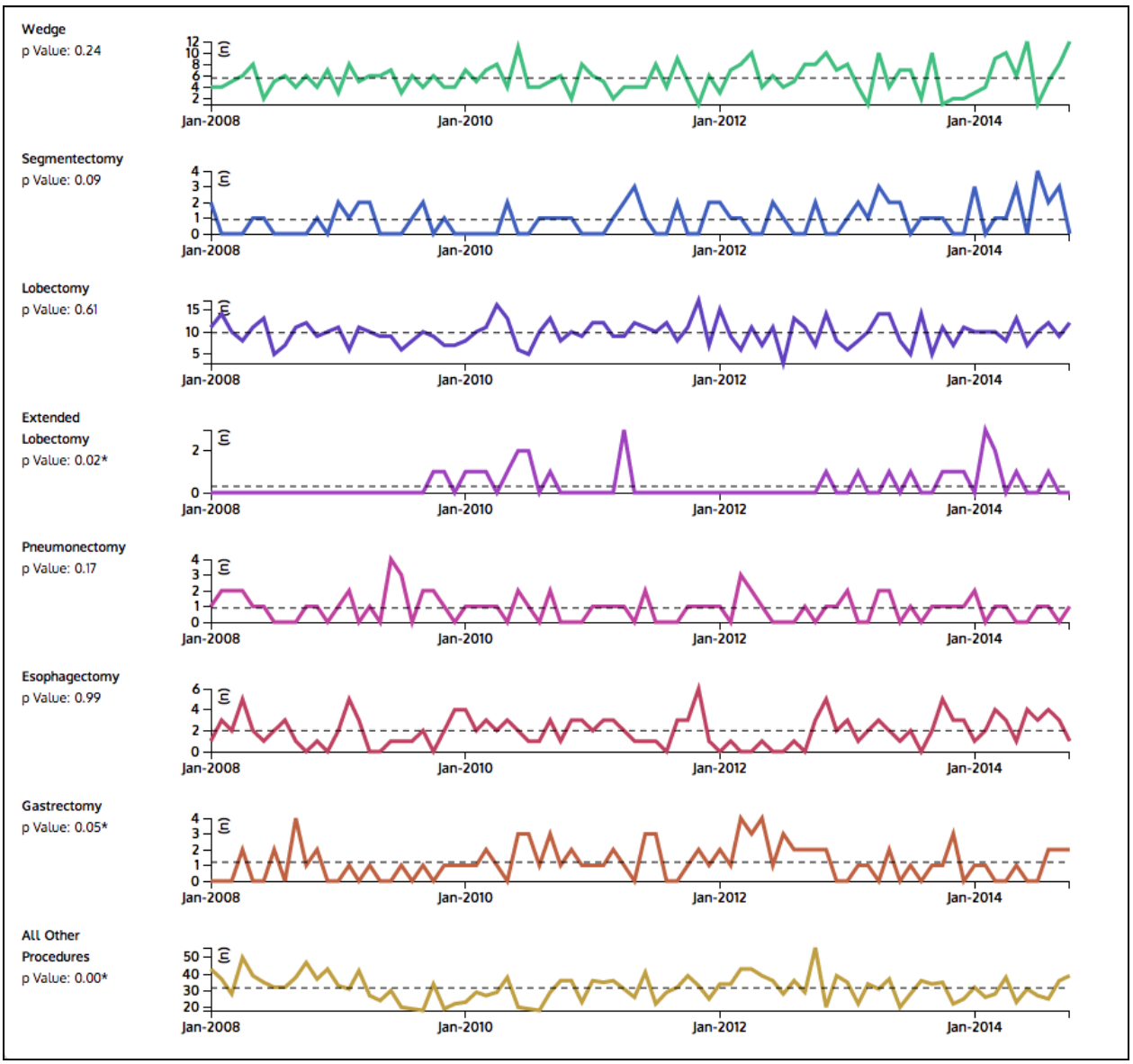


Figure 5. 2 Temporal trends in volume of procedures performed during study period, January 1, 2008 – October 31, 2014.

The total number of lobectomies surpassed the number of segmentectomies, wedge resections, extended lobectomies, and pneumonectomies combined (**Table 5.2**).

Table 5.2 Total patients and complications for all cases and for major thoracic surgical procedures, January 1, 2008 – October 31, 2014.

Complication Grade		Major Surgical Procedure							Total (n=3384 patients) (n=4293 cases)
		Wedge Resection (n=460)	Segmentectomy (n=73)	Lobectomy (n=898)	Extended Lobectomy (n=28)	Pneumonectomy (n=73)	Esophagectomy (n=162)	Gastrectomy (n=96)	
Minor	I	16 (3.5)	7 (9.6)	98 (12.2)	7 (25.0)	5 (6.8)	8 (4.9)	5 (5.2)	204 (4.8)
	II	54 (11.7)	14 (19.2)	256 (31.9)	15 (53.6)	38 (52.1)	96 (59.3)	22 (22.9)	832 (19.4)
Major	IIIa	17 (3.7)	1 (1.4)	69 (8.6)	3 (10.7)	8 (11.0)	49 (30.2)	9 (9.4)	298 (6.9)
	IIIb	7 (1.5)	0 (0)	13 (1.6)	1 (3.6)	5 (6.8)	14 (8.6)	9 (9.4)	129 (3.0)
	IVa	2 (0.4)	0 (0)	19 (2.4)	0 (0)	9 (12.3)	10 (6.2)	3 (3.1)	93 (2.2)
	IVb	2 (0.4)	0 (0)	3 (0.4)	0 (0)	0 (0)	5 (3.1)	1 (1.0)	23 (0.5)
Mortality	V	1 (0.2)	0 (0)	6 (0.7)	0 (0)	3 (4.1)	7 (4.3)	1 (3.1)	47 (1.1)
# Complications		99 (21.5)	22 (30.1)	464 (57.9)	26 (92.9)	68 (93.2)	189 (116.7)	52 (54.2)	1627 (37.9)
# Patients with Complications		75 (16.3)	15 (20.5)	294 (32.7)	17 (60.7)	36 (49.3)	84 (51.9)	31 (32.3)	1317 (30.9)

As expected, the frequency of a video-assisted thoracoscopic surgical approach increased markedly over time (p<0.001) (**Figure 5.3**).

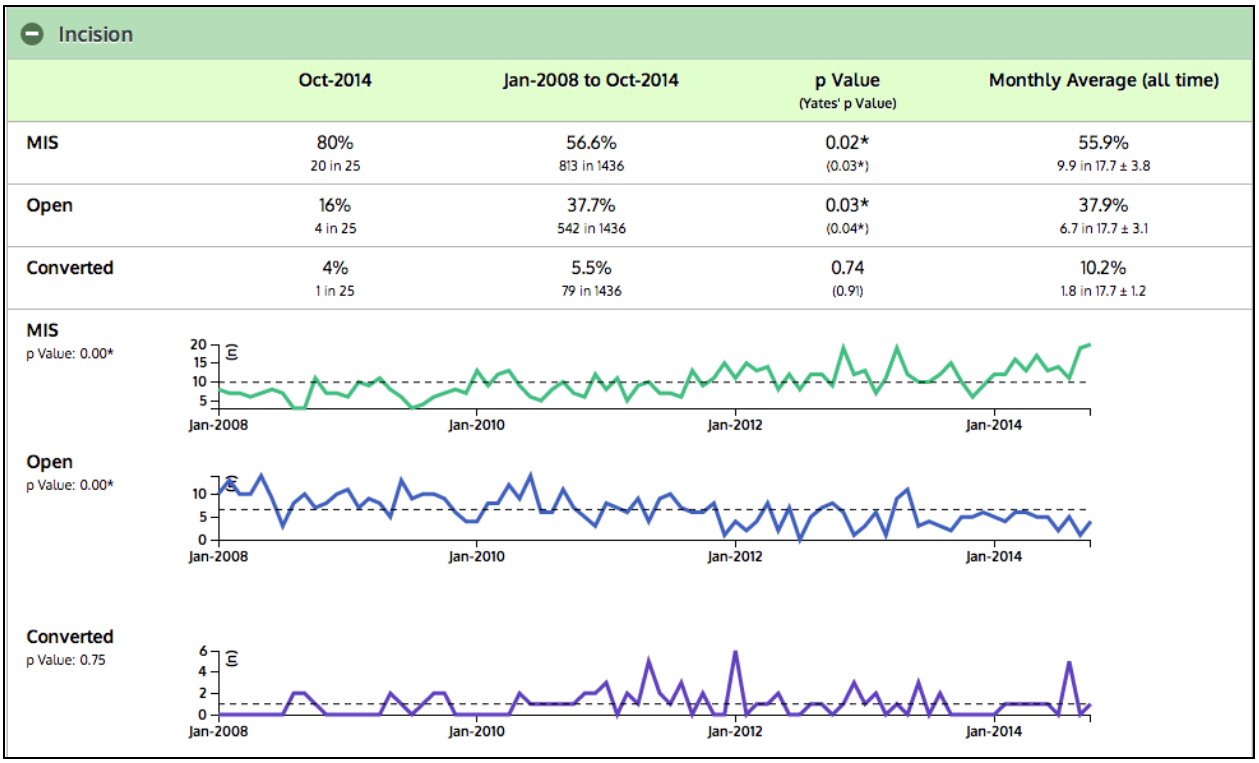


Figure 5.3 Temporal trends indicate that the frequency of a video-assisted thoracoscopic surgical approach increased markedly over time (p<0.001), January 1, 2008 – October 31, 2014.

A total of 780 foregut procedures were performed over 7 years. The proportion of patients who had undergone an esophagectomy remained stable over time (n=162; 20.8% of total number of foregut procedures; p=0.99), while the proportion of patients who had undergone a gastrectomy decreased over time (n=96; 12.3% of total number of foregut procedures; p=0.05).

Procedure-specific AE profiles

Table 5.2 provides data related to overall and procedure-specific AEs. During the study period, 1317 (30.9%) patients had at least one complication. Complications were assigned a severity grade from I-V. Grade I (n=204; 4.8%) and II (n=832; 19.4%) complications require no therapy or pharmacologic intervention only. Grade III (n=427; 9.9%) and IV (n=116; 2.7%) complications require surgical intervention or life support. Grade V (n=47; 1.1%) complications result in patient death.

The percentage of patients with complications varied between procedures. Major complication rates (grades III-IV) for lobectomy, pneunomenctomy, esophagectomy, and gastrectomy were 11.5% (n=104), 30.1% (n=22), 48.1% (n=78), and 22.9% (n=22); mortality rates were 0.7% (n=6), 4.1% (n=3), 4.3% (n=7), and 3.1% (n=1), respectively.

The most commonly recorded complications by system were pleural (n=437; 10.2%), cardiovascular (n=294; 6.8%), and pulmonary (n=277; 6.5%) for all cases (**Figure 5.4**).

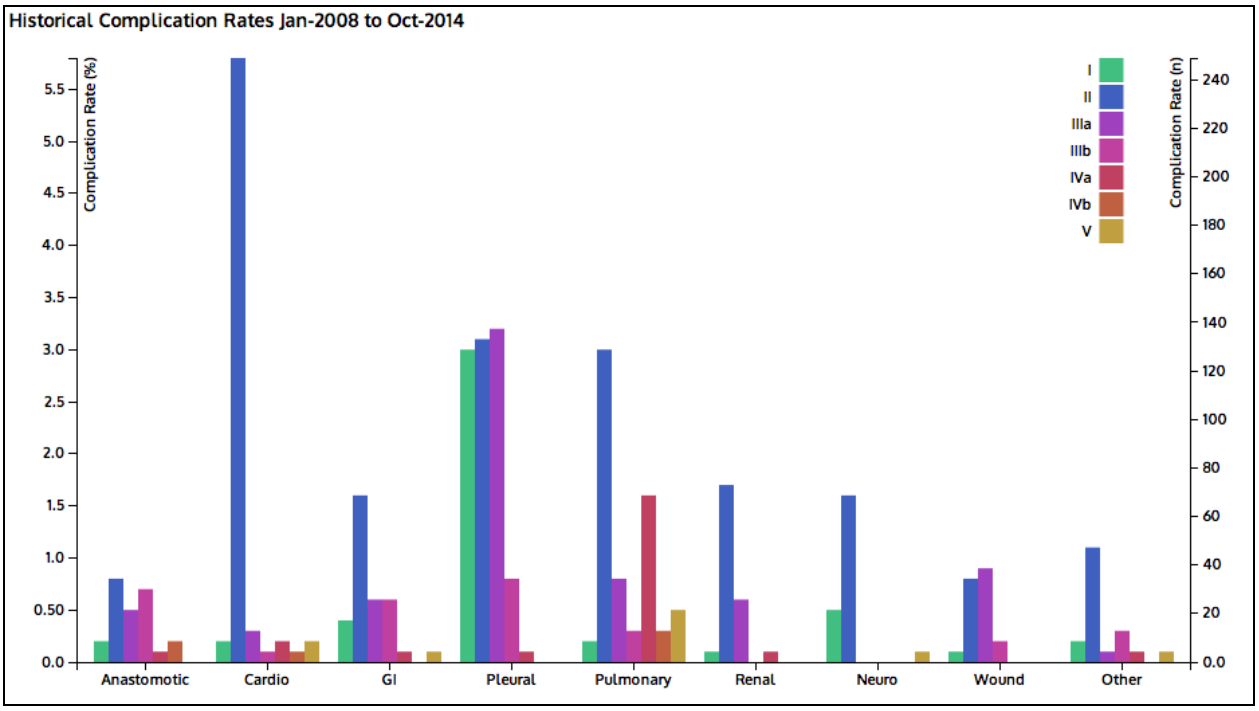


Figure 5.4 The most commonly recorded complications by organ system were pleural (n=437; 10.2%), cardiovascular (n=294; 6.8%), and pulmonary (n=277; 6.5%) for all cases.

Prolonged air leak (n=153; 10.7%), atrial fibrillation (n=121; 8.4%), and pneumonia (n=34; 2.4%) were identified as the most common complications after pulmonary resection (**Figure 5.5**).

AE Profile for Extended Lobectomy,Lobectomy,Pneumonectomy,Segmentectomy,Wedge Jan-2008 to Oct-2014

System	I	II	IIIa	IIIb	IVa	IVb	V	Total
Cardiovascular								**9.3%** 133 in 1436
Atrial Arrhythmia	0.1% 2 in 1436	7.9% 113 in 1436	0.1% 2 in 1436		0.2% 3 in 1436		0.1% 1 in 1436	8.4% 121 in 1436
Hypotension		0.2% 3 in 1436					0.1% 1 in 1436	0.3% 5 in 1436
MI		0.3% 4 in 1436			0.1% 1 in 1436	0.1% 2 in 1436		0.5% 7 in 1436
Pleural								**15.3%** 220 in 1436
Empyema		0.8% 11 in 1436	0.4% 6 in 1436	0.1% 1 in 1436				1.3% 18 in 1436
Prolonged Air Leak	4.9% 70 in 1436	4.7% 67 in 1436	0.7% 10 in 1436	0.4% 6 in 1436				10.7% 153 in 1436
SubQ Emphysema		0.5% 7 in 1436	1.4% 20 in 1436					3.4% 49 in 1436
Pulmonary								**3.9%** 56 in 1436
Atelactasis	0.1% 2 in 1436	0.7% 10 in 1436	0.4% 6 in 1436					1.3% 18 in 1436
Edema		0.2% 3 in 1436			0.1% 1 in 1436			0.3% 4 in 1436
Pneumonia		1.9% 28 in 1436	0.1% 2 in 1436		0.3% 4 in 1436			2.4% 34 in 1436
Total	**6.8%** 97 in 1436	**17.1%** 246 in 1436	**3.2%** 46 in 1436	**0.5%** 7 in 1436	**0.6%** 9 in 1436	**0.1%** 2 in 1436	**0.1%** 2 in 1436	**28.5%** 409 in 1436

Figure 5.5 Prolonged air leak (n=153; 10.7%), atrial fibrillation (n=121; 8.4%), and pneumonia (n=34; 2.4%) were identified as the most common complications after pulmonary resection.

When all of the cases are summed, 17.0% (n=728) of AEs led to prolonged length of stay, 6.8% (n=294) led to hospital readmission, 0.3% (n=12) led to a return to emergency, and 2.5% (n=108) of AEs were unrelated to the index surgical procedure (data not shown).

Temporal trends in thoracic morbidity & mortality

Unadjusted temporal trends in overall and procedure-specific AEs were analyzed. When all of the cases are summed (n=4293 cases), a statistically significant increase in Grade I complications ($p<0.001$), a statistically significant decrease in Grade II complications ($p<0.001$), and statistically significant increase in overall M&M rates ($p<0.001$) were noted, as shown in **Figure 5.6**.

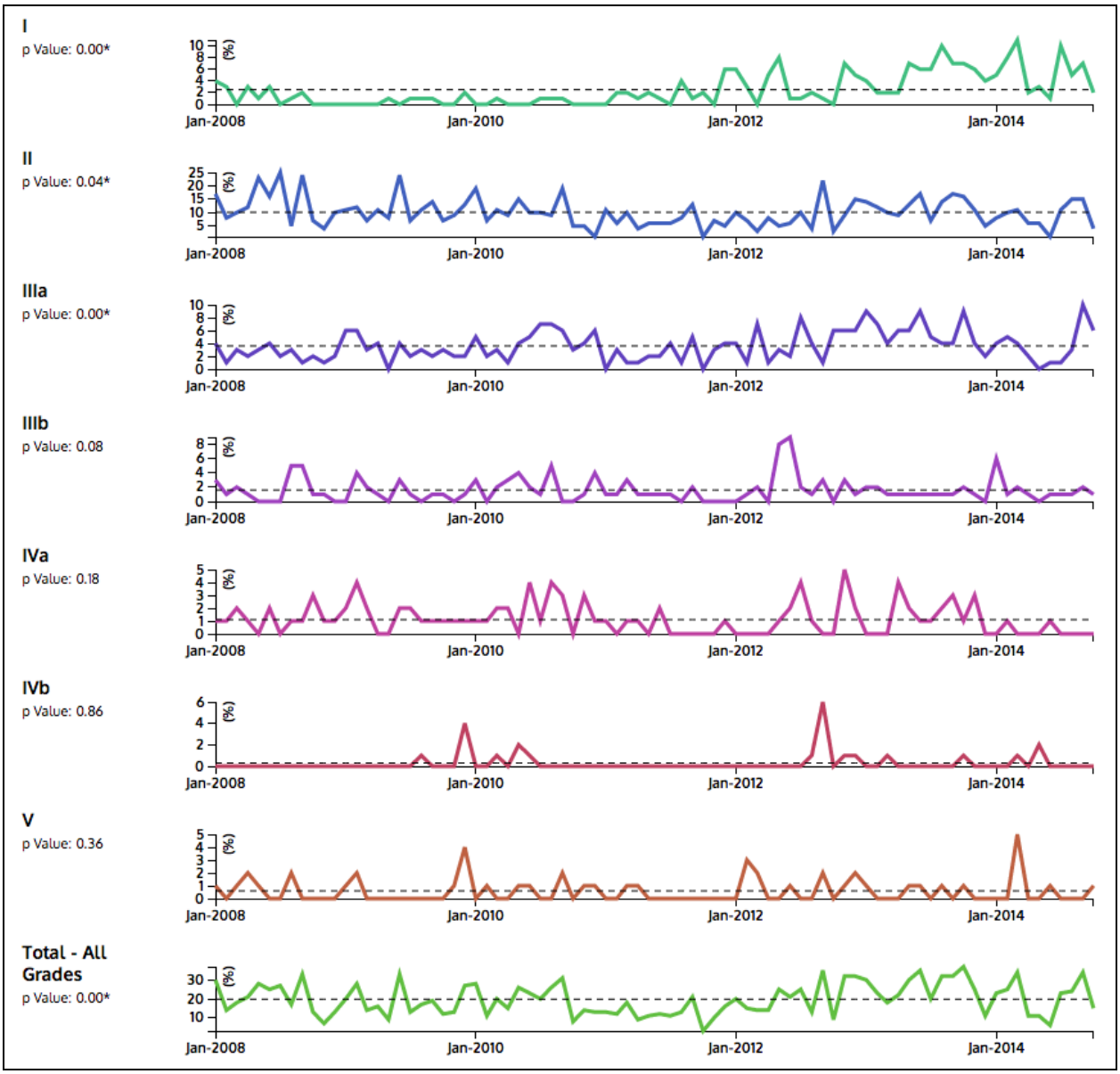


Figure 5.6 Unadjusted temporal trends in overall complication rates were analyzed. When all of the cases are summed (n=4293 cases), a statistically significant increase in Grade I complications (p<0.001), a statistically significant decrease in Grade II complications (p<0.001), and statistically significant increase in overall M&M rates (p<0.001) were noted during the study period.

A statistically significant decrease in Grade II complications (p=0.01), and statistically significant increase in overall M&M rates (p=0.01) was noted for patients undergoing lobectomy. There were no significant changes in respective complications or mortality rates for

patients undergoing pneumonectomy or gastrectomy. A statistically significant increase in observed overall M&M rates was noted for patients undergoing esophagectomy. Data not shown.

Post-implementation of surgeon-specific outcome reports and positive deviance seminars for continuous quality improvement

Preliminary data suggests a non-significant decrease in overall rates of post-operative atrial fibrillation following pulmonary resection (8.7% vs 7.0%; p=0.42) post-implementation of surgeon-specific outcome reports (SSORs) and quarterly positive deviance (PD) seminars (01/2012 to 11/2013 vs. 12/2013 to 02/2015) (**Figure 5.7**), a statistically significant decrease in overall rates of prolonged alveolar air leak following pulmonary resection (21.6% vs. 7.8%; p<0.001) (**Figure 5.8**), and a statistically significant decrease in grade II rates of anastomotic leak following esophagectomy (14.3% vs. 0%; p<0.04) (**Figure 5.9**).

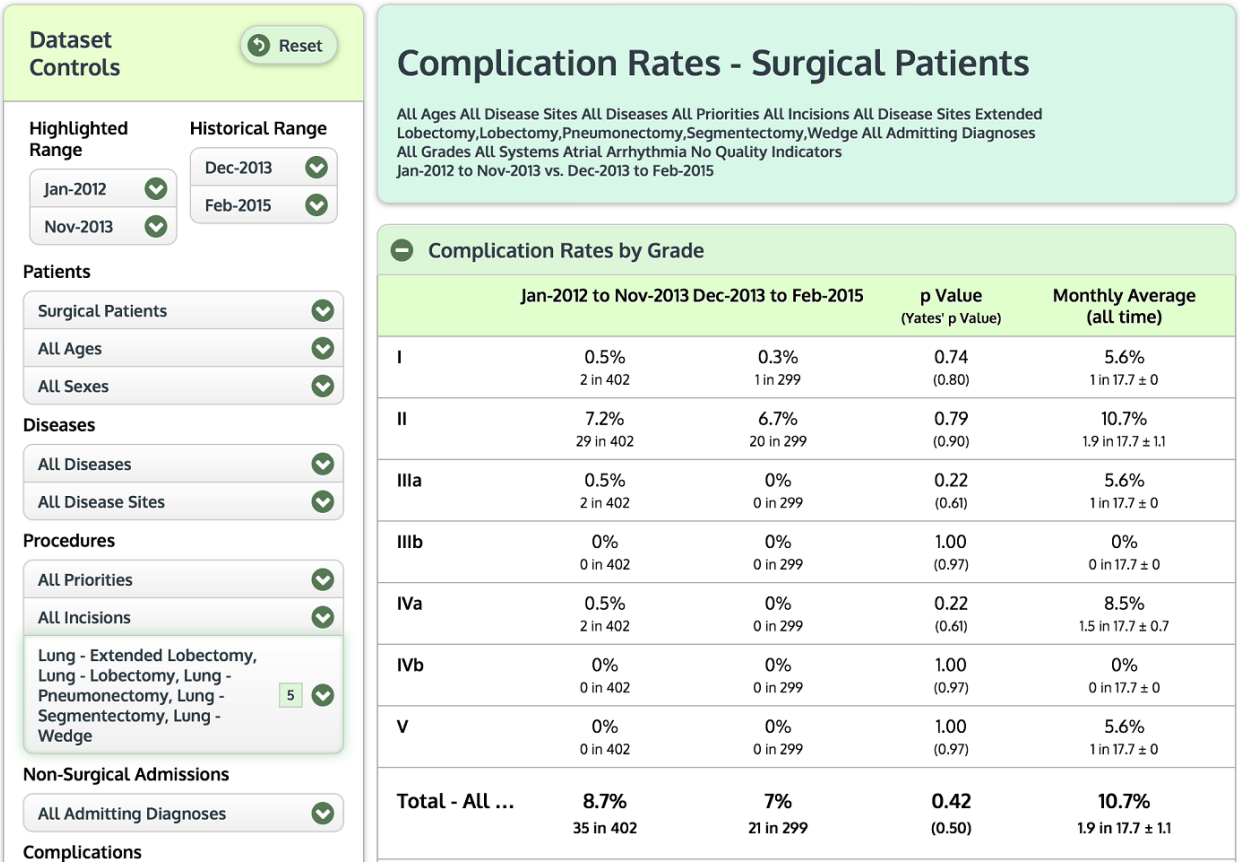
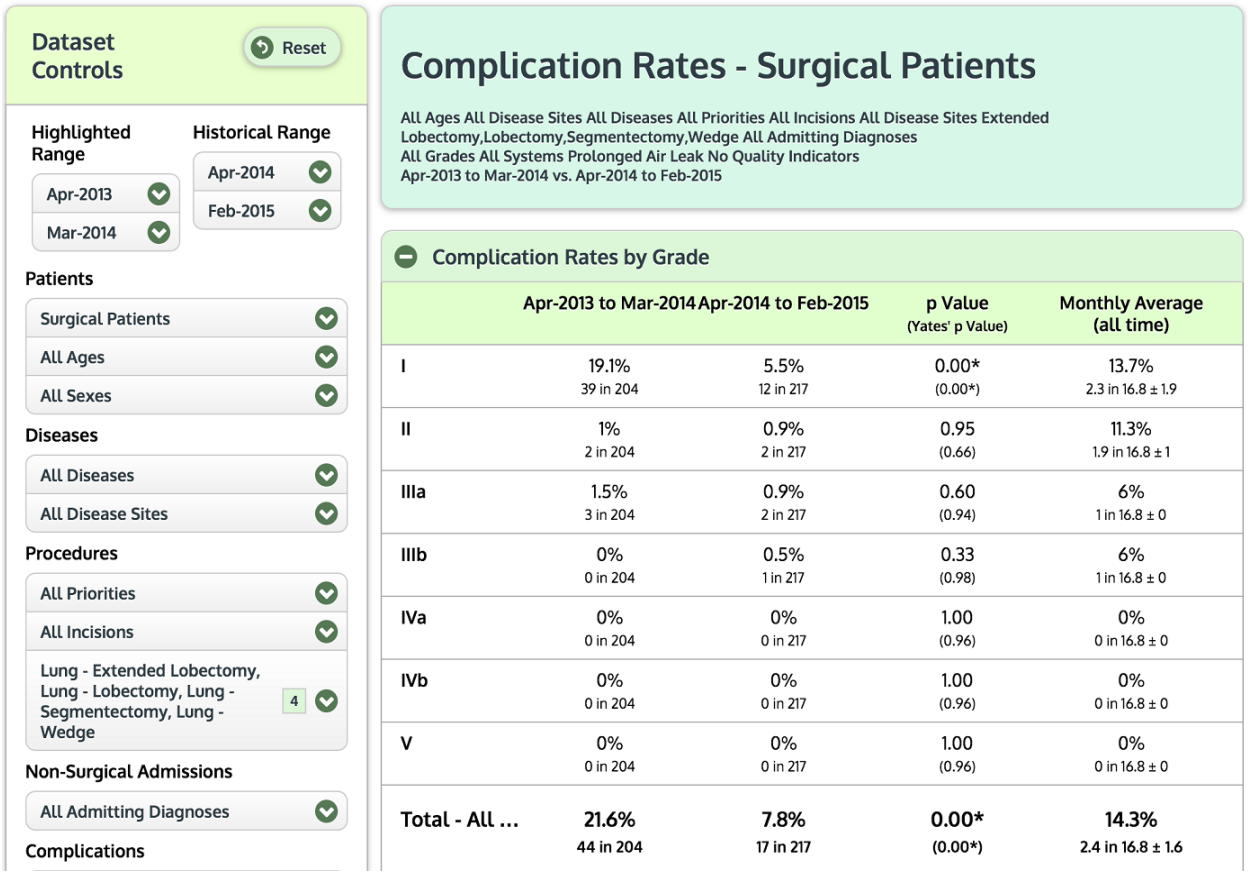
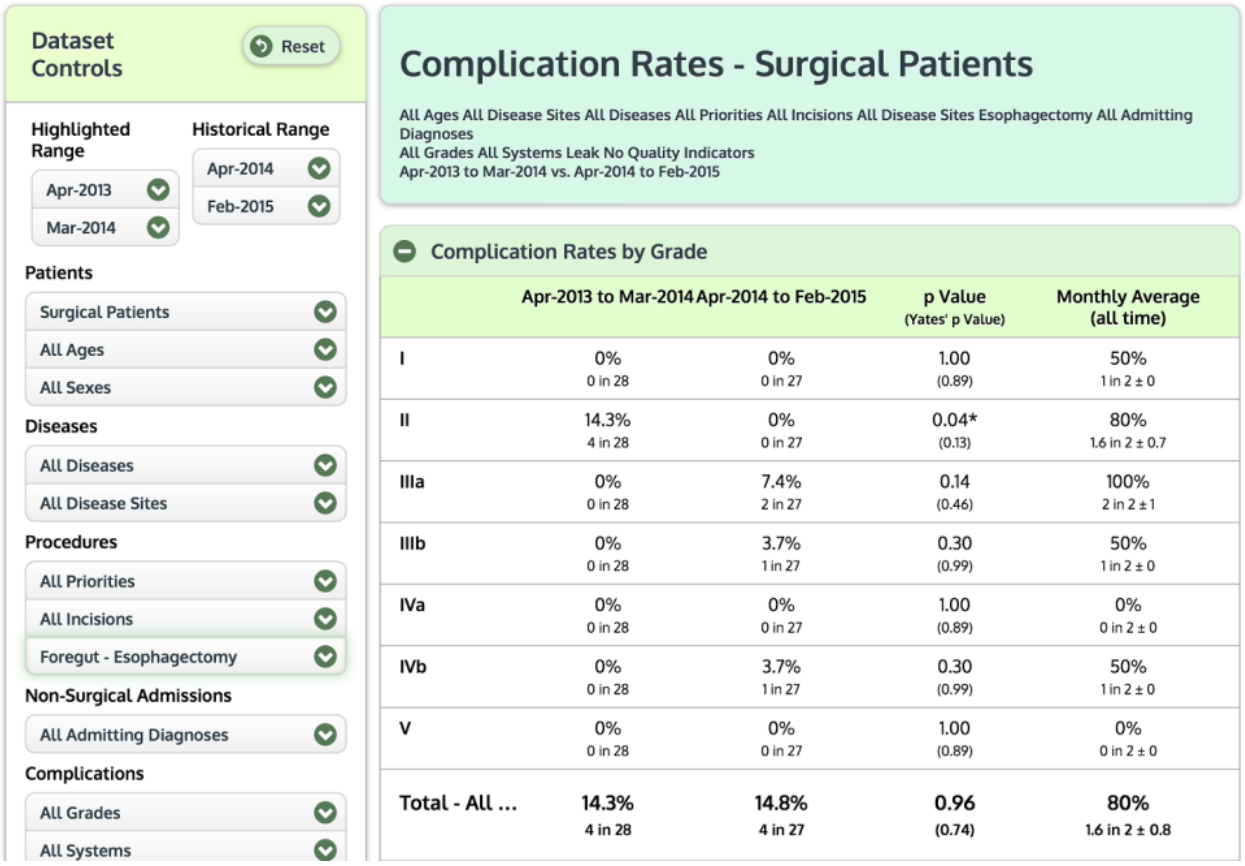


Figure 5.7 Preliminary data suggests non-significant decrease in overall rates of post-operative atrial fibrillation following pulmonary resection post-implementation of surgeon-specific outcome reports and positive deviance seminars.



	Apr-2013 to Mar-2014	Apr-2014 to Feb-2015	p Value (Yates' p Value)	Monthly Average (all time)
I	19.1% 39 in 204	5.5% 12 in 217	0.00* (0.00*)	13.7% 2.3 in 16.8 ± 1.9
II	1% 2 in 204	0.9% 2 in 217	0.95 (0.66)	11.3% 1.9 in 16.8 ± 1
IIIa	1.5% 3 in 204	0.9% 2 in 217	0.60 (0.94)	6% 1 in 16.8 ± 0
IIIb	0% 0 in 204	0.5% 1 in 217	0.33 (0.98)	6% 1 in 16.8 ± 0
IVa	0% 0 in 204	0% 0 in 217	1.00 (0.96)	0% 0 in 16.8 ± 0
IVb	0% 0 in 204	0% 0 in 217	1.00 (0.96)	0% 0 in 16.8 ± 0
V	0% 0 in 204	0% 0 in 217	1.00 (0.96)	0% 0 in 16.8 ± 0
Total - All ...	**21.6% 44 in 204**	**7.8% 17 in 217**	**0.00* (0.00*)**	**14.3% 2.4 in 16.8 ± 1.6**

Figure 5.8 Preliminary data suggests a statistically significant decrease in overall rates of post-operative prolonged alveolar air leak following pulmonary resection post-implementation of surgeon-specific outcome reports and positive deviance seminars.



	Apr-2013 to Mar-2014	Apr-2014 to Feb-2015	p Value (Yates' p Value)	Monthly Average (all time)
I	0% 0 in 28	0% 0 in 27	1.00 (0.89)	50% 1 in 2 ± 0
II	14.3% 4 in 28	0% 0 in 27	0.04* (0.13)	80% 1.6 in 2 ± 0.7
IIIa	0% 0 in 28	7.4% 2 in 27	0.14 (0.46)	100% 2 in 2 ± 1
IIIb	0% 0 in 28	3.7% 1 in 27	0.30 (0.99)	50% 1 in 2 ± 0
IVa	0% 0 in 28	0% 0 in 27	1.00 (0.89)	0% 0 in 2 ± 0
IVb	0% 0 in 28	3.7% 1 in 27	0.30 (0.99)	50% 1 in 2 ± 0
V	0% 0 in 28	0% 0 in 27	1.00 (0.89)	0% 0 in 2 ± 0
Total - All ...	**14.3% 4 in 28**	**14.8% 4 in 27**	**0.96 (0.74)**	**80% 1.6 in 2 ± 0.8**

Figure 5.9 Preliminary data suggests a statistically significant decrease in grade II rates of post-operative anastomotic leak following esophagectomy post-implementation of surgeon-specific outcome reports and positive deviance seminars.

DISCUSSION

Improving quality and patient safety is a fundamental priority given the pervasive problem of post-operative AEs; reported rates in the surgical literature have varied between 26% to 42% for inpatients undergoing general, vascular, and cardiothoracic surgery.[23] This variation can in part be traced to the lack of consistent or standardized definitions for procedure-specific complications, differences in case-mix, and differences in observation time. Accurate reporting of post-operative AEs is thus essential to effectively analyze and compare quality of surgical care.

In our center, we have developed a particular interest in surgical outcomes research for many years. In 2008, created in accordance to the Clavien-Dindo classification, we proposed and integrated a classification of thoracic surgical complications within our division allowing objective assessment of post-operative AEs, and embodying a means of quality control. However, the traditional paper-based approach which was originally used to support AE reporting was inefficient, and limited in its ability to rapidly and systematically analyze data.

Modern information technology has great potential for improving the safety of surgical care by overcoming the limitations posed by traditional paper-based methods. In addition, several influential reports on patient safety have highlighted the importance of the development of effective learning systems to reduce the occurrence of preventable AEs.[24] Over four years we developed, pilot-tested, and optimized a real-time, web-based, point-of-care Thoracic Surgery Quality Monitoring, Information management, and Clinical documentation (TSQIC) system to detect, capture, classify, analyze, and monitor all TM&M following all surgery. Beyond creating

an online AE data repository, however, we were also interested in using the TSQIC system to create procedure-specific AE profiles, and to identify and analyze temporal trends in rates of post-operative AEs based on our institution's 7-year evolutionary data.

Using the TSQIC system, we demonstrated that post-operative AEs occurred in 37.9% of all cases. About two thirds of all complications were grades I or II (i.e. they did not require invasive therapies). Prolonged air leak, atrial fibrillation, and pneumonia were ranked as the top three among the different complications evaluated in terms of incidence. Although our procedure-specific AEs mortality rates compare favourably with those reported in the surgical literature, our detection method differs in several important respects resulting in different incidence rates of post-operative morbidities. This observation is novel as other recent reports have suggested much lower incidence rates of overall post-operative AEs following major thoracic surgery, including 54.2% following pulmonary lobectomy,[25] 33.0% following pneumonectomy,[26] 24.4% following esophagectomy,[27] and 11.0% following gastrectomy.[28] Multiple studies have demonstrated that 65% to 91% of AEs fail to be reported.[29,30] Several reasons have been suggested regarding why surgeons do not report their complications, including time constraints and concerns of punitive actions, litigation, and diminished reputation in the opinion of their collegues.[31,32] Predefined definitions and severity grading, and a team-approach within our division have ensured that complications are entered in a uniform and accurate fashion. In addition, the use of the prospective, easily accessible, up-to-date, TSQIC system greatly facilitates accrual and analysis of complication data.

The evolutionary analysis of temporal trends is challenging to perform rigorously, as well as, challenging to interpret. We noted a significant increase in the observed rate of grade I complications, which may be explained by both an increasing capture of minor complications, and a reflection of the prospective nature of our data collection methods. We noted a significant reduction in the observed rate of grade II complications; while encouraging, the observation is not a definitive testament to a fundamental improvement in care. The incidence of more severe forms of complications (i.e. grades III, IV, and V) has remained unchanged over time. We noted a significant increase in the overall rate of complications during the study period. Again, increase in outcomes may be explained by chance, differences in case-mix over time, increasing capture of minor complications rather than deterioration in quality of care, or a combination of these.

Our evolutionary data shows that measurement is necessary but not sufficient for quality improvement[33] without also providing surgeons with information and the infrastructure on how to improve patient outcomes. Bridging this gap has been a top priority for our quality improvement program. To do so, we have used the TM&M classification and the TSQIC system as a foundation for several quality improvement initiatives.

First, the divisional M&M conference has been greatly enhanced by the improved quality of statistical reporting of all surgical cases and all post-operative AEs for the preceding month, allowing the division to follow trends over time, and to distinguish between random variation and a significant increase or decrease in the rate of post-operative AEs. The TSQIC system allows data about common complications to be presented in the context of the division's broader experience. Following the statistical overview, the chief thoracic surgical fellow then identifies

specific cases for discussion using a structured format.[34] Up to three cases are discussed at M&M conferences, and are selected according to the following criteria: i) AE resulting in death, disability, harm or injury; ii) preventability; iii) lessons to be learned about cognitive and system issues. Discussion is followed by creation and dissemination of bottom lines or action items related to each case.[34] The complementary nature of aggregated surgical volume and quality reporting, with structured case selection and analysis, leads to powerful quality assessment and improvement opportunities monthly.

Second, using the TSQIC platform, we have developed novel and dynamic surgeon-specific outcome reports (SSORs), allowing surgeons to compare their performance to that of their colleagues through automated, real-time, and risk-adjusted data monitoring. Ongoing participation in a national, regional, or local outcomes database or quality assessment program is mandatory for maintenance of certification (MOC) in surgery in support of lifelong learning needs. Given its inherent self-assessment nature leading to quality improvement, SSORs are eligible for MOC credits by the Royal College of Physicians and Surgeons.

A priority for the division has been to ensure such monitoring translates into tangible and reproducible improvements in surgical performance, for the benefit of surgeons, and ultimately patients. We have thus, implemented quarterly continuous quality improvement (CQI) seminars to provide an additional forum for discussion regarding collective results, utilizing the technique of positive deviance, to unmask best performers as a catalyst for discussing practice measures to improve specific AEs. The provision of timely feedback is congruent with the rapid-learning

health care initiative where real-time clinical data are used to support a focus on quality improvement and patient safety.[35]

Whether structured case analyses, SSORs and PD can lead to improvements in rates of post-operative AEs is a matter of ongoing research. Previous studies by our division have demonstrated that atrial fibrillation (AFIB), prolonged alveolar air leak (PAAL), and anastomotic leak represent significant morbidity in the post-operative period, and preliminary data suggests a non-significant decrease in overall rates of AFIB following pulmonary resection post-implementation of SSORs and quarterly PD seminars, a statistically significant decrease in overall rates of PAAL following pulmonary resection, and a statistically significant decrease in grade II rates of post-operative anastomotic leak following esophagectomy. While encouraging, these observations are not a definitive testament to a fundamental improvement in care; because our data lacks a control group, it is difficult to know whether such changes in complication rates represent longitudinal improvements in outcomes. Further sensitivity analyses are required in order to determine if the observed improvements occurred are related to SSOR use and adoption, PD seminar implementation, and changes in practice.

Last, while not a primary objective, embracing a culture of quality and safety has been an important outcome of this research program. Because the process of TM&M data collection and analysis is integrated with monthly M&M conferences, SSORs, and quarterly PD seminars, it has been viewed as a multifaceted approach for creating a continuous culture of quality improvement that is supportive of teamwork, safety, accountability, transparency, and open and collegial dialogue regarding the links among structure, process, and outcomes of surgical care.

It is important to recognize several limitations to this work. First, the TM&M classification of AEs relies on consistent vigilance and may be susceptible to reporting bias. Second, the current analysis shows that 30.9% patients had at least one AE during the seven-year study period. When the post-operative AEs are further subdivided into severity grades, the number of events in each severity grade is very small, and the power to draw accurate conclusions regarding surgical quality in each stratum is limited. Third, whenever we speak of surgical quality, the case-mix must be considered. We have yet to begin using risk-adjustment in the presentation of our overall complication data. Future work will be needed to further refine the risk-adjustment strategy that accounts for the heterogeneity of patient populations and the baseline risks of different categories of procedures. Last, this is a single institution study; as such, it is unclear if this approach may be successfully re-created elsewhere. Furthermore, what remains to be evaluated is whether detailed and real-time AE monitoring, structured M&M conferences, SSORs, and PD seminars can demonstrably enhance patient outcomes.

Conclusion

The TM&M classification schema offers prospective, standardized, and reliable definitions to accurately report all post-operative AEs over time and in a consistent manner, while the TSQIC system provides an effective method for real-time post-operative AE data entry, and analysis. Together, these applications offer a potentially powerful initiative that has the ability to improve quality and safety of surgical care. The evidence relating to the use of this program to date: confirms ease of use for both surgical fellows and staff surgeons; documents all post-operative AEs following all surgeries to identify patterns of AEs that otherwise may have gone unreported;

supports collection of surgeon-specific performance data for self-assessment and benchmarking purposes; and suggests a positive cultural change of the opinions of surgeons to data collection and outcomes monitoring. It is possible that adding other interventions to the measurement and monitoring of post-operative AEs, such as structured M&M conferences, and efforts to identify and disseminate best practices, will lead to improved outcomes.

ACKNOWLEDGEMENTS

We would like to acknowledge the dedication of past and current thoracic surgical fellows for ensuring efficient and consistent AE reporting, and for their incessant commitment to embracing a culture of quality and patient safety, including: Dr. Derar Alshahab, Dr. Ahmed Alhussaini, Dr. Saleh Abu-Daff, Dr. Anas Alshuhayeb, Dr. Anna McGuire, Dr. Ramzi Addas, and Dr. Daniel French.

DISCLOSURES

No relevant financial relationships with commercial interests to disclose.

AUTHOR CONTRIBUTIONS

Study Conception and Design: Ivanovic, Jelena; Seely, Andrew; Anstee, Caitlin

Acquisition of Data: Anstee, Caitlin; Ivanovic, Jelena

Analysis and Interpretation of Data: Ivanovic, Jelena; Seely, Andrew

Drafting of Manuscript: Ivanovic, Jelena

Critical Revision: Ivanovic, Jelena; Seely, Andrew; Ramsay, Tim; Gilbert, Sebastian; Maziak, Donna; Shamji, Farid; Sundaresan, Sudhir; Finley, Christian

REFERENCES

1. de Vries EN, Ramrattan MA, Smorenburg SM, et al. The incidence and nature of in-hospital adverse events: A systematic review. Qual Saf Health Care 2008;17:216–23 .
2. Vonlanthen R, Slankamenac K, Breitenstein S, et al. The impact of complications on costs of major surgical procedures: A cost analysis of 1200 patients. Ann Surg 2011;254:907–13.
3. Chorney JM, McGrath P, Finley GA. Pain as the neglected adverse event. CMAJ 2010;182: 732.
4. Khan NA, Quan H, Bugar JM, et al. Association of post-operative complications with hospital costs and length of stay in a tertiary care center. J Gen Intern Med 2006;21:177–80.
5. Khuri SF, Henderson WG, DePalma RG, et al. Determinants of long-term survival after major surgery and the adverse effect of post-operative complications. Ann Surg 2005;242**:** 326–43.
6. Martin RC, Brennan MF II, Jaques DP. Quality of complication reporting in the surgical literature. *Ann Surg* 2002;235:803–13.
7. Clavien PA, Sanabria JR, Strasberg SM. Proposed classification of complications of surgery with examples of utility in cholecystectomy. Surgery 1992;111:518–26.
8. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: A new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205–13.
9. Seely AJE, Ivanovic J, Threader J, et al. Systematic classification of morbidity and mortality after thoracic surgery. Ann Thor Surg 2010;90:936–42.
10. Ivanovic J, Threader J, Alhussaini A, et al. Testing the reliability and reproducibility of the Ottawa Hospital Thoracic Morbidity and Mortality classification system. Ann Thor Surg 2011;91:387–93.
11. Liang S, Ivanovic J, Gilbert S, et al. Quantifying the incidence and impact of post-operative prolonged alveolar air leak after pulmonary resection. J Thorac Cardiovasc Surg 2013;145:948–54.
12. Ivanovic J, Ramzan S, McGuire AL, et al. Incidence, severity and perioperative risk factors for atrial fibrillation following pulmonary resection. Interact Cardiovasc Thorac Surg 2014;18:340–46.
13. Ivanovic J, Seely AJE, Anstee C, et al. Measuring Surgical Quality: Comparison of post-operative adverse events with the American College of Surgeons NSQIP and the Thoracic Morbidity and Mortality Classification System. J Am Coll Surg 2014;218;1024–31.
14. Salati M, Pompili C, Refai M, et al. The use of the Thoracic Morbidity and Mortality system for the internal analysis of performance: A case-matched temporal audit. Eur J Cardiothorac Surg 2014;45:859–63.
15. Salati M, Refai M, Pompili C, et al. Major morbidity after lung resection: A comparison between the European Society of Thoracic Surgeons Database system and the Thoracic Morbidity and Mortality system. J Thorac Dis 2013;5:217–22.
16. Infante M, Chiesa G, Solomon D, et al. Surgical procedures in the DANTE trial: A randomized study of lung cancer early detection with spiral computed tomography: Comparative analysis in the screening and control arm. J Thorac Onc 2011;6:327–35.

17. Bates DW. Using information technology to improve surgical safety. British J Sur 2004;91**:** 939–40.
18. Schneider L, Shargall Y, Schieman C, et al. Design of a consensus-derived synoptic operative report for lung cancer surgery. Ann Thorac Surg 2014;97:1163–68.
19. Cantrill JA, Sibbald B, Buetow S. Indicators of the appropriateness of long-term prescribing in general practice in the United Kingdom: Consensus development, face and content validity. Qual Health Care 1998;7:130–35.
20. Campbell SM, Cantrill JA, Roberts D. Prescribing indicators for UK general practice: Delphi consultation study. BMJ 2000;321:425–28.
21. Bilimoria KY, Liu Y, Paruch JL, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: A decision aid and informed consent tool for patients and surgeons. J Am Coll Surg 2013;217:833–42.
22. Ferguson MK, Durkin AE. A comparison of three scoring systems for predicting complications after major lung resection. Euro J Cardiothorac Surg 2003;23:35–42.
23. Healy M, Shackford S, Osler T. Complications in surgical patients. Arch Surg 2002;137:611–18.
24. Kohn LT, Corrigan JM, Donaldson MS: To Err Is Human: Building a Safer Health System. Washington, DC: National Academy Press; 1999.
25. Rueth NM, Parsons HM, Habermann EB, et al. The long-term impact of surgical complications after resection of stage I non-small cell lung cancer: A population-based survival analysis. Ann Surg 2011;254:368–74.
26. Thomas PA, Berbis J, Baste JM, et al. Pneumonectomy for lung cancer: Contemporary national early morbidity and mortality outcomes. J Thorac Cardiovasc Surg 2015;149:73−83.
27. Noble F, Curtis N, Harris S, et al. Risk assessment using a novel score to predict anastomotic leak and major complications after oesophageal resection. J Gastrointest Surg 2012;16:1083–95.
28. Miki Y, Tokunaga M, Tanizawa Y, et al. Perioperative risk assessment for gastrectomy by Surgical Apgar Score. Ann Surg Oncol 2014;21:2601–07.
29. Feldman L, Barkun J, Barkun A, et al. Measuring post-operative complications in general surgery patients using an outcomes-based strategy: Comparison with complications presented at morbidity and mortality rounds. Surg 1997;122:711–20.
30. Hutter MM, Rowell KS, Devaney LA, et al. Identification of surgical complications and deaths: An assessment of the traditional surgical morbidity and mortality conference compared with the American College of Surgeons National Surgical Quality Improvement Program. J Am Coll Surg 2006;203:618–24.
31. Leape LL. Reporting of adverse events. N Engl J Med 2002;347:1633−38.
32. Weissman JS, Annas CL, Epstein AM, et al. Error reporting and disclosure systems: Views from hospital leaders. JAMA 2005;293:1359−66.
33. Berwick DM, James B, Coye MJ. Connections between quality measurement and improvement. Med Care 2003;41:I30-38.
34. Calder LA, Kwok ES, Cwinn AA, et al. Enhancing the quality of morbidity and mortality rounds: The Ottawa M&M Model. Acad Emerg Med 2014;21:314–21.
35. Abernethy AP, Etheredge LM, Ganz PA, et al. Rapid-learning system for cancer care. J Clin Oncol 2010;28:4268–74.

CONCLUSIONS & IMPLICATIONS

Assessing adequacy of surgical care begins with the traditional Donabedian standards for performance measurement, which are the measures of structure, process, and outcomes. Outcomes, which are inherently patient-centered, have been the most measured, and indeed fundamental to evaluating the quality of surgical care. Specifically, post-operative adverse events (AEs) remain the most frequently measured and reported outcomes. AEs lead to substantial physical, emotional and financial damage, and drive patient mortality.

In 2008, created in accordance to the Clavien-Dindo classification, we proposed, developed, and iteratively refined a classification of Thoracic Morbidity & Mortality (TM&M) within The Ottawa Hospital's Division of Thoracic Surgery allowing objective and standardized assessment of all post-operative AEs.

The TM&M classification system was rigorously evaluated over the first two years of its implementation at The Ottawa Hospital, and results showed that this standardized classification system for identifying presence and severity of thoracic surgical complications is feasible, facilitates objective comparison, identifies burden of illness of specific complications, and provides an effective method for continuous surgical quality assessment within an individual division of thoracic surgery.

However, lists and rates of specific complications, although useful, cannot be analyzed in a meaningful way, nor compared across multiple studies or institutions. In a subsequent study, an attempt was made to validate the application of the TM&M classification system across multiple

institutions through a national survey of the members of the Canadian Association of Thoracic Surgeons (CATS). Using kappa statistics, the degree to which different members gave consistent grades of the same case scenario were assessed. Results of CATS membership survey revealed that that the TM&M classification system offers high inter-rater reliability or agreement on the classification of post-operative complications. The classification system was further regarded as straightforward by 98% of the survey respondents, reproducible by 94%, logical by 92%, and useful by 98% of respondents.

In this thesis, we have continued and extended the evaluation and the clinical use of the Thoracic Morbidity & Mortality (TM&M) classification system. In the first two studies, an attempt is made to provide an overview of the burden and impact of the two most pervasive post-operative AEs have on the thoracic surgical patient population, and identify high-risk patient sub-groups. Prolonged alveolar air leak (PAAL) is a common complication following pulmonary resection. PAAL increases the hospital length of stay, increases hospital costs, and is associated with elevated rates of other complications. A contemporary, practical definition of PAAL is an air leak that lasts beyond post-operative day 5, and formed the basis of the study in **Chapter I**. The use of the TM&M classification system further allowed for the determination of the severity and burden, and risk factors of PAAL. These factors include severe radiologic emphysema, histopathologic emphysema, percentage of predicted value for forced expiratory volume in 1 second less than 80%, and lobectomy; while treatment options of PAAL include watchful waiting with continuous drainage through a thoracostomy tube, pleurodesis, surgical procedures, as well as the use of endoscopic techniques. The use of these different treatment options could be monitored for efficacy using the TM&M classification system. Another frequent complication

(overall second most common AE after pulmonary resection) for pulmonary diseases is atrial fibrillation. Like PAAL, post-operative atrial fibrillation (PAF) has a significant clinical impact on both patients and resources. Using the same methodology as outlined above, the incidence, severity, and risk factors for PAF were identified in **Chapter II**. Results indicate that one-third of PAF cases lead to persistence or major intervention; while increasing age, coronary artery disease, and extent of surgery further exacerbate the risk of PAF following pulmonary resection. Identifying patients with elevated risk for both complications leads to targeted prophylaxes to reduce their incidence. These targeted prophylaxes are the subject of future research.

The use of standardized, valid and reliable definitions is fundamental to the accurate measurement and monitoring of surgical AEs. The American College of Surgeons-National Surgical Quality Improvement Program (ACS-NSQIP) has been considered as the most prominent surgical quality improvement effort for standardizing and comparing AEs after surgery. However, the accuracy of AE reporting specific to thoracic surgical specialty has never been examined in depth. In **Chapter III**, we explored the inter-system reliability of reported AEs following thoracic surgery from the ACS-NSQIP database and the TM&M classification system in order to better understand to what extent the methods used to collect data may be impacting study results. Our results demonstrated that the method of data collection can significantly alter the reported rate of post-operative AEs. For example, although overall rates of post-operative AEs were similar between the two systems, rates of reported cardiac complications were higher in TM&M vs. ACS-NSQIP (5.8% vs. 1.1%; $p<0.01$), and wound complications were lower (2.5% vs 6.0%; $p<0.01$). We acknowledge that both systems have strengths and limitations, and offer complementary value. Quality improvement systems built on partial and retrospective

sampling, such as ACS-NSQIP, may be useful for global inter-institutional analyses and benchmarking, but are of limited use for providing a foundation for meaningful quality improvement initiatives when analyzing data for specific thoracic surgical operations and related post-operative AEs within a single surgical division. Across-the-board standardization of morbidity definitions between the two systems may further help to facilitate future quality improvement endeavors and generate significant new knowledge.

Emphasis on quality and patient safety has generated significant interest in performance measurement of individual surgeons. However, literature on individual surgeon performance remains scarce, with an absence of qualitative evidence on the credibility of surgeon-specific outcome reporting and its impact on the quality of surgical care. Using the TM&M classification to document all AEs after every thoracic surgical procedure, we created risk-adjusted surgeon-specific outcome reports (SSORs) to enable individualized performance measurement and feedback.

A priority for the division has been to ensure such measurement translates into tangible and reproducible improvements in surgical performance, for the benefit of surgeons, and ultimately patients. To do so, we implemented quarterly continuous quality improvement (CQI) seminars to provide an additional forum for discussion regarding collective results, utilizing the technique of positive deviance, to unmask best performers as a catalyst for discussing practice measures to improve specific AEs.

In **Chapter IV**, we identified the benefits and limitations surrounding SSORs and a complementary program of CQI through structured in-person interviews. Interviews revealed enthusiastic support for both SSORs and a complementary program of CQI as a means to improve patient care through awareness of personal outcomes with blinded divisional comparison for similar operations and diseases, with application to continuous professional development and maintenance of certification. Several concerns do persist, however, regarding the use of SSORs, including resistance to change and the many ways information may be misinterpreted and misused. Whether SSORs and CQI can lead to improvements in rates of post-operative AEs is a matter of ongoing research.

Lastly, an evolutionary understanding of the heterogeneity of TM&M data is a critical step to following improvements in care, and to developing novel and successful approaches to quality and patient safety research. Recognizing that software was necessary to efficiently record and review TM&M data, as well as to report the data back to surgeons, iterative development led to an evolution of a real-time, point-of-care, iPad-optimized Thoracic Surgery Quality monitoring, Information management, and Clinical documentation (TSQIC) software system. The TSQIC system enables bedside clinical data recording (including clinical data, operative data, and post-operative TM&M), data storage, and automated dynamic analysis and reporting of surgical volume and quality of care, with the option of using the system for clinical documentation of the clinical encounter (replacing need for voice dictation).

Chapter V summarises our division's progress to date, including both successes and challenges with the use of the TM&M classification system and its integration into a software application,

and lists our recommendations for the next steps that will shape the future of patient safety and quality improvement in surgery. The evidence relating to the use of this program to date: confirms ease of use for both surgical fellows and staff surgeons; documents all post-operative AEs following all surgeries; supports collection of surgeon-specific performance data for self-assessment and benchmarking purposes; and suggests a positive cultural change of the opinions of surgeons to data collection and outcomes monitoring. It is possible that adding other interventions to the measurement and monitoring of post-operative AEs, such as structured morbidity and mortality conferences, and efforts to identify and disseminate best practices, will lead to improved patient outcomes.

Future quality improvement directions

Citations of the TM&M classification system since its introduction in the literature in 2010 have steadily increased. All articles have originated from the thoracic surgical specialty and recommended the routine use of the TM&M classification system to report outcomes.

The Ottawa Hospital's Department of Surgery is poised to expand the breadth and scope of the TM&M program of research in surgical quality and performance improvement. Thus, in keeping with hospital's emphasis on quality and safety of patient-care, the department has announced the creation of a Departmental Quality and Patient Safety (QPS) program. The vision of this forthcoming program is that the rigorous analytical components of TM&M will be transposed onto other surgical disciplines throughout the Department of Surgery at The Ottawa Hospital; allowing more accurate tracking and reporting of AEs, analysis of risk factors underlying these events; and enhancing and standardizing divisional morbidity & mortality (M&M) conferences.

Finally, the aggregate data will allow creation of surgeon-specific outcome reports, providing us feedback on key quality metrics, and promoting dialogue and sharing of best-practices facilitated via the positive deviance approach. This Departmental QPS program will interface closely with the already-existing National Surgical Quality Improvement Program (NSQIP) and Comprehensive Unit-Based Safety Program (CUSP) efforts at The Ottawa Hospital. If this system of surgeon-led, divisionally-focused, recording and reporting of all AEs after all surgeries is scalable to other divisions within The Ottawa Hospital, there is widespread potential for quality improvement.

Future research directions

While the use of a reliable and continuous system of evaluation of presence and severity of AEs after surgery is necessary, it is not sufficient for a comprehensive evaluation of surgical quality. Future research will focus on the impact of post-operative AEs and their severity on patient experience with surgical care, as well as, the impact of post-operative AEs on long-term oncologic outcomes of surgical care.

Second, quantitative estimates of how anesthesia management impacts perioperative AEs are limited. Future studies will link data from the post-anaesthesia care unit with the TM&M database to evaluate the impact of multidisciplinary strategies on surgical outcomes with the intent to perform innovative individualized inter-provider feedback, coupled with positive deviance workshops to help improve care.

Last, but not least, we plan to objectively evaluate the conceptualization, implementation, and reception of our quality improvement enterprise in the Department of Surgery at The Ottawa Hospital. It is possible that a number of valuable lessons can be learned from the observed successes within the Division of Thoracic Surgery to help determine what the generalizable factors might be that can be adapted when engineering cultural change in the other divisions or units of care at The Ottawa Hospital.

ACKNOWLEDGEMENTS

First and foremost, I offer my enduring gratitude to my fellow classmates, administrative staff (Fay Draper and Ginette Robitaille), and faculty members (Dr. Beth Potter, Dr. Yue Chen, Dr. Doug Coyle, Dr. Julien Little, and Dr. Brenda Wilson), in the Department of Epidemiology and Community Medicine, at the University of Ottawa.

I would like to thank my colleagues at the Ottawa Hospital Research Institute (OHRI) and The Ottawa Hospital's (TOH) Cancer Assessment Clinic (CAC), including: Caitlin Anstee, Jennifer Molson, Salome Shin, and Jennifer Smylie; and the administrative staff of TOH's Division of Thoracic Surgery, including: Nathalie Bedard, Shantel Mudge, Kelly White, and Jo-Anne Villeneuve. You have contributed immensely to my professional time at OHRI, CAC, and TOH by being a strong source of collaboration, advice, and support.

I would like to acknowledge the dedication of past and current thoracic surgical fellows for ensuring efficient and consistent adverse event reporting, and for their incessant commitment to embracing a culture of quality and patient safety, including: Dr. Derar Alshahab, Dr. Ahmed Alhussaini, Dr. Saleh Abu-Daff, Dr. Anas Alshuhayeb, Dr. Anna McGuire, Dr. Ramzi Addas, and Dr. Daniel French.

Dr. Donna Maziak, Dr. James Villeneuve, Dr. Sebastien Gilbert, Dr. Farid Shamji, and Dr. Sudhir Sundaresan of TOH's Division of Thoracic Surgery, thank you for your guidance, and feedback, and for bearing with me the brunt of thesis frustrations, and likewise, sharing in the joy of the successes. These successes reflect all of the Division's contributions and commitment to this exciting quality improvement research.

I would like to express my sincerest appreciation and gratitude to my thesis advisors. Dr. Tim Ramsay, thank you for your dedication, commitment, and your brilliant comments and suggestions throughout my doctoral thesis work. Dr. Andrew Seely, you have been a tremendous mentor for me; your joy, enthusiasm, and passion for research were, and continue to be, contagious and inspiring. I appreciate all your contributions of time and ideas, to make my doctoral experience productive and stimulating.

Grandma, thank you for your endless supplies of hummus, labneh, toum, tabbouleh, fattoush, wara'enab, falafel, kibbeh, shish taouk, mulukhiyah, bamieh, manaeesh, za'atar, zeitoun, ahweh, ka'ak, and ma'amoul, all served with a side order of compassion and love. This thesis could not have been possible without you.

A special thank you is owned to my mother, my father, and my brother, who have been a tremendous source of love, support, and inspiration throughout the entirety of my education.

Last, Alexy Hajjar – my husband, my love, my honorary doctor - thank you for making this entire journey worthwhile and possible.