Evaluating the Reliability and Reproducibility of the Ottawa Thoracic Morbidity and Mortality Classification System

Jelena Ivanovic, BS, Ahmed Al-Hussaini, MD, Derar Al-Shehab, MD, Jennifer Threader, BS, Patrick James Villeneuve, MDCM, PhD, Tim Ramsay, PhD, Donna E. Maziak, MDCM, MS, Sebastian Gilbert, MD, FRCSC, Farid M. Shamji, MBBS, FRCSC, R. Sudhir Sundaresan, MD, FRCSC, and Andrew J.E. Seely, MD, PhD

Division of Thoracic Surgery, Department of Surgery, University of Ottawa, the Ottawa Hospital–General Campus, and Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

Background. **Minimizing adverse events after surgery is widely recognized as an important indicator of quality; yet no consensus has been reached on how to standardize the reporting of adverse events after surgical procedures. Our objectives were to develop a standardized classification system to monitor both the presence and severity of thoracic morbidity and mortality, and to evaluate its reliability and reproducibility among a national cohort of thoracic surgeons.**

Methods. **To assess the Thoracic Morbidity and Mortality classification system (based on the Clavien-Dindo classification of adverse events), a 31-item questionnaire was sent to all members of the Canadian Association of Thoracic Surgeons in August 2009, consisting of a general description of the Thoracic Morbidity and Mortality severity grades, 20 case-based questions of postoperative adverse events to be classified, and questions regarding personal judgments. We derived descriptive and quantitative information using weighted Kappa statistics.**

Results. **Fifty-two (54.7%) thoracic surgeons completed the questionnaire; 41 (78.8%) of the respondents were affiliated with an academic teaching hospital. A total of 1,326 individual weighted Kappa statistics were calculated for all distinct pairs of raters, of which 1,152 (87%) were greater than 0.81, a range that is interpreted as "almost perfect agreement." A further 174 (13%) were in the range between 0.61 and 0.8, interpreted as "substantial agreement." All results were statistically significant ($p < 0.0001$). The classification system was regarded as straightforward (98% of the respondents), reproducible (94%), logical (92%), and useful (98%).**

Conclusions. **The modified classification system appears to offer objective, reliable, and reproducible reporting of thoracic morbidity and mortality, and thus may assist continuous quality improvement in thoracic surgery.**

(Ann Thorac Surg 2011;91:387–93)

Morbidity and mortality (M&M) rates are frequently used as indicators of quality in surgical care [1]. Thus, accurate and objective reporting of postoperative M&M is essential to implement and follow improvements in quality of surgical care [2]. Reporting of postoperative adverse events has traditionally been accomplished at M&M conferences and through retrospective case series within the surgical literature. Unfortunately, these approaches are susceptible to selection bias and have resulted in underreporting [3]. To date, no consensus has been reached among surgeons on how to quantify presence or severity of postoperative adverse events.

In 1992, Clavien and colleagues [4] were the first to introduce the idea of severity grading of surgical adverse events as a way of assessing the degree of injury caused by that event. This classification system was modified in 2004 to increase its accuracy and acceptability in surgical practice and renamed as the Clavien-Dindo classification system [5]. The principle underlying the Clavien-Dindo classification system assumes the severity of an adverse event is proportional to its impact on a patient and the degree of effort to correct it. Since its modification, the Clavien-Dindo classification system has been avidly used as a tool for quality assessment in surgery, and has broad applicability in clinical practice [6–10]. To date, however, no attempts have been made to apply the approach of standardizing surgical adverse events after thoracic surgery.

Therefore, the goals of the current study were to first develop a classification of surgical adverse events based on our experience of thoracic morbidity and mortality (TM&M) and to assess the acceptability of the classification through peer review and questionnaire in a single academic thoracic surgery center (The Ottawa Hospital); and second, to assess the reproducibility and reliability of

Accepted for publication Oct 18, 2010.

Address correspondence to Dr Seely, Ottawa Hospital–General Campus, 501 Smyth Rd, Box 708, Ottawa, Ontario, K1H 8L6; e-mail: aseely@ottawahospital.on.ca.

Published by Elsevier Inc

0003-4975/$36.00
doi:10.1016/j.athoracsur.2010.10.035

Table 1. Classification of Complications After Thoracic Surgery

Grade	Definition
Complication	Any deviation from the normal postoperative course.
Minor	
Grade I	Any complication without need for pharmacologic treatment or other intervention.
Grade II	Any complication that requires pharmacologic treatment or minor intervention only.
Major	
Grade III	Any complication that requires surgical, radiologic, or endoscopic intervention.
Grade IIIa	Intervention does not require general anesthesia.
Grade IIIb	Intervention requires general anesthesia.
Grade IV	Any complication requiring ICU management or life support.
Grade IVa	Any complication leading to single-organ dysfunction.
Grade IVb	Any complication leading to multiorgan dysfunction.
Mortality	
Grade V	Any complication leading to the death of the patient.

ICU = intensive care unit.

the classification through a survey of all active members of the Canadian Association of Thoracic Surgeons (CATS). We hypothesize that a graded classification system for evaluating M&M after thoracic surgery is an objective, reproducible, and practical tool to monitor and analyze adverse events.

Material and Methods

Single-Center (Internal) Development of the Thoracic Morbidity and Mortality Classification System

The TM&M classification system was developed in a single academic thoracic surgery center, the Ottawa Hospital, according to the Clavien-Dindo classification proposed in 2004, which grades an adverse event on a severity scale from grades I to V based on the effort required to treat the event (Table 1). Grades I and II include events that deviate from the normal postoperative course but require either no intervention or pharmacologic therapy, respectively. A grade III event requires medical intervention, without general anesthesia (IIIa), and with general anesthesia (IIIb). Grade IV events are life-threatening and require intensive care unit management owing to single-organ dysfunction (IVa) or multiorgan dysfunction (IVb). Grade V events result in death of the patient. For each of the following systems—pulmonary, pleural, cardiac, renal, gastrointestinal, neurologic, and wound—adverse events were defined associated with a specific grade (Table 2). The Common Terminology Criteria for Adverse Events (CTCAE version 3.0) [11] was also used to refine a number of definitions. Specific definitions were further refined by peer review and questionnaire, and modified according to potential adverse events in patients after thoracic surgery. This was an iterative process that required consensus among the 5 practicing thoracic surgeons and the 2 thoracic surgical residents, who form a large thoracic oncology program in Ontario.

Multicenter (External) Testing of Reproducibility and Reliability of the Thoracic Morbidity and Mortality Classification System: Survey of the Members of the Canadian Association of Thoracic Surgeons

The CATS was approached for a multicenter evaluation of the TM&M classification system. The membership of CATS includes full-time practitioners of general (noncardiac) thoracic surgery, along with qualified general and cardiovascular surgeons whose practice includes more than 50% thoracic surgery [12].

To assess the reproducibility and reliability of the modified classification, an electronic questionnaire was designed with 31 items (refer to online appendix). The CATS master file, provided by the executive committee, was used in developing a mailing list of the target surgeons. This list was based on data collected from individual members and included all thoracic surgeons with a valid e-mail address. An initial e-mail was sent with a link to the survey at the start of August 2009, and three reminder e-mails were sent each week thereafter. Eligible candidates who did not have valid e-mail addresses were sent a questionnaire package by postal service. The questionnaire was estimated to take less than 15 minutes to complete. The questionnaire was voluntary and elicited opinions and did not qualify as human subject research. The Ottawa Hospital Research Ethics Board approved this study. As a token of gratitude, each survey respondent received a $10 gift certificate to any Tim Horton's coffee shop across Canada. The questionnaire consisted of three parts including (1) an information sheet with the TM&M classification system along with definitions of the severity grades; (2) 20 case-based questions asking respondents to classify postoperative adverse events in accordance to the proposed classification system (Table 3 shows several case examples placed in order from lowest to highest grade of severity); and (3) questions regarding personal judgments about the classification system. In the second part of the questionnaire, the 20 case-based scenarios were placed randomly with regard to their complication grade. The 20 case scenarios were chosen to have an even representation of minor (grades I and II) and major case examples (grades IIIa to V). Respondents were asked to choose the most severe grade of complication for each case (one choice per case). A pilot study was performed using the survey with 5 thoracic surgeons and 2 thoracic surgical residents at the Ottawa Hospital before mailing. At the time the survey was conducted, CATS had a total of 95 members.

Table 2. Clinical Examples of Postoperative Adverse Event Based on Severity Grades

Grade	Organ System	Examples
I	Pulmonary	Atelectasis: requiring no intervention other than additional chest physiotherapy
	Pleural	Effusion: asymptomatic, no intervention indicated
	Anastomotic	Anastomotic leak: transient, no therapy added
	Cardiac	Atrial fibrillation: converting after correction of electrolytes
	Renal	UTI: no intervention required other than removal of Foley
	Gastrointestinal	Constipation: no intervention required
	Neurologic	Confusion: transient, new or worsened; no intervention required
	Wound	Hematoma: no intervention required
II	Pulmonary	Atelectasis: requiring endotracheal suctioning
	Pleural	Effusion: requiring medical therapy (eg, diuretics) for CHF and drainage
	Anastomotic	Anastomotic leak: medical therapy (eg, antibiotics) added
	Cardiac	Atrial fibrillation: requiring medications (eg, beta-blockers) for heart rate control
	Renal	UTI: requiring medical therapy only (eg, antibiotics)
	Gastrointestinal	Constipation: nasogastric intubation, stool softeners, laxatives, dietary modification, or enema
	Neurologic	Confusion: requiring medical intervention
	Wound	Hematoma: transfusion, evacuation or aspiration; opening of wound at bedside
IIIa	Pulmonary	Atelectasis: endoscopic or radiologic intervention, or noninvasive ventilation for <24 h
	Pleural	Effusion: endoscopic, radiologic, or bedside pleural interventions performed
	Anastomotic	Anastomotic leak: intervention required (opening of wound)
	Cardiac	Atrial fibrillation: symptomatic, requiring a cardioversion or a device (eg, pacemaker)
	Renal	UTI: endoscopic, radiologic, or bedside interventions performed
	Gastrointestinal	Constipation: obstipation with manual evacuation indicated
	Neurologic	Confusion: prolonged hospitalization indicated; danger to self or others
	Wound	Hematoma: interventional radiology indicated
IIIb	Pulmonary	Atelectasis: operative intervention required under general anesthesia
	Pleural	Effusion: surgical intervention performed under general anesthesia
	Anastomotic	Anastomotic leak: intervention required under general anesthesia
	Cardiac	Atrial fibrillation: intervention required under general anesthesia
	Renal	UTI: surgical intervention performed
	Gastrointestinal	Constipation: surgical intervention performed under general anesthesia
	Neurologic	Confusion: N/A
	Wound	Hematoma: surgical intervention performed
IVa	Pulmonary	Atelectasis: respiratory compromise requiring intubation and positive-pressure ventilation
	Pleural	Effusion: life-threatening; debilitating; organ failure present
	Anastomotic	Anastomotic leak: leads to single-organ failure
	Cardiac	Atrial fibrillation: single-organ failure (eg, CHF, hypotension, syncope, shock)
	Renal	UTI: life-threatening; debilitating; organ failure present
	Gastrointestinal	Constipation: life-threatening consequences (eg, obstruction, toxic megacolon)
	Neurologic	Confusion: N/A
	Wound	Hematoma: life-threatening consequences; major urgent intervention indicated
IVb	Pulmonary	Atelectasis: concomitant multiorgan complications
	Pleural	Effusion: concomitant multiorgan complications
	Anastomotic	Anastomotic leak: leads to multiorgan failure
	Cardiac	Atrial fibrillation: concomitant multiorgan complications
	Renal	UTI: concomitant multiorgan complications
	Gastrointestinal	Constipation: concomitant multiorgan complications
	Neurologic	Confusion: N/A
	Wound	Hematoma: concomitant multiorgan complications

CHF = congestive heart failure; N/A = not applicable; UTI = urinary tract infection.

Statistical Analysis

All survey data collected were carefully entered using quality control and verification measures into a secure database. Replies were kept anonymous. Weighted kappa statistics were calculated to assess the interrater reliability among the survey respondents. For each inter-

Table 3. Examples of Clinical Cases and Their Respective Severity Grades

Grade	Clinical Case
I	A RUL lobectomy was performed on a 54-year-old patient for lung cancer; the patient exhibited a prolonged air leak but required no further intervention beside observation.
II	A 47-year-old patient experienced atrial fibrillation on the second day after pneumonectomy for malignant mesothelioma. IV metoprolol and digoxin were given; the patient reverted to sinus rhythm.
IIIa	A 22-year-old patient underwent video-assisted thoracoscopic surgery bullectomy and pleurectomy for recurrent spontaneous pneumothorax. The experienced prolonged air leak and persistent pneumothorax, which was treated by insertion of another chest tube.
IIIb	An 84-year-old patient underwent total thyroidectomy for a retrosternal goiter. The patient exhibited stridor and hypoxia. An immediate operation was done to evacuate the hematoma and control the bleeding.
IVa	A 69-year-old patient experienced ARDS after right lower lobectomy for NSCLC. The patient was transferred to ICU and required intubation for 2 weeks; eventually the patient recovered and was extubated.
IVb	An esophagectomy and gastric pull up was performed on a 50-year-old patient. On day 8, the patient showed signs and symptoms of sepsis and anastomotic leak. The patient was taken to the OR for drainage and lavage. Thereafter, the patient was transferred to the ICU, where he exhibited ARDS and acute renal failure that required dialysis. Eventually the patient recovered.
V	A 62-year-old patient underwent a redo fundoplication for recurrent hiatal hernia. The patient exhibited a massive pulmonary embolism, which was treated with tPA. The patient was transferred to the ICU where he died after a cardiac arrest.

ARDS = acute respiratory distress syndrome; ICU = intensive care unit; IV = intravenous; NSCLC = non-small cell lung cancer; OR = operating room; RUL = right upper lobe; tPA = tissue plasminogen activator.

rater comparison, calculations were performed to establish the proportion of observed agreement (Po), the proportion of expected agreement (Pe), and the kappa value. A distribution of kappa scores was calculated for the five grades. This score yielded a weighted calculation of all of the paired kappa scores for each case scenario and was used to indicate general agreement among the surgeons. The same calculation was applied for the minor and major case scenarios. The level of agreement among the raters was evaluated using the system put forth by Landis and Koch in 1977 [13], in which a kappa value of 0.21 to 0.4 reflects fair agreement, a value of 0.41 to 0.60 reflects moderate agreement, a value of 0.61 to 0.80 reflects substantial agreement, and a value of 0.81 or more reflects almost perfect agreement. Data were analyzed using R statistical software (Version 2.9.2).

Results

Response Rate

From the 95 members, 52 surveys were completed (54.7%) within the designated timeframe permitted for response, which was 4 weeks. The 52 survey respondents represented 28 Canadian centers. The majority of respondents were affiliated with a university teaching hospital (78.8%, n = 41) and practiced in Ontario (32.7%, n = 17) or Quebec (15.4%, n = 8). Ontario and Quebec are the most densely populated provinces in Canada with populations of approximately 12.2 million and 7.6 million, respectively. In terms of geographic distribution, the characteristics of the sample were representative of the whole. Of the 52 completed surveys, 8 (15.4%) were completed by members practicing outside of Canada. Most surgeons had been in practice for less than 10 years (51.0%, n = 26). Fifty-one respondents (98.1%) indicated that they were a full-time staff member at their institution; one resident completed the survey.

Interrater Agreement

The weighted kappa statistic assesses agreement between two raters on an ordered scale. With 52 raters, a total of 1,326 individual weighted kappa statistics were calculated for all distinct pairs of raters (Table 4). Of those 1,326 weighted kappa statistics, 1,152 (87.0%) were greater than 0.81, a range that is interpreted as "almost perfect agreement." Furthermore, 173 (13.0%) were in the

Table 4. Inter-Rater Agreement[a]

	Overall Agreement		Minor Complications		Major Complications	
Range	N	%	n	%	n	%
0.81 to 1.0 "almost perfect agreement"	1152	87.0	428	32.2	992	74.8
0.61 to 0.8 "substantial agreement"	173	13.0	455	34.3	265	20.0
0.41 to 0.6 "moderate agreement"	0	0	265	20.0	47	3.5
0.21 to 0.4 "fair agreement"	0	0	178	13.4	22	1.7
Total	1326	100	1326	100	1326	100

[a] All results were statistically significant ($p < 0.0001$).

range between 0.61 and 0.8, interpreted as "substantial agreement." Thus, all of the statistics indicated at least substantial agreement. All results were statistically significant ($p < 0.0001$), indicating that we should reject the null hypothesis that the ratings are independent (ie, kappa = 0) and accept the alternative hypothesis that agreement is better than one would expect by chance alone.

The level of interrater agreement was also calculated among minor and major clinical case examples. A factor that has to be considered in the interpretation of kappa coefficients is the number of categories [14]. For the minor clinical case examples (grades I and II), a total of 428 (32.2%) individual weighted kappa statistics were greater than 0.81. A further 455 (34.3%) individual weighted kappa statistics were in the range between 0.61 and 0.8; and 265 (20.0%) in the range between 0.41 and 0.6, which is a range interpreted as "moderate agreement." The remaining 178 (13.4%) individual weighted kappa statistics were in the range between 0.21 and 0.4, which is a range interpreted as "fair agreement."

A more pronounced increase of kappa coefficients was observed with increasing numbers of categories. Thus, for the clinical case examples with major complications (grades IIIa, IIIb, IVa, IVb, and V), a total of 992 (74.8%) individual weighted kappa statistics were greater than 0.81. A further 265 (20.0%) individual weighted kappa statistics were in the range between 0.61 and 0.8; and 47 (3.5%) in the range between 0.41 and 0.6. The remaining 22 (1.7%) individual weighted kappa statistics were in the range between 0.21 and 0.4.

Personal Judgments

Respondents were asked to agree or disagree with several statements regarding their personal judgments of the TM&M classification system. Of the 52 respondents, 49 (98.0%) considered the TM&M classification system as straightforward to understand. A total of 48 respondents (94.1%) considered the TM&M classification system as reproducible; that is, different surgeons would tend to agree on the classification of individual patient events. A total of 47 respondents (92.2%) considered the TM&M classification system to be logical; that is, it accurately reflects level of severity of adverse events. Lastly, 50 respondents (98.0%) considered the TM&M classification system to be useful in their patients; that is, it will be helpful to evaluate both presence and severity of surgical adverse events.

Comment

In the surgical community, M&M rates have been established as critical outcome measures and indicators of quality [2]. However, conflicting approaches of reporting postoperative adverse events make the use of these rates unreliable for quality assessment. In addition, the absence of standardized definitions and a generally accepted classification scheme to grade surgical adverse events has further hindered appropriate evaluation of surgical outcome data [15]. Specifically, surgeons and medical institutions have inconsistently applied expressions such as minor, moderate, major, or severe to classify surgical adverse events [16]. An objective system for monitoring and accurately reporting postoperative adverse events is fundamental to advance performance in thoracic surgery and collect reliable data for benchmarking.

To evaluate the reproducibility of the TM&M classification system, clinical case examples were created by the thoracic oncology team at the Ottawa Hospital and sent to all members of the CATS. The consistency of a surgeon's rating is an important consideration in outcome assessment. These ratings often fall on an ordinal scale, making the kappa coefficient an appropriate measure of reliability for such data [14]. A high level of agreement was calculated among the 52 survey respondents for the 20 case scenarios, indicating that the TM&M classification system is consistent among surgeons' opinion and can be applied to multifaceted case examples. One explanation for the lower proportion rate of kappa scores among the minor case scenarios (ie, grades I and II) can be attributed to the number of categories. The proportion of kappa statistics that was greater than 0.81 was lower when two categories were presented, as agreement by chance was more likely among the raters. A more pronounced increase of kappa coefficients was observed among the major case scenarios (ie, grades IIIa, IIIb, IVa, IVb, and V) because of increasing numbers of categories. As the number of categories increased, the likelihood of agreement by chance was reduced among the raters. The increase of kappa coefficients with the number of categories is a preferred outcome, because as the number of categories increases, so does the proportion of the variability in the true variable captured by the imperfect ordinal variable [17]. Through the application of severity grades, the TM&M classification system has provided standardized measures for discriminating what may represent a minor as opposed to major adverse event after thoracic surgery.

The presented results demonstrate that the TM&M classification system can be used in its current state in clinical research and for quality improvement. Since initiating the TM&M classification system in January 2008 at the Ottawa Hospital, daily data collection of M&M has been carried out by a senior thoracic surgical resident and the thoracic surgery research coordinator using the TM&M form (refer to online appendix). Weekly lists of operative procedures along with related adverse events are compiled and further validated by all attending staff present. Complications are then presented and discussed at monthly M&M conferences—allowing for regular and active reporting of adverse events. A secure database for adverse event reporting was developed and has since been used to compare surgical procedures and subgroups of patients, allowing us to evaluate the feasibility of the system during the first 2 years of its implementation at the Ottawa Hospital [18]. The TM&M classification system advocates for a practice of continuous quality improvement, advances the development of quality improvement programs, can be used to improve the

quality of retrospective studies (as it does not rely on the original surgeon to grade the adverse event as such), and facilitates an open forum for ongoing medical education on surgical quality assurance. Moreover, our 2-year experience indicates that the TM&M classification system is feasible, facilitates objective comparison, accurately identifies burden of illness of individual adverse events, and provides an effective method for continuous surgical quality assessment [18].

There are many other potential applications of this system. For example, video-assisted thoracoscopic surgery is a relatively new technology that has become the standard of care for minimally invasive thoracic procedures. However, controversy surrounds the safety, reproducibility, and oncologic efficacy of video-assisted thoracic surgery for lobectomy [19]. Awareness of this controversy inspired us to initiate a comparison between video-assisted thoracoscopic surgery and open lobectomy procedures at our institution. The TM&M classification system was used for reporting the difference in presence, severity, and types of postoperative morbidity in patients undergoing video-assisted thoracoscopic surgery versus open lobectomy [20].

We further plan to use this continuous TM&M classification and reporting system as a backbone for prospective monitoring of essential surgical information, upon which to add additional clinical data collection tools. The TM&M classification system provides a strong base with which we can build a system to continuously monitor and improve the overall quality of thoracic surgical care. Expanding the TM&M classification system to include clinical data on all time points on the continuum of care, starting with patient referral to at least a 2-year follow-up after surgery would help improve continuous assurance of care. A reduction in adverse events and death can be achieved by the exchange of information and analysis of M&M among hospitals [21]. Additionally, standardized collection and electronic storage of patient information can provide a data set to enable prospective clinical research [22].

Other future refinements to the TM&M classification system are planned, including a measure of the etiology of surgical adverse events and determination of whether the event was preventable. Additionally, patients with prolonged hospital stay and those who are readmitted to the hospital form a small proportion of thoracic patients at our institution, and may highlight quality of care problems. To monitor and further our understanding of those patients who are readmitted to the hospital, the TM&M classification system can be made capable of linking hospital readmissions and, eventually, hospital readmissions with out-of-hospital services as well [23].

We recognize important limitations of the use of a survey instrument for evaluating the reproducibility of the proposed system. Our overall response rate was 54.7%. As such, it is possible that our group of respondents was not representative of thoracic surgeons as a whole. Although the use of a reliable and continuous system of evaluation of the presence and severity of adverse events after thoracic surgery is necessary, it is not sufficient for a comprehensive evaluation of surgical quality. Monitoring of wait times, efficient resource utilization, and patient experience and satisfaction are all dimensions of surgical care quality improvement.

We conclude that the TM&M classification system appears reliable and reproducible and may represent a feasible tool for quality evaluation in surgery. The objective evaluation of both the presence and severity of TM&M and the prospective monitoring of thoracic surgical volume represents an important means of standardizing evaluation of outcomes, enabling comparison among centers and surgeons, and represents a crucial component to ensuring best practice of care. We hope that utilization of this system in future studies can enable improvements in thoracic surgical quality.

Web Archive

The Ottawa TM&M system is freely available online and can be found at the following Web site: https://sites.google.com/site/ottawatmmtool/home. This online appendix includes the definitions of all the complications according to grade and to the affected system. A copy of the survey can be found as well.

The authors would like to gratefully acknowledge the entire membership of the Canadian Association of Thoracic Surgeons for their participation in the questionnaire, in addition to their feedback and encouragement of this project.

References

1. Dimick JB, Weeks WB, Karia RJ, et al. Who pays for poor surgical quality? Building a business case for quality improvement. J Am Coll Surg 2006;202:933–7.
2. Veen EJ, Steenbruggen J, Roukema JA. Classifying surgical complications: a critical appraisal. Arch Surg 2005;140:1078–83.
3. Grobmyer SR, Pieracci FM, Allen PJ, et al. Defining morbidity after pancreaticoduodenectomy: use of a prospective complication grading system. J Am Coll Surg 2007;204:356–64.
4. Clavien PA, Sanabria JR, Strasberg SM. Proposed classification of complications of surgery with examples of utility in cholecystectomy. Surgery 1992;111:518–26.
5. 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.
6. 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;244:931–9.
7. Gonzalgo ML, Pavlovich CP, Trock BJ, et al. Classification and trends of perioperative morbidities following laparoscopic radical prostatectomy. J Urol 2005;174:135–9.
8. Kocak B, Koffron AJ, Baker TB, et al. Proposed classification of complications after live donor nephrectomy. Urology 2006;67:927–31.
9. Tamura S, Sugawara Y, Kaneko J, et al. Systematic grading of surgical complications in live liver donors according to Clavien's system. Transplant Int 2006;19:982–7.
10. Mazeh H, Samet Y, Abu-Wasel B, et al. Application of a novel severity grading system for surgical complications after colorectal resection. J Am Coll Surg 2009;208:355–61.

11. 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.
12. Darling GE, Maziak DE, Clifton JC, et al. The practice of thoracic surgery in Canada. Can J Surg 2004;47:438–45.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
14. Brenner H, Kliebsch U. Dependence of weighted kappa coefficients on the number of categories. Epidemiology 1996; 7:199–202.
15. Martin RC 2nd, Brennan MF, Jaques DP. Quality of complication reporting in the surgical literature. Ann Surg 2002;235:803–13.
16. Clavien PA, Barkun J, DeOliveira ML, et al. The Clavien-Dindo classification of surgical complications: five-year experience. Ann Surg 2009;250:187–96.
17. Sim J, and Wright CC. The Kappa statistic in reliability studies: Use, interpretation, and sample size requirements. Phys Ther 2005;85:257–68.
18. Seely AJE, Ivanovic J, Threader J, et al. Systematic classification of morbidity and mortality after thoracic surgery. Ann Thorac Surg 2010;90:936–42.
19. Gharagozloo F, Tempesta B, Margolis M, et al. Video-assisted thoracic surgery lobectomy for stage I lung cancer. Ann Thorac Surg 2003;76:1009–15.
20. Villeneuve PJ, Alhussaini A, Al-Shehab D, et al. Video assisted thoracoscopic lobectomy: short-term outcomes and safety. 13th Annual Meeting of the Canadian Association of Thoracic Surgeons, Quebec City, September 2010.
21. Campbell DA Jr, Dellinger EP. Multihospital collaborations for surgical quality improvement. JAMA 2009;302:1584–5.
22. Mittnacht AJ, Wax DB, Srivastava S, et al. Development and implementation of a pediatric cardiac anesthesia/intensive care database. Semin Cardiothorac Vasc Anesth 2008;12: 12–7.
23. Corrigan JM, Kazandjian VA. Characteristics of multiple admissions. Am Med Rec Assoc 1991;62:37–47.

Notice From the American Board of Thoracic Surgery Regarding Trainees and Candidates for Certification Who Are Called to Military Service Related to the War on Terrorism

The Board appreciates the concern of those who have received emergency calls to military service. They may be assured that the Board will exercise the same sympathetic consideration as was given to candidates in recognition of their special contributions to their country during the Vietnam conflict and the Persian Gulf conflict with regard to applications, examinations, and interruption of training. If you have any questions about how this might affect you, please call the Board office at (312) 202-5900.

Valerie W. Rusch, MD
Chair
The American Board of Thoracic Surgery