Physician Estimates of Perioperative Cardiac Risk in Patients Undergoing Noncardiac Surgery

P. J. Devereaux, MD; William A. Ghali, MD, MPH; Neil E. Gibson, MD, MSc; Neil M. Skjodt, MD; David C. Ford, MD, MSc; Hude Quan, MD, PhD; Gordon H. Guyatt, MD, MSc

Background: We know little about how physicians assess perioperative cardiac risk in patients undergoing noncardiac surgery.

Objectives: To evaluate preoperative medical consultations and determine the extent to which consultants used validated cardiac risk indices and specialized noninvasive cardiac tests, and to assess agreement between physician ratings of cardiac risk (low, moderate, or high) and risk estimates derived using validated cardiac risk indices or, in the case of vascular surgery, a risk index.

Methods: This observational study was conducted at 5 Canadian teaching hospitals affiliated with 2 universities. We retrospectively evaluated 308 preoperative consultations performed in 297 patients and examined the frequency with which consultants recorded the use of validated cardiac risk indices. We used $\kappa$ statistics to quantify the extent to which physician ratings of cardiac risk agreed with risk estimates derived using validated cardiac risk indices.

Results: Physicians recorded use of a risk index in 31% of the consultations, but the index used was almost always the suboptimal classification of the American Society of Anesthesiologists. The agreement between physician estimates of cardiac risk and the validated cardiac risk indices was only fair, with a weighted $\kappa$ of 0.38 (95% confidence interval, 0.28-0.49). Overestimation and underestimation of cardiac risk occurred in 16% and 13% of the consultations, respectively. Consultants did not order dipyridamole thallium imaging or dobutamine stress echocardiography for any moderate-risk patients undergoing vascular surgery.

Conclusions: Physicians underuse validated cardiac risk indices, and the agreement between the cardiac risk estimates and risk as determined by validated cardiac indices is suboptimal. Physicians are also underusing dipyridamole thallium imaging and dobutamine stress echocardiography for moderate-risk patients undergoing vascular surgery.

Arch Intern Med. 1999;159:713-717

In 1990, it was estimated that 1 million of the 27 million patients undergoing noncardiac surgery in the United States would suffer a perioperative cardiac complication.1 As a result of this risk, surgeons frequently request medical evaluations to assess perioperative cardiac risk. Preoperative medical consultations are so common that members of the American College of Physicians (ACP) rated perioperative cardiac risk assessment as a top-10 topic needing guideline development.2 Recently, the ACP provided guidelines for assessing and managing the perioperative risk from coronary artery disease associated with major noncardiac surgery.3,4

Currently, physicians can assess perioperative cardiac risk using cardiac risk indices, noninvasive cardiac tests, and invasive cardiac tests. Of the cardiac risk indices, the subjective, nonquantitative American Society of Anesthesiologists (ASA) perioperative risk scoring system has proved inferior to other cardiac indices, especially for predicting major postoperative cardiac complications.5,6 The indices of Goldman et al7 (herinafter, Goldman) and Detsky et al8 (herinafter, Detsky) are 2 validated and prospectively studied clinical indices that reliably identify high-risk patients. Of these 2 indices, the Detsky index has a higher area under the receiver operating characteristic curve and uses a Bayesian model.7,8 A new validated and prospectively studied Bayesian index recently described by L’Italien et al9 (hereinafter, L’Italien) estimates perioperative cardiac risk in patients undergoing vascular surgery. Since the other indices listed above were developed for general surgical populations and tend to underestimate risk for patients undergoing vascular surgery, the L’Italien index adds important information.10,11 Of the available cardiac tests, there is strong evidence that only dipyridamole thallium imaging and dobutamine stress echocardiography add useful information for risk stratification beyond the clini-
METHODS

STUDY SITES

Five hospitals affiliated with 2 Canadian universities—the University of Calgary in Calgary, Alberta, and McMaster University in Hamilton, Ontario—provided the setting for this study.

SAMPLING SCHEME

Starting on September 1, 1996, we identified potential study patients by consulting each hospital’s operating room schedule for consecutive patients who had surgery conducted, postponed, or canceled. We reviewed patients’ charts and entered those who fulfilled the following entry criteria: (1) 40 years of age or older, (2) anesthesia administered for scheduled noncardiac surgery or surgery canceled or postponed due to a perioperative cardiac event or concern about cardiac status, and (3) a preoperative consultation by a general internist, cardiologist, or intensivist recorded on their hospital chart. This approach to patient selection captured both inpatient and outpatient consultations.

The focus of this study was the consultation process. Accordingly, if a patient underwent preoperative consultations with more than one physician, we evaluated each consultant’s assessment. Consultations with anesthesiologists were excluded from our study because anesthesiologists are frequently consulted preoperatively for noncardiac issues, and we wanted to minimize the number of consultations in which there were no cardiac concerns.

DATA COLLECTION

We reviewed eligible patients’ charts and recorded information pertaining to the perioperative cardiac risk assessment including the following: (1) location of preoperative consultation (ie, inpatient or outpatient, hospital, or university); (2) specialty of consultant; (3) explicit use a cardiac risk index (ie, yes or no and specific index); (4) preoperative ordering of any specialized cardiac test(s) and the results; (5) physician estimates of perioperative cardiac risk (ie, low, moderate, or high), and (6) chart-derived estimates of cardiac risk using the Detsky and L’Italien risk indices (see below).

Four reviewers performed chart abstraction, 2 in Calgary (P.J.D. and N.M.S.) and 2 in Hamilton (N.E.G. and D.C.F.). Initially, reviews examined charts in duplicate to ensure interobserver reliability of the chart abstraction process. When reviewers demonstrated acceptable interobserver reliability, each subsequent chart was evaluated by a single reviewer, alternating between all 4 reviewers listed above.

The interobserver $\kappa$ statistics were consistent with substantial or near-perfect agreement and were similar in Calgary and Hamilton. The interobserver $\kappa$ statistics for the chart reviewers in Calgary were as follows: 1.00 for physicians’ ordering of specialized cardiac tests; 0.72 for physicians’ cardiac risk estimates (low, moderate, and high); and 0.69 for reviewers’ Detsky and L’Italien risk group calculations. The interobserver $\kappa$ statistics for the chart reviewers in Hamilton were as follows: 1.00 for physicians’ explicit use of a cardiac risk index; 1.00 for physicians’ ordering of specialized cardiac tests; 0.78 for physicians’ cardiac risk estimates (low, moderate, and high); and 0.65 for reviewers’ Detsky and L’Italien risk group calculations.

DATA ANALYSIS

The main analysis compared agreement between physician estimates of perioperative cardiac risk (low, moderate, and high) and cardiac risk estimates as determined by validated cardiac risk indices. We calculated both simple agreement and the $\kappa$ statistic (which quantifies agreement over and above that expected by chance) including the 95% confidence intervals.

RESULTS

CONSULTATION AND PATIENT INFORMATION

We evaluated 308 preoperative consultations in 297 patients. There were 49 medical consultants (in the 5 hospitals studied), of whom 30 (61%) were general internists, 14 (29%) cardiologists, and 5 (10%) intensivists. Of the 308 consultations, 92% were performed by general internists, 6% by cardiologists, and 2% by intensivists. When general internists assessed patients preoperatively, they requested a cardiology consultation 2% of the time.

Patient preoperative evaluations occurred in the outpatient setting 73% of the time and in the inpatient setting 27% of the time. The distribution of planned surgical procedures was 59% orthopedic surgery, 16% head and neck or minor surgery, 15% intrathoracic or intraperitoneal surgery, and 10% vascular surgery.

USE OF INDICES

Of the 308 consultations, 31% explicitly mentioned use of a perioperative cardiac index. When physicians used an index, it was the subjective nonquantitative ASA classification 98% of the time and the Detsky index 2% of the time. The use of the ASA index occurred entirely at one university and only by general internists.
We derived a standard using the Detsky and L’Italien indices for patients undergoing nonvascular and vascular surgery, respectively. For patients undergoing nonvascular surgery, we chose the Detsky modified cardiac index for the comparison with physician risk estimates because it is based on a Bayesian model (allowing determination of pretest and posttest probabilities), meets the standards for clinical prediction rules as outlined by Wasson et al, has been prospectively studied and validated, and has the highest area under the receiver operating characteristic curve of all the general cardiac risk indices. When using the Detsky index, the type of surgery represents a pretest probability that is modified by a patient’s index score (ie, sum of points representing the patient’s clinical characteristics) to yield a posttest probability. Fortunately for clinicians, Detsky et al developed a simple likelihood ratio nomogram that facilitates the determination of posttest probabilities, thus avoiding the need for calculation. For our study, the Detsky posttest probability of a cardiac event was converted into categories of low, moderate, and high risk for the comparison to physician risk estimates. To make this conversion, we used the ACP guidelines for assessing and managing the perioperative risk from coronary artery disease associated with major noncardiac surgery. In these guidelines, the perioperative cardiac outcomes of interest are cardiac death and nonfatal myocardial infarction. Using these 2 outcomes, the guidelines define low risk as less than 3%, moderate risk as 3% to 15%, and high risk as greater than 13% chance of having 1 or both of these 2 outcomes. In the Detsky index, the outcomes were cardiac death, nonfatal myocardial infarction, and alveolar pulmonary edema. Since cardiac death and nonfatal myocardial infarction accounted for 80% of the outcomes in the original Detsky cardiac index study, the posttest probabilities would need to be 3% or less, 4% to 19%, and 20% or greater to correspond to the ACP’s definition of low, moderate, and high perioperative cardiac risk.

For patients undergoing vascular surgery, the L’Italien Bayesian model provided the reference standard. This index is based on a Bayesian model and performs well in prospective validation testing. As with the Detsky index, the L’Italien index uses the type of vascular surgery as a pretest probability. This factor is then combined with a patient’s risk score (determined through the presence of individual risk factors, such as age, diabetes, or angina) to yield a posttest probability of cardiac events. In our study, the L’Italien posttest probability of a cardiac event (cardiac death or myocardial infarction) was converted into categories of low, moderate, and high risk using the corresponding posttest probabilities of less than 3%, 3% to 15%, and greater than 13% for the reasons discussed above. Hereinafter, we will refer to these definitions of low, moderate, and high risk for nonvascular and vascular surgery as the Detsky/L’Italien standard.

In light of the subjective nature of these cutoff points, we performed a number of sensitivity analyses. These sensitivity analyses included using Detsky posttest probabilities of 10% or less (low), 11% to 30% (moderate), and greater than 30% (high). In another analysis, we used 5% or less (low), 6% to 30% (moderate), and greater than 30% (high). Two other sensitivity analyses used Detsky index points rather than the posttest probabilities. One such analysis used index point scores, which were in multiples of 5, of 5 points or less (low), 10 points (moderate), and greater than 10 points (high), while another used less than 5 points (low), 5 to 15 points (moderate), and greater than 15 points (high). Another similar sensitivity analysis used Goldman index point scores of less than 6 points (low), 6 to 25 points (moderate), and greater than 25 points (high), while yet another used Goldman index point scores of 3 points or less (low), 6 to 12 points (moderate), and greater than 12 points (high).

We performed other analyses assessing the proportion of physicians recording explicit use of a cardiac risk index and the proportion of physicians ordering specialized cardiac tests and tested whether there were differences between universities and specialists. These analyses used χ² or Fisher exact tests depending on cell frequencies.

PHYSICIAN RISK ESTIMATES

The Table presents physician estimates of cardiac risk contrasted with the Detsky/L’Italien standard. Physician estimates of perioperative cardiac risk (low, moderate, and high) agreed with our predetermined Detsky/L’Italien standard in 71% of the cases. When disagreement occurred it was due to physicians overestimating risk 16% of the time and underestimating risk 13% of the time. The observed agreement corresponds to a weighted κ of 0.38 (95% confidence interval, 0.28-0.49), indicating only fair agreement beyond chance. The third row of the Table highlights the lack of agreement indicated by the low κ score. This row represents high-risk patients as determined by the Detsky/L’Italien standard. Physicians rated 13 (42%) of these 31 patients at low risk for a perioperative cardiac event when the indices indicated high risk.

The κ values generated by the sensitivity analyses described above yielded similarly low κ values (range, 0.26-0.41), and there was no significant difference between university site or type of specialist (ie, general internist, cardiologist, and intensivist) when comparing physicians’ risk estimates and the Detsky/L’Italien standard.

CARDIAC TEST ORDERING

Figure 1 shows physicians’ noninvasive cardiac test ordering practices stratified by their estimates of cardiac risk. Overall, physicians ordered a specialized cardiac test in 7% of the consultations. When physicians estimated perioperative cardiac risk to be low, moderate, and high, they ordered a specialized cardiac test 6%, 7%, and 19% of the time, respectively (P = .05, for the difference across physician estimates of risk). Among these ordered tests, echocardiography was the most common (77%), followed by dipyridamole thallium imaging (9%), stress thallium imaging (9%), and Holter monitoring (5%).

Figure 2 demonstrates physicians’ thallium imaging ordering practices in patients undergoing vascular surgery. Physicians only ordered thallium imaging (either stress or dipyridamole) 3% of the time for patients undergoing vascular surgery. However, they never ordered thallium imaging for any of the 23 moderate-risk
patients undergoing vascular surgery (ie, the patient population in whom there is strong evidence supporting di-pyridamole thallium imaging\(^\text{4,12}\)). Also, there were no dobutamine stress echocardiography studies ordered despite the evidence supporting the use of this test and the recommendations of the ACP guidelines.\(^\text{4,12}\) Invasive cardiac testing (ie, coronary angiography) was ordered in 1 of the 297 patients.

The agreement between physicians’ risk estimates and the Detsky/L’Italien standard was suboptimal, and our sensitivity analyses indicate that agreement remains suboptimal despite the use of a number of different standardized definitions of low, moderate, and high risk. Physicians overestimated risk 16\% of the time, possibly leading to unnecessary anxiety for patients and inappropriate test ordering. Indeed, test ordering was shown to increase in relation to higher physician estimates of perioperative risk. Physicians in this study also underestimated risk 13\% of the time compared with the Detsky/L’Italien standard. Avoiding underestimation of risk is important because it allows patients to make more informed decisions about whether to proceed with surgery. It also allows identification of patients who have the most potential to benefit from close perioperative monitoring and attempted risk modification.

Considering our findings of overestimation and underestimation of risk, we strongly suggest that clinicians in the 2 universities studied increase their use of cardiac risk indices to more accurately assess baseline perioperative cardiac risk. The indices are easy to use and require minimal time.

We cannot be confident that a more accurate assessment of risk will improve outcomes because our ability to modify cardiac risk perioperatively remains uncertain. For example, it is uncertain that the cumulative risk of coronary artery surgery followed by noncardiac surgery is any different than that of noncardiac surgery followed by coronary artery surgery.\(^\text{16}\) On this point, Mason et al\(^\text{17}\) conducted a decision analysis that suggested that short-term outcomes may be worse in patients undergoing cardiac procedures before vascular surgery. However, uncertainty persists because this decision analysis assumed a low mortality rate for the baseline case and did not examine long-term outcomes. Likewise, despite recent encouraging evidence on \(\beta\)-adrenergic blocking agents,\(^\text{18}\) it is currently undetermined if \(\beta\)-adrenergic blocking agents\(^\text{19,20}\) or any perioperatively administered medications definitively modify cardiac risk.

Even with this uncertainty of how to modify risk, accurate risk estimates remain useful. As discussed herein, these estimates allow for more informed decision mak-
The evaluation of physician cardiac test ordering also raises concerns. Our findings suggest that physician test ordering does not follow the best available evidence as summarized in the ACP guidelines. Specifically, physicians underused dipyridamole thallium imaging and dobutamine stress echocardiography in moderate-risk patients undergoing vascular surgery. Consultants did not order either of these tests for any of the 23 moderate-risk patients undergoing vascular surgery.

Physicians ordered a cardiac test preoperatively 7% of the time, but it was usually echocardiography, a test that does nothing to refine clinical estimates of risk. It is possible that physicians were ordering this test preoperatively for specific questions about the general care of the patient beyond perioperative concerns. One example of this is ordering an echocardiogram to assess a patient’s ejection fraction to determine whether an angiotensin-converting enzyme inhibitor may be of benefit to the patient’s overall long-term health. Although we cannot be certain as to why physicians ordered a cardiac test preoperatively, our study nonetheless raises the possibility that physicians may be using echocardiography more than they should be.

We call the following study limitations to the readers’ attention. The first limitation is the study’s retrospective design, which only allowed for assessment of explicit use of cardiac risk indices. It is possible that some physicians used an index but did not explicitly mention its use in the medical record. However, our analysis of physician risk estimates compared with the Detsky/L’Italien standard indicates that physician estimates were often inaccurate, making it unlikely they used indices to determine risk. A second limitation is that the information we used to determine the cardiac index scores came from the consultation record in patient charts. It is possible that patients had other characteristics that were not recorded in the consultation but that would have increased their posttest probabilities using the Detsky/L’Italien standard. Such underdetection of clinical information would have led us to record posttest probabilities using the Detsky/L’Italien standard that were falsely low and could have explained some of the instances in which physicians overestimated risk. However, this would not explain the instances in which physicians underestimated risk. This underestimation of risk is one of our most concerning findings because consultants may be missing opportunities for risk modification and failing to accurately inform patients and referring surgeons of risk. The third potential limitation of our study is our use of the published pretest probabilities of the Detsky index, which may differ somewhat from the true pretest probabilities for the hospitals studied. However, our sensitivity analysis using the Detsky index points (and not probabilities) reached the same conclusion, thus minimizing the significance of this limitation. A final point of note is that our findings may not be applicable to anesthesiologists because their preoperative consultations were excluded from our study.

We conclude that, at the 2 universities studied, (1) consultants underused cardiac indices and often made inaccurate cardiac risk estimates and (2) the use of specialized noninvasive cardiac tests preoperatively was suboptimal. Increased use of the indices of Detsky and L’Italien and improved use of cardiac tests have the potential to refine cardiac risk estimation for patients undergoing noncardiac surgery.

Accepted for publication December 21, 1998.

This study was supported by the Foothills Hospital Foundation, Calgary, Alberta, Research and Development Project 1653/8297. Dr Ghali is funded by the Alberta Heritage Foundation for Medical Research, Edmonton (Population Health Investigator).


Reprints: William Ghali, MD, MPH, Health Science Center Room 1432, 3330 Hospital Dr NW, Calgary, Alberta, Canada T2N 4N1 (e-mail: wghali@ucalgary.ca).

REFERENCES


©1999 American Medical Association. All rights reserved.