Dr Sundt. And I agree with you. I didn't mean to take a shot at you for the 4% rate in the primary cases or the 8% in the redos. The real question, though, is, would the stroke rate have been less if they had been perfused from the ascending aorta rather than from the groin? But I appreciate that it is difficult to answer that question.

Dr Bolman. Well, they are perfused antegrade. The patients receive percutaneous venous but central aortic cannulation.

Dr Sundt. It was only percutaneous femoral?

Dr ElBardissi. They are cannulated centrally and perfused centrally except in reoperations, in which case they are perfused peripherally through the femoral artery or they are perfused through the axillary artery, but in all primary operations they are perfused through the aorta.

Dr Sundt. I am sorry I misunderstood. Thank you very much.

Dr Hillel Laks (*Los Angeles, Calif*). One of the factors that was not discussed in the article on robotic mitral valve surgery that we just heard or in this article on MiniAVR was the effects of these approaches on the brain, on neurocognitive function. One of the great concerns that many surgeons have is the issue of de-airing of the heart at the end of these procedures, which has been shown by transcranial Doppler to shower the brain with microemboli, and

the literature is replete with studies retrospectively that have shown 30% of patients having some neurocognitive dysfunctions. Do you know of any studies that have been done to compare these approaches, and particularly this one, in terms of neurocognitive function recovery, which is particularly pertinent to the old age group?

Dr ElBardissi. As far as neurocognitive function, if we are talking about these transient, diffuse neurocognitive dysfunction episodes, I don't know of any studies. If we are talking about acute events such as strokes or transient ischemic attacks, I can tell you that our institution did a retrospective study that was published in the *Annals of Surgery* and found no significant difference between the minimally invasive approach and the open approach as far as those acute events are concerned.

Dr Laks. I would like to emphasize that the incidence of clinically apparent strokes underrepresents diffuse brain injury that can result in neurocognitive dysfunction and that can be present as long as 6 months after surgery, and it is thought that many of those may be permanent. I think before we declare that these approaches are superior or equal, a randomized or equivalent type of study needs to be done looking at neurocognitive function both for robotic mitral valve surgery and for other minimally invasive approaches.

COMMENTARY

Predictive value of surgical scoring systems in determining operative risk for octogenarians undergoing aortic valve replacement

Marc R. Moon, MD

In a retrospective, single-center study, ElBardissi and associates¹ report the results of isolated aortic valve replacement (AVR) in 249 octogenarians from 1996 to 2009, during which time they performed a minimally invasive approach in almost all patients. This study does not compare the minimally invasive approach with the traditional AVR via median sternotomy, so any conclusions as to the impact of a particular surgical approach would not be reasonable. Therefore, I will make no reference to the specific surgical technique, but refer

only to AVR in general. The authors' specific aims included the following: (1) demonstrate that AVR can be performed safely in high-risk patients; (2) document long-term survival as a benchmark for future studies; and (3) assess the reliability of two current risk-prediction models, The Society of Thoracic Surgeons Predicted Risk of Mortality (STS-PROM) and modified European System for Cardiac Operative Risk Evaluation (EuroSCORE) algorithms,^{2,3} hypothesizing that these current risk-prediction models "overestimate risk in appropriately chosen and optimized patients." Although the authors were successful in addressing specific aims 1 and 2, they may have fallen short in their assessment of the STS-PROM and EuroSCORE risk score algorithms.

SPECIFIC AIM 1

The authors report operative mortality of $3\% \pm 2\%$ (95% confidence limit) for octogenarians undergoing AVR during this modern era. Mean ventilator duration was 16 ± 27

From the Division of Cardiothoracic Surgery, Washington University School of Medicine, St Louis, Mo.

Disclosures: Author has nothing to disclose with regard to commercial support.

Address for reprints: Marc R. Moon, MD, Joseph C. Bancroft Professor of Surgery, Division of Cardiothoracic Surgery, Washington University School of Medicine, 660 S. Euclid Ave, Box 8234, St Louis, MO 63110-1013 (E-mail: moonm@wustl.edu).

J Thorac Cardiovasc Surg 2011;141:335-7
0022-5223/\$36.00

Copyright © 2011 by The American Association for Thoracic Surgery
doi:10.1016/j.jtcvs.2010.12.015

hours, and major complications included stroke in 4%, pneumonia in 1%, and renal failure requiring dialysis in 1%. These operative results are outstanding, especially considering that this group had a median predicted surgical risk of 10.5% and 11% using STS-PROM and EuroSCORE, respectively. Thus, the authors' first hypothesis was true—AVR can be performed safely in high-risk patients.

SPECIFIC AIM 2

The authors report long-term survival after AVR at 1, 5, and 10 years to be 91%, 77%, and 56%, respectively (no measure of variability was provided). Compared with 2002 US Census data, survival was similar to that of an age-matched, gender-matched population ($P = .88$), both of which were substantially better than reported survival of octogenarians with aortic stenosis who had been denied surgery (52% 1-year and 22% 5-year survivals).⁴ Thus, the authors successfully demonstrated that AVR in octogenarians can restore normal life expectancy, a benchmark against which future studies can be compared.

SPECIFIC AIM 3

The authors report that although the STS score and EuroSCORE were well correlated ($r = 0.40$, $P < .0001$), the c-index (area under the receiver operator curve, which quantifies discriminatory ability) approached random chance (c-index = 0.5) rather than perfect discrimination (1.0) for both algorithms in this cohort of patients, 0.67 ($P = .18$) for STS-PROM and 0.53 ($P = .74$) for EuroSCORE. In addition, the Hosmer–Lemeshow statistic, which evaluates model calibration, confirmed that neither scoring system effectively predicted mortality in this selected patient cohort ($P = .38$ for STS-PROM and $P = .50$ for EuroSCORE). From these data, the authors concluded that “current risk-prediction models are not representative of the outcomes after cardiac surgery.” Caution needs to be expressed when evaluating the validity of this conclusion. The conclusion that the STS and European databases do not accurately predict operative risk can only be made after an adequately powered, randomized study, neither of which was present in the current retrospective analysis.

Population-based studies, although helpful in assessing the global community, cannot reliably predict the response of individual subjects or selected cohorts. Although one center may have an operative risk that is higher or lower than the model norms, the STS score and EuroSCORE do not overestimate risk for the entire study population. In fact, if the risk factor algorithms were applied to the entire database from which they were developed, their c-index would likely approach unity or perfect discrimination. According to the 2009 STS Executive Summary, the average center performs 301 cardiac surgery procedures, of which only 45 are AVRs⁵; surely these numbers pale in comparison with those reported from the Brigham.

Although 12 surgeons contributed patients to this study, the majority of the operations were performed by the senior surgeon, whom many consider to be one of the elite surgeons in the world during the past quarter century. I would speculate that the senior surgeon's results with coronary bypass grafting, mitral valve repair, and other cardiac procedures may also outperform the STS database, but this, in and of itself, does not invalidate the applicability of STS predicted risk to the “everyday” surgeon. Brown,⁶ Dewey,⁷ Wendt,⁸ and their associates also identified discrepancies between the STS score and operative results at their individual centers, albeit not quite as striking as the current report. Regardless, mathematical principle dictates that for each center like the Brigham that outperforms predictive risk within the STS or European database, there is a corresponding center (or centers) that underperforms predictive risk. Risk-prediction models simply cannot overestimate the risk of surgery in the population in which they were developed, assuming, of course, consistent, standardized definitions and quality data entry.⁹

One of the secondary goals of the Brigham study was to evaluate outcome in octogenarians who might otherwise be considered candidates for transcatheter aortic valve implantation (TAVI). This is reflected in the title, describing the cohort as “TAVI candidates.” However, according to current criteria for inclusion in the surgical arm of most randomized studies, a predicted operative risk of 8% to 10% is required for TAVI consideration, which would exclude over 25% of the patients included in this analysis. Although the authors did not report operative mortality for each predictive risk quartile, I suspect that mortality rates were quite low, even among the highest risk subset (17% and above), such that at the Brigham, very few patients, potentially only those at extreme risk, should ever be offered TAVI. In contrast, at centers whose operative mortality rates for AVR exceed STS norms, even patients with low predicted risk (4%–6%) may best be served with a percutaneous approach (based on the results of pending randomized studies, of course). This type of discussion could easily meld into a shift toward Pay for Performance and Centers of Excellence, but the reality remains that Centers of Excellence like the Brigham do not have the capacity to perform every AVR in the United States, and patient management decisions are going to continue to be individualized at the local level.

The beauty of the STS-PROM model is that it falls prey to very little selection bias (ie, elective participation, quality of follow-up data) and generates a starting point from which a surgeon can predict surgical risk for John Q. Public, with adjustments based on local modifiers. Only with caution can we extrapolate the findings of a single-center study to the general population inasmuch as published studies are generally lopsided toward success. Although the Brigham group may be eager to publish their outstanding results in high-risk patients, the centers that balance the curve may

be reticent to publicize their results. Results reported by the world's experts can be something for which to strive, but cannot reasonably be extrapolated to "Anytown, USA."

The STS score and EuroSCORE remain outstanding tools for preoperative patient counseling and quality improvement initiatives for individual centers and surgeons, so long as those using the algorithms understand the essential need to consider local modifiers that may affect outcomes when applying these data to individual patients or selected patient cohorts. ElBardissi and coauthors¹ express concern that "current scoring systems alone should not be used" to predict survival and that "current models for determining inoperability are unreliable, ... and should not unilaterally determine clinical practices," but this is nothing new. Such has never been the intent of population-based risk modeling protocols, and the STS clearly states, "These risks are solely statistical estimates, they should be supplemented by the professional judgment of the patients' healthcare provider, particularly their cardiac surgeon."¹⁰ Furthermore, there are surely patient factors that can affect operative risk that are difficult to characterize and or quantify (frailty, porcelain aorta, etc). The "eyeball test" can go a long way to help a surgeon, who understands his own limitations and those of his institution, predict a suboptimal outcome and identify patients whose operative risk may be substantially higher than standard algorithms predict.

In summary, ElBardissi and colleagues¹ are to be commended for their excellent results in this high-risk, elderly group undergoing AVR, but until randomized studies debunk the results from STS and European database-generated analyses, these algorithms will remain critical

for the preoperative assessment and postoperative evaluation of cardiac surgical patients and will best represent the expected outcomes for the "everyday surgeon," more so than data that emanate from individual, retrospective, single-center studies.

References

1. ElBardissi AW, Shekar P, Couper GS, Cohn LH. Minimally invasive aortic valve replacement in octogenarian, high-risk, transcatheter aortic valve implantation candidates. *J Thorac Cardiovasc Surg.* 2011;141:81-8.
2. O'Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2—isolated valve surgery. *Ann Thorac Surg.* 2009;88(1 Suppl):S23-42.
3. Nissinen J, Biancari F, Wistbacka JO, Lopenen P, Teittinen K, Tarkkiainen P, et al. Is it possible to improve the accuracy of EuroSCORE? *Eur J Cardiothorac Surg.* 2009;36:799-804.
4. Vaadarajan P, Kapoor N, Bansal RC, Pai RG. Survival in the elderly patients with severe aortic stenosis is dramatically improved by aortic valve replacement: results from a cohort of 277 patients aged > or = 80 years. *Eur J Cardiothorac Surg.* 2006;30:722-7.
5. <http://www.sts.org/sections/stsnationaldatabase/publications/executive/article.html>; The Society of Thoracic Surgeons: Executive Summaries and Sample Reports; accessed December 12, 2010.
6. Brown ML, Schaff HV, Sarano ME, Li Z, Sundt TM, Dearani JA, et al. Is the European System for Cardiac Operative Risk Evaluation model valid for estimating the operative risk of patients considered for percutaneous aortic valve replacement? *J Thorac Cardiovasc Surg.* 2008;136:566-71.
7. Dewey TM, Brown D, Ryan WH, Herbert MA, Prince SL, Mack MJ. Reliability of risk algorithms in predicting early and late operative outcomes in high-risk patients undergoing aortic valve replacement. *J Thorac Cardiovasc Surg.* 2008;135:180-7.
8. Wendt D, Osswald BR, Kayser K, Thielmann M, Tossios P, Massoudy P, et al. Society of Thoracic Surgeons score is superior to the EuroSCORE determining mortality in high risk patients undergoing isolated aortic valve replacement. *Ann Thorac Surg.* 2009;88:468-75.
9. Brown ML, Lenocho JR, Schaff HV. Variability in data: the Society of Thoracic Surgeons National Adult Cardiac Surgery Database. *J Thorac Cardiovasc Surg.* 2010;140:267-73.
10. <http://www.sts.org/sections/stsnationaldatabase/riskcalculator>; The Society of Thoracic Surgeons: STS National Database Risk Calculator; accessed December 12, 2010.