The Journal of
Thoracic and Cardiovascular Surgery

Ranking institutions

David Spiegelhalter

J Thorac Cardiovasc Surg 2003;125:1171-1173

phy in a patient with renal failure: first clinical report. *Cathet Cardiovasc Interv.* 2001; 54:68-9.
3. Adding LC, Bannenberg GL, Gustafsson LE. Basic experimental studies and clinical aspects of gadolinium salts and chelates. *Cardiovasc Drug Rev.* 2001;19:41-56.
doi:10.1067/mtc.2003.194

Risk adjustment in analysis of surgery for congenital heart disease

To the Editor:

The article by Jenkins and associates,[1] in the July 2002 issue of the *Journal,* is clearly an important advancement in risk adjustment when analyzing mortality outcome in the field of surgery for congenital heart disease. Further statistical analysis is needed in centers with high case volumes as the RACHS (Risk Adjustment in Congenital Heart Surgery) methodology evolves.

The concern regarding the large spectrum of mortality among risk categories 2 to 4, with some institutions displaying a threshold increase or decrease in mortality as higher-risk procedures, is disturbing for a potential severity model. This mismatch of observed/expected ratios is not surprising for severity scores in surgery in general.[2] Several factors involved that stem from inherent surgical practice and original logistic regression model of 5 variables were not discussed in the article.

The inherent surgical practice could be divided into physiologic and operative features. The operative variables, including in-hospital redo cases and estimated blood loss, should be evaluated. The first can have a dual knock effect on mortality. The estimated blood loss factor can have direct effects by contributing to child hypoxia and later the side effects of blood transfusion. Multivariate preoperative physiologic variables were not described in the original article by Jenkins and coworkers,[3] for example, hemoglobin, plasma sodium, and potassium. Recently, a hemoglobin concentration of 100 g/L or less had a 5-fold higher in-hospital mortality rate that those with higher concentrations.[4]

A widely recognized European severity score, POSSUM (a *P*hysiological and *O*perative *S*everity *S*core for the en*U*meration of *M*ortality and morbidity), used in vascular and colorectal surgery, is a simple scoring system incorporating only a total of 18 variables, of which preoperative electrolytes determined outcome.[5] There may be room for more specific and relevant variables to be included in RACHS-1.

Another aspect that could lead to the diversity of outcome in the intermediate risk category factors is whether a linear versus exponential model was adopted in the original equation. A linear method of analysis of the risk for each mortality group is artificially taken to the median as opposed to the mean in that risk category group. Although it may apply well to high-risk patients with smaller n values, it may overpredict death in the low-risk population, as was illustrated in centers A, F, O, S, and U in category 2 versus 3 and 4. Reanalyzing the data may produce different trends.

It is possible that these factors were considered; however, a discussion of these factors would be necessary for the reader. This would reduce chance-related outcome and might revisit the question regarding institutional inquiries, especially with regard to assignment of certain types of cases to specific surgeons and location of postoperative care.

Jeffrey H. Shuhaiber, MD
University of Illinois at Chicago
Department of Surgery
840 Southwood St
CSB Suite 518-E
Chicago, IL 60612

References

1. Jenkins KJ, Gauvreau K. Center-specific differences in mortality: preliminary analyses using the Risk Adjustment in Congenital Heart Surgery (RACHS-1) method. *J Thorac Cardiovasc Surg.* 2002;124:97-104.
2. Jones HJ, de Cossart L. Risk scoring in surgical patients. *Br J Surg.* 1999;86:149-57.
3. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH, Iezzoni LI. Consensus-based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg.* 2002;123:110-8.
4. Zindrou D, Taylor KM, Bagger JP. Preoperative haemoglobin concentration and mortality rate after coronary artery bypass surgery. *Lancet.* 2002;359:1747-8.
5. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br J Surg.* 1991;78:355-60.
doi:10.1067/mtc.2003.270

Ranking institutions

To the Editor:

Jenkins and Gauvreau[1] illustrated the use of a novel risk adjustment method in congenital heart surgery and chose to present their results largely in terms of institutional rankings. However, ranks are a notoriously inaccurate comparator for performance—someone always has to be bottom and top of a league table, no matter how much the play of chance may have contributed to their performance. Figure 1 shows the risk-adjusted standardized mortality ratios (SMRs) with 95% confidence intervals, as ranked by Jenkins and Gauvreau[1] according to outcomes from 22 institutions in 1996. We first note that a formal test that all the centers have SMRs of precisely 1 is barely significant ($\chi^2 = 35.6$, $df = 22$, $P = .03$, after transformation of all values to power 0.3 to bring to approximate normality), so there is not even strong evidence of any heterogeneity among centers. We can also estimate the "true rank" of each center. This requires the methodology described by Marshall and Spiegelhalter,[2] in which the "true SMRs" are repeatedly simulated from the confidence intervals in Figure 1 and then ranked at each iteration of the simulation. The resulting estimated true ranks and their 95% confidence intervals

TABLE 1. Probabilities of being "true best" and "true worst" centers for the 8 highest and lowest ranking centers

Center	Probability that "true best" center	Center	Probability that "true worst" center
C	0.27	J	0.01
D	0.16	Q	0.01
H	0.20	L	0.06
E	0.06	A	0.15
M	0.03	O	0.09
F	0.11	N	0.05
G	0.00	P	0.27
S	0.12	B	0.21

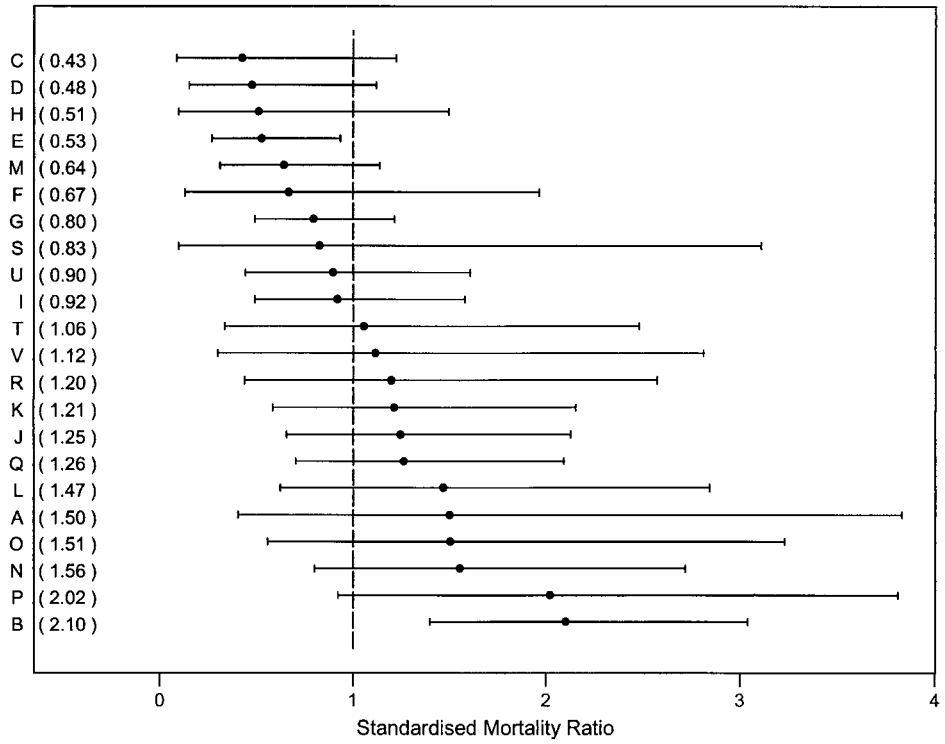

Figure 1. Risk-adjusted SMRs *(data points)* with 95% confidence intervals *(horizontal bars)* for 22 centers performing pediatric cardiac surgery in 1996, as ranked by Jenkins and Gauvreau.[1]

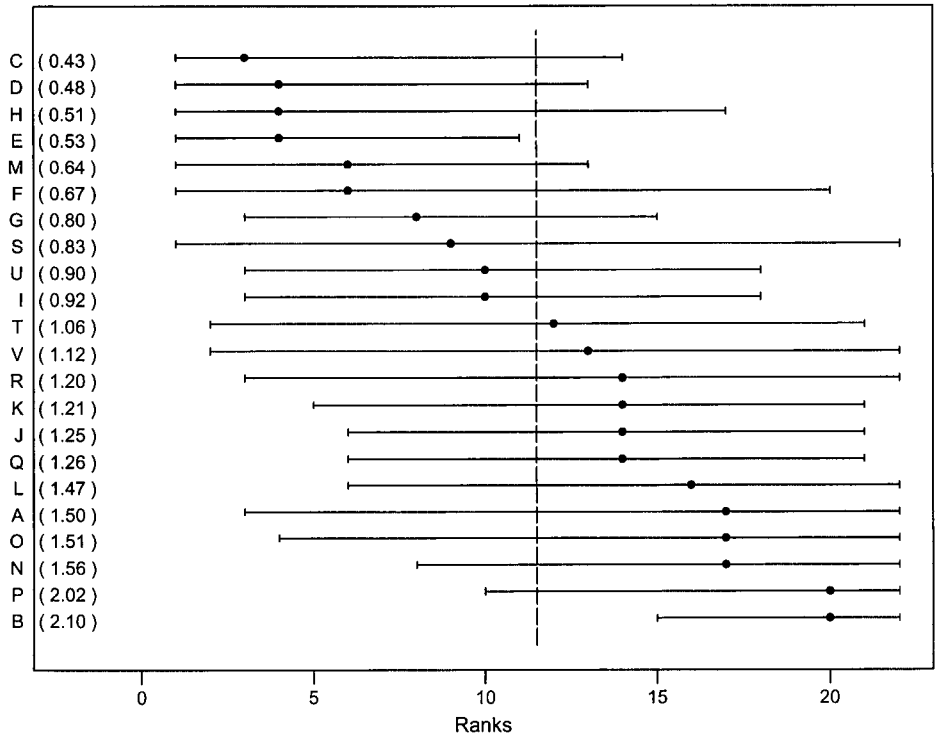

Figure 2. Estimated true ranks *(data points)* with 95% confidence intervals *(horizontal bars)* for 22 centers from Figure 1, showing great uncertainty associated with ranks ascribed to individual institutions.

are shown in Figure 2. There is considerable uncertainty about all the centers' true ranks, which naturally arises from the high degree of overlap of the confidence intervals in Figure 1. We can only state with confidence that center E is in the top half (despite being ranked fourth) and center B is in the bottom half; any further attempt at detailed ranking is spurious. Table 1 presents the probabilities that centers near the top or bottom of the league table truly are the best or worst centers. No center receives more than 30% chance of being either the winner or loser, although center P turns out most likely to be the worst by a small margin.

Such an analysis illustrates the grave dangers of institutional ranking unless there is clear heterogeneity among centers. It also explains why there are generally such radical changes in rankings from year to year when profiling institutions. Presentations that do not emphasize rankings, such as the funnel plots of Stark and colleagues,[3] are thus to be preferred.

David Spiegelhalter, PhD, Senior Scientist
MRC Biostatistics Unit
University Forvie Site
Cambridge, United Kingdom

References

1. Jenkins KJ, Gauvreau K. Center-specific differences in mortality: preliminary analyses using the Risk Adjustment in Congenital Heart Disease (RACHS-1) method. *J Thorac Cardiovasc Surg*. 2002;124:97-104.
2. Marshall EC, Spiegelhalter DJ. Reliability of league tables of in-vitro fertilisation clinics: retrospective analysis of live birth rates. *BMJ*. 1998;317:1701-4.
3. Stark J, Gallivan S, Lovegrove J, Hamilton JR, Monro JL, Pollock JC, et al. Mortality rates after surgery for congenital heart defects in children and surgeons' performance. *Lancet*. 2000;355:1004-7.
doi:10.1067/mtc.2002.265

Reply to the Editor:

We thank Shuhaiber and Spiegelhalter for their important questions regarding our article. Both questions relate to whether the differences in institutional outcomes for mortality after congenital heart surgery demonstrated in our article represent true differences in performance or were an artifact of our methodology, the Risk Adjustment in Congenital Heart Disease (RACHS-1) method.

Shuhaiber questions whether differences might have been mitigated had a more comprehensive method of risk adjustment been used. We agree that improved methods of risk adjustment would have increased our ability to compare outcomes accurately. However, methods including physiologic variables would have required validation in a population with complex congenital heart problems and would require extensive data collection. The RACHS-1 method was derived from a formal, consensus-based process and has been validated with two diverse data sets, with favorable performance characteristics. The consensus committee that developed RACHS-1 specifically sought to create a method of risk adjustment useful to understand group outcomes using data elements that are frequently available. To clarify, risk categories were incorporated into the risk adjustment model as binary covariates, which do not impose a linear or exponential relationship among categories. We agree that although most centers in the analysis had similar relative ranks across risk categories or worse performance for higher risk procedures, in 5 centers a surprising pattern of worse performance for higher risk performance was observed. Explorations by centers of why these patterns emerged should include a search for unmeasured risk factors but should also evaluate more programmatic possibilities, such as surgical referral patterns, location of postoperative care, and so on.

Spiegelhalter questions whether ranking institutions is an appropriate way to judge relative performance. Although we agree in general about the imprecision inherent in using ranks, especially when numbers of cases are small, we are attempting to guide quality improvement efforts in a field where considerable variability in institutional surgical mortality has been demonstrated by many investigators but annual caseloads are small and are unlikely to increase substantially. Although there may be uncertainty about a center's exact rank or about how large a difference in ranks is clinically important, program directors trying to guide improvement efforts should find it more useful to know their observed rank than to be informed that their center's relative performance did not reach statistical significance.

We would like to emphasize that we would never suggest "profiling" an institution on the basis of any single analysis, especially one derived from administrative data in a single calendar year. However, these analyses may prove useful to illuminate potential quality problems that need to be explored further, preferably by the institution itself.

Kathy Jenkins, MD, MPH
Kimberlee Gauvreau, ScD
Department of Cardiology
Boston Children's Hospital
300 Longwood Ave
Boston, MA 02115

Reference

1. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH, Jezzoni LI. Consensus-based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg*. 2002;123:10-8.
doi:10.1067/mtc.2003.462

Symmetry aortic connector system

To the Editor:

I read with interest the article by Donsky and associates[1] in which they outlined several misadventures with the Symmetry aortic connector system (St Jude Medical, Inc, St Paul, Minn). In our practice, we have an extensive series of off-pump coronary artery bypass operations in which the Symmetry connector has been used. Although the manufacturer has not recommended any anticoagulation regimen postoperatively, my colleagues and I routinely administer clopidogrel postoperatively for 6 weeks.

We justified this therapy after we demonstrated, at least by thrombelastography, a relative state of hypercoagulation after off-pump operations when compared with conventional cardiopulmonary bypass.[2] Furthermore, after deployment of an intracoronary stent, it is standard to prescribe a postprocedure course of clopidogrel therapy (ie, CLASSICS trial[3]). Since some of these stents are also composed of nitinol (ie, Scimed Radius stent, Boston Scientific, Boston, Mass), the management of a patient with an aortic connector should be no different from the documented protocol well described in the cardiology literature.

To date we have not experienced any complications with the aforementioned aortic connector and agree with the authors that the 2 cases that they described

Ranking institutions

David Spiegelhalter

J Thorac Cardiovasc Surg 2003;125:1171-1173

The Journal of
Thoracic and Cardiovascular Surgery