of pacemakers and implantable cardioverter defibrillators. Anesthesiology 2015; 123:1024–32

(Accepted for publication January 26, 2016.)

In Reply:

We thank you for your interest in our program for the perioperative management of pacemakers and internal cardioverter defibrillators (ICDs).[1] You and your colleagues are to be commended for taking on this often overlooked task.

Learning how to use the programming boxes was not a trivial process, at least for us. The screen appearance and the method of performing tasks not only differ among the device manufacturing companies, but also may vary from model to model. Videotaping the session for future use, including as a refresher, is a great idea that should be considered by programs taking on this task. Learning how to make basic programming changes is certainly possible with modest training, especially if all that is to be performed is the disabling of tachycardia sensing. The decision to convert to asynchronous pacing is sometimes more complicated, as it may require programming changes to determine the underlying rhythm, in addition to the consideration of the location and the extent of electrocautery. Additional device features, such as noise reversion and the mode switch response, may require further decision-making if demand pacing is used during the procedure. The individual providing the programming may have to reevaluate intraoperatively, as well. For example, our assessment of pacing dependency has occasionally proved wrong despite careful preoperative assessment, making intraoperative programming changes necessary.

An overarching goal of our program was to avoid making programming errors, especially with regard to restoration of the original device settings after surgery. Our caution was justified when we discovered that altered settings sometimes occur with restoration of demand pacing. It became quickly apparent that acquisition of complete device settings (baseline printout of all device settings) before making any programming changes is absolutely mandatory. We sincerely hope that any program performing this service, whether by anesthesiologists or cardiologists, takes such precautions. Our program faces the additional challenge of caring for surgical procedures that are not common to all practices, notably ventricular assist device implantation, during which pacing capture may fail due to lead dislodgement or tissue trauma. In such situations, intraoperative lead impedance, sensitivity, and threshold testing must be performed expeditiously to determine a solution. We wanted to be able to handle these more advanced tasks, something that may not be necessary in all practices.

Your process, whereby devices are evaluated well in advance of surgery, is an important aspect of appropriate management. Our preoperative clinic sees only a fraction of the patients with devices. Although all patients with these devices should have regular follow-up, that is not always the case. In consequence, when we interrogate the device in the preoperative holding area, we occasionally discover some level of device malfunction or low battery.

Finally, we completely agree with you that all practices should devise a system for managing devices that involves more than just "placing a magnet." What works best for a given institution will depend on the institution's patient population, case mix, system of preoperative assessment, and level of institutional support. No system, including ours, is likely to be perfect. But it will be better than no plan at all.

Competing Interests

Drs. Rooke and Poole were supported by Medtronic, Inc., Minneapolis, Minnesota, for a study of electromagnetic and cardiovascular implantable electronic devices. Dr. Poole has received honoraria for educational speaking from Biotronik, Lake Oswego, Oregon; Boston Scientific, Marlborough, Massachusetts; Medtronic, Inc.; and St. Jude Medical, St. Paul, Minnesota. Dr. Poole is also on the Medical Advisory Board for Boston Scientific. The other authors declare no competing interests.

G. Alec Rooke, M.D., Ph.D., Stefan A. Lombaard, M.B.Ch.B., F.A.N.Z.C.A., Gail A. Van Norman, M.D., Jörg Dziersk, M.D., F.R.C.A., Krishna M. Natrajan, M.B.B.S., F.R.C.A., Lyle W. Larson, Ph.D., P.A.-C., Jeanne E. Poole, M.D. University of Washington, Seattle, Washington (G.A.R.). rooke@uw.edu

Reference

1. Rooke GA, Lombaard SA, Van Norman GA, Dziersk J, Natrajan KM, Larson LW, Poole JE: Initial experience of an anesthesiology-based service for perioperative management of pacemakers and implantable cardioverter defibrillators. Anesthesiology 2015; 123:1024–32

(Accepted for publication January 26, 2016.)

Predilection for Poor Prediction with the Surgical Apgar Score

To the Editor:

I enjoyed the recent article by Terekhov *et al.*,[1] "Preoperative Surgical Risk Predictions Are Not Meaningfully Improved by Including the Surgical Apgar Score" (SAS). I value the contributions of these authors to this field of investigation, including their pioneering work with the SAS.[2] The authors made two methodological choices that may have contributed to the study concluding no improvement in prediction, so I humbly offer two suggestions to permit a more definitive test of their hypothesis.

First, would the authors consider performing their analyses using an alternative sampling interval for vital signs? The authors constructed the vital signs components of the SAS,

lowest heart rate, and lowest mean arterial pressure, using "instantaneous" measures, or the true lowest heart rate and lowest mean arterial pressure in the record. These "instantaneous" values for the SAS are the least useful option for predicting outcomes when compared with alternatives such as moving median values over 5- and 10-min windows.[3] In essence, the choice of instantaneous values biases the assessment to no benefit of the SAS.

Second, would the authors consider adding a calculation of risk reclassification to better test the clinical utility of the SAS? The authors reported the c-statistic and Brier score to evaluate the utility of the SAS. Although statistically robust, neither of these measures provides clinical insight. Moreover, the c-statistic is known to change minimally even when important improvements are made with risk prediction.[4] For this reason, the use of a reclassification measure may be applied to provide a more clinically meaningful assessment of change in risk prediction.[5] Reclassification approaches can be problematic, but the concept of categorizing patients into high- and low-risk groups is clinically intuitive and actionable, because we treat high-risk patients differently such as with admission to the intensive care unit.

The potential for real-time risk revision is not known, and with these suggestions, the authors may be able to more robustly test its potential.

Competing Interests

Although the author is a current awardee of the Anesthesia Patient Safety Foundation and Anesthesia Quality Institute, this letter was written on separate time. "No Funding Received" is the most accurate description of the funding details. The author receives no funding from industry or honoraria and has no conflicts of interest to disclose.

Joseph A. Hyder, M.D., Ph.D., Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, Minnesota. joseph.a.hyder@gmail.com

References

1. Terekhov MA, Ehrenfeld JM, Wanderer JP: Preoperative surgical risk predictions are not meaningfully improved by including the Surgical Apgar Score: An analysis of the risk quantification index and present-on-admission risk models. Anesthesiology 2015; 123:1059–66
2. Regenbogen SE, Ehrenfeld JM, Lipsitz SR, Greenberg CC, Hutter MM, Gawande AA: Utility of the surgical Apgar score: Validation in 4119 patients. Arch Surg 2009; 144:30–6; discussion 37
3. Hyder JA, Kor DJ, Cima RR, Subramanian A: How to improve the performance of intraoperative risk models: An example with vital signs using the surgical Apgar score. Anesth Analg 2013; 117:1338–46
4. Cook NR: Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007; 115:928–35
5. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW: Net reclassification improvement: Computation, interpretation, and controversies: A literature review and clinician's guide. Ann Intern Med 2014; 160:122–31

(Accepted for publication January 28, 2016.)

In Reply:

We thank Dr. Hyder for his interest in our recent article published in Anesthesiology, "Preoperative Surgical Risk Predictions Are Not Meaningfully Improved by Including the Surgical Apgar Score: An Analysis of the Risk Quantification Index and Present-On-Admission Risk Models."[1]

As suggested by Dr. Hyder, we performed additional analyses using an alternative sampling interval for vital signs and added a calculation of risk reclassification to better test the clinical utility of the Surgical Apgar Score (SAS) when combined with preoperative risk stratification models.

A sampling method for slowest heart rate (HR) and lowest mean arterial pressure (MAP) was established before initiating data analyses. The method was based on "windows" or intervals of data and was established as follows: 10-min nonoverlapping windows, with windows beginning at the time of incision (0 to 10 min, 11 to 20 min, 21 to 30 min, *etc.*). Within each window, a median value was determined. Median values for HR and MAP were the basis for the original SAS investigations, and median values were chosen for this investigation. Estimated blood loss as recorded by the in-room anesthesia provider was calculated for the entire case.[2]

We also added a calculation of risk reclassification to better test the clinical utility of the SAS. The use of a reclassification measure may be applied to provide a more clinically meaningful assessment of change in risk prediction. A concept of categorizing patients into high- and low-risk groups is clinically intuitive and actionable, as we treat high-risk patients differently, such as with admission to the intensive care unit. Traditionally, risk prediction models have been evaluated using the area under the receiver operating characteristic curve, along with model calibration, Brier score, information criteria, *etc.*, but this can be an insensitive measure for model comparison in a healthcare setting, providing little direct clinical relevance. Since its description in 2006, much interest has been generated in reclassification, which assesses the ability of new models to more accurately classify individuals into higher or lower risk strata. This has led to new methods of evaluating and comparing risk prediction models, including the reclassification calibration test and the net reclassification index (NRI). Pencina *et al.*[3] developed the NRI and the integrated discrimination improvement (fig. 1).

After performing analyses using alternative sampling for vital signs and calculating risk reclassification, the Risk Quantification Index and present-on-admission preoperative risk models were not meaningfully improved by adding intraoperative risk using the SAS, as determined by the NRI value of 0.02 ($P = 0.10$). These analyses supported the original findings: adding the SAS did not substantively improve predictions. In addition to the estimated blood loss, lowest HR, and lowest MAP, other dynamic clinical parameters from the patient's intraoperative course may need to be combined with procedural risk estimate models to improve risk stratification.