Evaluating Anesthesiology Residency Training Programs:
What Criteria Should Medical Students Use?

Catherine Barden, M.D.
Peter Rock, M.D., M.B.A., FCCM, President-Elect
Jeffrey R. Kirsch, M.D., President
David A. Zvara, M.D., Member
Ronald G. Pearl, M.D., Ph.D., Member
Charles W. Whitten, M.D., Member
Society of Academic Anesthesiology Associations

Attempts have been made to rank institutes of higher education for decades. In 1983, *U.S. News and World Report* (*U.S. News*) published its first rankings of undergraduate programs, based solely on the opinions of college presidents.[1] Because of the poor relevance of ranking based on opinion, the methodology of these rankings evolved to include statistical data and a shift toward evaluating colleges by the success they have in graduating students (outcomes data). In 1990, "America's Best Graduate Schools" was published, including annual listings of medical, engineering, law, business and education schools and has become the default method by which these schools are compared.[2,3] Last year, *U.S. News* and Doximity, a social media network for physicians, launched Residency Navigator, the first ranking of graduate medical education programs. With more than half of U.S. physicians as members, Doximity is the largest site of its kind.[4] In 2014 and again recently, its board-certified physician members were invited to submit peer nominations for up to five residency programs within their specialty. Frequency counts of the number of residency program nominations were used to rank training programs.

Applicants to anesthesiology residency programs utilize several resources when comparing programs, such as departmental websites, Web-based comments from other applicants (The Student Doctor Network [SDN], Scutwork), advice from faculty advisors and student affairs offices, and databases of program demographics (FRIEDA) (Table 1). The

Catherine Barden, M.D. is Associate Professor and Program Director, Department of Anesthesiology and Pain Management, UT Southwestern Medical Center, Dallas.

Jeffrey R. Kirsch, M.D. is Professor and Chair, Anesthesiology and Perioperative Medicine, Associate Dean, Clinical and Veteran's Affairs, Oregon Health & Sciences University, Portland, Oregon.

Peter Rock, M.D., M.B.A, FCCM is Martin Helrich Professor and Chair, and Professor of Anesthesiology, Medicine and Surgery, Department of Anesthesiology, University of Maryland School of Medicine, Baltimore.

David A. Zvara, M.D. is Professor and Chair, Department of Anesthesiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina.

difficult professional and personal process of selecting a residency program leads most applicants to utilize most, if not all, of these resources. Unfortunately, the validity of the information presented is widely variable, as some resources represent anecdotal experiences of commentators (Scutwork, SDN), and others are databases of demographics, presenting no measure of program quality or outcomes (FRIEDA). A recent report from the Institute of Medicine issued a call for transparency into the outcomes-based performance of residency programs.[2] In an era of increasing importance of outcomes metrics, programs should be evaluated on the same principles.

Table 1: Available Resources for Medical Students Applying to Anesthesiology Training Programs

FRIEDA online: Information for 9,000-plus AGCME-accredited graduate medical education programs. Program-reported data are collected by the AAMC via an annual survey.

The Student Doctor Network, Scutwork.com: Student-driven online resource. Comments and opinion from site members only.

SAAA site for Research and Education Innovation.

Mentorship from advisors within the specialty: Useful for "intangibles."

U.S. News and World Report/Doximity: Ranks residency programs by specialty and region based upon information self-reported by board-certified physician members and a proprietary database of all U.S. physicians. Member physicians can submit peer nominations for up to five residency programs. Filters are available to narrow programs based upon location, preferences (research, urban or rural setting) and fellowships.[4]


Prior to the availability of program demographic data on the Internet, the primary resource for applicants to the specialty was a program director, chairperson or trusted mentor from whom to seek advice about the quality and stability of a training program's environment. While still a cornerstone of advice for students, the information provided may be biased, is sometimes out of date and may not adequately take into account the personal needs of the applicant. In addition, the advice may be based on the reputation of the department rather than on the quality of the training program.

As a specialty, we are attracting an increasingly competitive group of applicants who want to make an informed decision about where to apply, interview and how to construct a Match rank list. Thus, the idea of a more formal GME "ranking" system is particularly alluring to this high-quality applicant pool. Rankings also have great appeal to institutions as a recruitment tool, not only to attract top medical students but also junior faculty members and philanthropy.

Weakness of opinion-based ranking systems such as *U.S News* or Doximity include:

- The number of responding physicians represents a small percentage of those practicing, and these individuals often have limited firsthand knowledge of any institution other than the ones at which they trained or work.[2]
- Halo effect of parent institution with which the program is affiliated that does not necessarily translate into program quality. The emphasis on "name recognition-type popularity" may be misleading to applicants to the specialty.
- Quality and diversity of affiliated hospital partners within which residents train (of vital importance in a hospital-based specialty) is not evaluated.[4,5]
- Data, which for anesthesiology is mostly self-reported by Doximity members, may be severely skewed by programs "stuffing the ballot box" by encouraging votes from current department members and alumni.
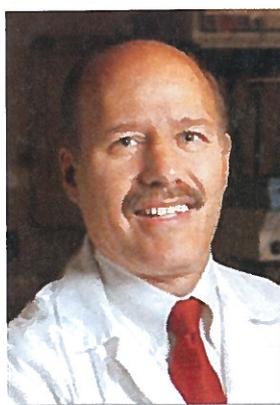
Ronald G. Pearl, M.D., Ph.D. is Professor and Chair, Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University Medical Center, Stanford, California.

Charles W. Whitten, M.D., is Professor and Chairman, Margaret Milam McDermott Distinguished Chair In Anesthesiology and Pain Management, UT Southwestern Medical Center, Dallas.

- Rankings are by for-profit social media sites that generate revenue by selling access to physician-users to clients that include pharmaceutical companies, market research and hedge funds and other investors, so there may be a conflict of interest between the revenue generating aspects of their site and the information they provide. The dual nature of these sites may undermine integrity and transparency as forums for exchange of medical opinion and presents an ethical conflict for physicians who use the sites.[6]
- Objective data is buried and not related to ranking.

A comprehensive review of program quality would ideally be based upon data from reliable, well-established third-party sources such as the AGCME Residency Review Committees and/or the American Board of Anesthesiology (ABA), both of which use quality metrics to evaluate programs.[2,5,7] Other sources of information relevant to the applicant and potentially available from the ABA could include:

- Program-level specialty data (Maintenance of Certification in Anesthesiology (MOCA®) performance.
- Practice demographics, scope-of-practice information.
- Entering resident demographics (academic data, AOA status, publications, couples match data).
- ACGME data (case logs, work hours, aggregated resident and faculty survey data).

A combination of these quality metrics, if available in a transparent form, is a stronger and more objective method by which the end consumer (medical students) can evaluate program quality. Such resources would also help dispel myths and inaccuracies about programs, which are propagated in student-driven online forums and by mentors who have no specific knowledge of programs about which they are advising. It is also important for applicants to know which programs have graduates who always perform well on examinations (i.e., programs that have strong graduates because they recruit the strongest applicants) as compared to those programs that are able to transform residents with a weaker entering portfolio into anesthesiologists who ultimately become outstanding anesthesiologists with excellent performance on their ABA certification process.[7]

Recent publications from other specialties have explored novel methods of comparing programs based upon quality and outcomes-based metrics. An example from the general surgery literature describes a sample ranking system that relies on input, process and outcomes measures and controls for program and resident characteristics (size, residents' entering aptitude and research requirements) to offer a more valid measure of a program's ability to generate high-quality surgeons.[5]

The host of intangibles that helps construct the environment in which a resident trains cannot possibly be measured by a series of data points; however, for naïve medical students looking for the right "fit," a relevant and valid tool for comparison of programs is essential. The future of our specialty deserves better benchmarks than those recently released by Doximity, which are subjective rankings of anesthesiology training programs. It underscores the need for the specialty and its affiliated organizations to establish and publish measures that define program quality in an era of pay for performance. A solution should begin with a move away from commercial entities whose motives may not be aligned with our specialty and which compare programs based upon historical reputation, dominance and size, and instead focus on quality of education, learning environment, diversity of clinical experience and innovative programs.

Formal ranking of anesthesiology residency training programs would be divisive to the specialty and not particularly helpful to applicants. A significant advance for applicants and programs would be a mechanism by which applicants can have access to relevant, objective data about programs so that he or she can decide which factors are most important to his or her own individual needs.

References:
1. Boyington, B. Infographic: 30 editions of the U.S. News Best Colleges rankings. U.S. News Education website. http://www.usnews.com/education/best-colleges/articles/2014/09/09/infographic-30-editions-of-the-us-news-best-colleges-rankings. Published September 9, 2014. Accessed September 15, 2015.
2. Mellinger JD, Damewood R, Morris JB. Assessing the quality of graduate surgical training programs: perception vs reality. *J. Am Coll Surg.* 2015;220(5):785-789.
3. Goldstein MJ, Lunn MR, Peng L. What makes a Top Research medical School? A call for a new model to evaluate academic physicians and medical school performance. *Acad Med.* 2015;90(5):603-608.
4. Doximity. https://www.doximity.com. Accessed September 15, 2015.
5. Wilson AB, Torbeck LJ, Dunnington GL. Ranking surgical residency programs: reputation survey or outcomes measure? [Published online ahead of print May 14, 2014.] *J Surg Educ.* doi: http://dx.doi.org/10.1016/j.jsurg.2015.03.021.
6. Landa AS, Elliott C. From community to commodity: the ethics of pharma-funded social networking sites for physicians. *J Law Med Ethics.* 2013;41(3):673-679.
7. Peterson LE, Carek P, Holmboe ES, Puffer JC, Warm EJ, Phillips RL. Medical specialty boards can help measure graduate medical education outcomes. *Acad Med.* 2014;89(6):840-842.