A Solution Worth Trying: Providing Program Preference in Residency Applications

Amber Lin, MS
Lalena M. Yarris, MD, MCR

In this issue of the *Journal of Graduate Medical Education*, Whipple and colleagues examine a theoretical model for a possible modification to the current residency application process: allowing medical students to provide their program preferences early in the application process.[1] They argue that the prevailing system results in an abundance of applications, which prevents programs from comprehensively reviewing all candidates. The authors suggest that their modification could provide an additional criterion to screen and reduce the pool of interested applicants. Whipple and colleagues provide preliminary theoretical evidence that their intervention may conserve resources with little impact on the highest-performing students while potentially benefiting programs and all other students.

From the perspective of a program director in a competitive specialty (L.M.Y.), this article's conceptual framework hits home. Our emergency medicine residency program received applications from more than 1300 students this year for 11 positions, and we were able to interview approximately 10% of applicants. The assumptions made by the authors based on their experience in otolaryngology (detailed in the article's online supplemental material) approximate my experience: a program can holistically review a maximum of 40 applications per residency slot, can interview a maximum of 10 applicants per residency slot, and a student can interview at a maximum of 20 residency programs during the interview season.

In the article's discussion section the authors describe their model's prediction of "a counterintuitive situation where a competitive specialty could have both unmatched and unfilled programs." While this situation may be counterintuitive, in my experience, it represents the current reality. The most desirable programs are likely to have a common pool of top applicants. Because the metrics best suited to screening, such as United States Medical Licensing Examination (USMLE) scores, medical school status, and student ranking, are not subject to program interpretation, we find ourselves with a common pool of applicants *invited* to interview as well. However, as the season progresses, the combination of interview trail fatigue and positive feedback regarding their competitiveness can result in applicants withdrawing from interview appointments at programs lower on their preference list. This practice can be good for students and good for the program, as long as they match and the program fills. However, as a program director, I have had last-minute interview day openings (or even "no shows") that I have been unable to fill, and as an advisor, I have seen students with strong, but not top, applications struggle to secure interviews and even go unmatched.

Thus, this article is timely and the conversations it will spark are necessary. It is in the best interest of trainees and programs to modify a system that currently encourages indiscriminate, if inadvertent, saturation of programs' abilities to holistically review applications and determine interview day selections. The current system does not provide each applicant the same opportunity to match at a program that may be an excellent mutual fit.

The authors begin their study by substantiating a previously peer-reviewed analysis[2] as well as an analysis by the Association of American Medical Colleges.[3] These studies found that when students apply to many programs, they receive a slightly higher number of interview invitations; thus, the practice of applying to as many programs as possible is advantageous, yet only marginally.

Whipple and colleagues then examine—using theoretical models and simulation—the effect of their suggested intervention (offering students the option to reveal program preferences) on the number of subsequent interviews offered. While the authors should be commended for their candid discussion of the limitations of this analysis, there are significant issues with the generalizability and validity of their approach.

It is important to restate the authors' acknowledgement that their analysis is based on assumptions and simplifications, as is required in simulation analyses. This simplification is multifaceted: it analyzes only one method for implementing statements of preference, relies on a potentially inaccurate random

DOI: http://dx.doi.org/10.4300/JGME-D-18-01092.1

distribution of student and program characteristics, and lacks an investigation across the combination of programs' competitiveness and student application quality.

First, this analysis considers only one possible method by which a program interprets a conveyance of preference. In all simulations the authors use a statement of preference as an additive metric, to enhance students' easy to review scores (eg, USMLE scores and class rank). For the top 10 programs ranked, the authors equate a student's preference to an approximately 10-point increase on USMLE Step 1. However, implementation of preference "points" will likely differ from program to program (based on the quality and size of the applicant pool as well as other yet unknown factors) and potentially from student to student within programs.

This leads to additional limitations, as the authors assume a random distribution of "easy to review" and "hard to review" characteristics across students and programs, which is largely unrealistic in practice. This is another area for further research. It also is important to consider how students would choose to use preference rankings, particularly for average or lower-performing students. Should they use all of their rankings of preference for their "reach" programs or should they use their rankings at schools for which they are stronger candidates? The investigators could analyze this by stratifying by the interaction between low- and average-performing students and highly, moderately, and less competitive programs. As an important aside, when comparing the number of interviews offered among differing scenarios regarding program preference, the authors include a simulation in which only the index student provides program preferences. This scenario is highly unlikely and the results may be misleading, especially since the student who provided preference receives dramatically more interview offers. Further analyses should limit simulations to more realistic scenarios (eg, half of applicants express program preference).

Some crucial considerations when performing simulation-based research include transparency in methods, as well as verification and validation. For a sufficient evaluation of a simulation to occur, it is imperative for authors to include their code for review, ideally as a supplement accompanying the research manuscript, in order to aid in evaluation and foster reproducible research.[4] Methods of verification and validation for simulation studies may include formal examination of the conceptual model by content experts from a variety of residency programs, detailed reviews and descriptions of the results of all intermediate steps of the simulation, sensitivity analysis to determine the robustness of the models to small deviations in assumptions and parameters, and measures of prediction error for the final analysis (eg, confidence bands around all estimates).[5,6]

One final consideration for this application model is that the number of interviews offered was the chosen outcome of interest. This is a convenient and logical outcome as it is the component of the residency application process that is most directly affected by this intervention. However, the most student-centered outcome would be a successful match, preferably at their first-ranked program. Since preference is already included in the postinterview assessment, it would be interesting to see this analysis extended to the probability of a match. The authors could supplement their current model by adding a step incorporating the interview assessment into their hard to review characteristics. In the same fashion as they determined the estimated number of interviews offered, they could plot the estimated proportion of students matched at their preferred program. This analysis would certainly be of interest to applicants and programs, especially when combined with an analysis of interaction between student application quality and competitiveness of a program.

Even with the models' simplifications and assumptions, the authors make a satisfactory case that the introduction of student preferences could have a mitigating effect on the mismatch between the excessive number of residency program applications and available interview slots, with a low potential for negative impact on students and residency programs. While the authors have focused on the highly competitive specialty of otolaryngology, this is a solution that could be applied across a wide range of specialties.

The authors comment that other proposed solutions to the excessive number of residency applications add undue burden and potentially disadvantage lower-income students. We agree and would emphasize that this new model presents a better solution. We believe this may be a more equitable approach than increasing the cost of applications, less burdensome than requiring additional essays for each program, and likely more palatable than limiting the total number of applications. Thus, we encourage the National Resident Matching Program (NRMP) to carefully review the possibility of further testing of this intervention. If this intervention is implemented, we urge the NRMP to provide best practices as to how programs will interpret student preferences and to carefully examine—both qualitatively and quantitatively—the impact of the intervention. Finally, we urge programs to be transparent with their methods of model implementation.

References

1. Whipple ME, Law AB, Bly RA. A computer simulation model to analyze the application process for competitive residency programs. *J Grad Med Educ*. 2019;11(1):30–35.
2. Weissbart SJ, Hall SJ, Fultz BR, Stock JA. The urology match as a prisoner's dilemma: a game theory perspective. *Urology*. 2013;82(4):791–797. doi:10.1016/j.urology.2013.04.061.
3. Apply Smart: New Data to Consider. https://students-residents.aamc.org/applying-residency/article/apply-smart-data-consider. Accessed December 9, 2018.
4. Extending transparency to code. *Nature Neuroscience*. 2017;20(6):761. doi:10.1038/nn.4579.
5. Anderson AE, Ellis BJ, Weiss JA. Verification, validation and sensitivity studies in computational biomechanics. *Comput Methods Biomech Biomed Engin*. 2007;10(3):171–184. doi:10.1080/10255840601160484.
6. Hicks JL, Uchida TK, Seth A, Rajagopal A, Delp SL. Is my model good enough? Best practices for verification and validation of musculoskeletal models and simulations of movement. *J Biomech Eng*. 2015;137(2):0209051–02090524. doi:10.1115/1.4029304.

Both authors are with Oregon Health & Science University. **Amber Lin, MS,** is an Instructor, School of Public Health, and a Staff Biostatistician, Center for Policy and Research in Emergency Medicine; and **Lalena M. Yarris, MD, MCR,** is a Professor and Emergency Medicine Residency Director.

Corresponding author: Amber Lin, MS, Oregon Health & Science University, Mailcode CR114, 3181 SW Sam Jackson Park Road, Portland, OR 97239, 503.494.6924, lauriea@ohsu.edu