The Decline in Family Medicine in-Training Examination Scores: What We Know and Why It Matters

Warren P. Newton, MD, MPH, Ting Wang, PhD, and Thomas R. O'Neill, PhD

(J Am Board Fam Med 2023;36:423–426.)

Since the first Board examination in Ophthalmology in 1916, assessment of cognitive expertise has been foundational to Board Certification. Since then, there have been dramatic improvements in the methodology of high-stakes examinations, including psychometric techniques, the writing of questions, and differential item functioning analysis. In 1969, ABFM introduced the requirement of periodic recertification to the Board Certification Community, and subsequent research in the cognitive sciences has confirmed the importance of independent assessment of cognitive expertise across many fields. To support family medicine residencies, the ABFM conducts an in-training examination (ITE) every fall; as of 2008, the ITE uses the same psychometric scale as the certification examination, making the scores comparable to certification scores. With the aid of an easy to use, web-based app (https://rtm.theabfm.org/bayesian/predictor), residents and their residency directors can estimate their likelihood of passing the Board certification examination. In both of the past 2 years, this app has been accessed more than 200,000 times by more than 15,000 users.

Since 2020, however, scores on in-training examinations in Family Medicine have dropped significantly. Figure 1 depicts the major changes, with drops in average scores for each class each each of the past 3 years. The data from almost all family medicine residents are included. Given the large numbers of residents, the confidence intervals are very small. Moreover, controlling for baseline USMLE score, gender, proportion of minorities underrepresented in medicine, international medical graduates and Doctors of Osteopathy does not change the relationship. The overall drop is clinically meaningful. In general, the growth of knowledge during residency averages 30 to 40 points per year of residency, and this seems to be true for all subgroups of residents;[1] the aggregate drop we've seen in the past 3 years suggests that current interns have test scores approximately 1.25 years lower than interns in 2019.

Intriguingly, however, the score and pass rate on the high-stakes certification examination have *not* changed meaningfully over the same 3 years. It is reassuring that certification scores and pass rates have not changed, but the immediate question is why has a gap emerged between in-training and certification scores. One major possibility is that it may be that what we are seeing is the difference between a low-stakes examination (the in-training examination) and the high stakes certification examination. Perhaps, when it counts, residents (and their faculty and program directors) take the high stakes certification examination more seriously, by responding to the low PGY-3 ITE scores in the fall and putting more effort into studying in advance of the certification examination, which most people take in the spring. Alternately, however, it could be that certification scores and pass rates are a "lagging indicator" and that we may see a decline in the next years in which the full effect of COVID on both medical school and residency is felt.

Importantly, these changes in knowledge assessments are not limited to family medicine. Review of 2022 data across specialties is pending, but other specialties, including Emergency Medicine, Psychiatry, and some of the Pediatric and Medicine subspecialties—

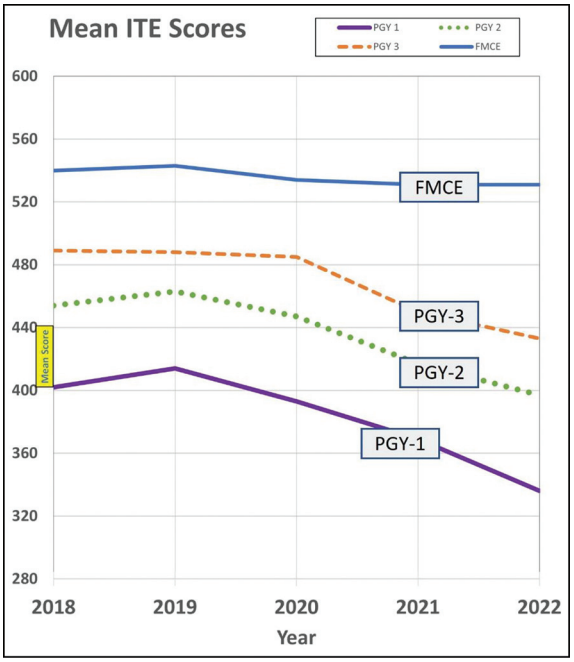
From the American Board of Family Medicine, Department of Family Medicine, University of North Carolina (WPN); American Board of Family Medicine (TW); American Board of Family Medicine (TRO).

Conflict of interest: The authors are employees of the ABFM.

Corresponding author: Warren P. Newton, MD, MPH, American Board of Family Medicine, 1648 McGrathiana Pkwy, Ste 550, Lexington, KY 40511-1247 (E-mail: wnewton@theabfm.org).

J Am Board Fam Med: first published as 10.3122/jabfm.2023.230092R0 on 7 April 2023. Downloaded from http://www.jabfm.org/ on 10 November 2026 by guest. Protected by copyright.

Figure 1. Family medicine in-training examination (ITE) and certification trends 2018 to 2022. Abbreviations: PGY, Postgraduate Year; FMCE, Family Medicine Certification Examination.



but not most surgical disciplines—also report significant drops in either in-training or certification examinations in 2021. In addition, as is well known, standardized test scores have dropped nationally in K-12.[2] Why would residency education be different from education at other levels?

ABFM believes that it is important to take these data seriously: they suggest an important decline in clinical knowledge acquired by Family Medicine residents. Given the time course, it is also likely that the COVID pandemic, with its many impacts on clinical care, education and the well-being of residents and faculty, may be responsible. It is well known that Family Medicine residencies have been impacted greatly by the pandemic; given their generalist skills, family medicine residents were often deployed to a wide variety of settings in hospitals, giving up their usual curricula to serve in diagnostic tents, hospital wards, and ICUs. This happened even as the ACGME 2019 reductions in required faculty time deeply impacted many Family Medicine residency programs, with conversion of educational time to clinical time[3] and, in some cases, dismissal of faculty. Family Medicine residents also received less clinical experience, as documented in the decline of the continuity visit count, and a decline in the number of patients seen changed as some symptom, such as cough, were often managed elsewhere, and some rotations, such as nursing homes, were changed or eliminated.

There is also good evidence that didactic sessions—which average 4 to 6 hours a week in Family Medicine—have often had poor attendance, rarely required prework, and were interactive only approximately half the time during the pandemic.[4] Didactics represent a sizable investment of time for both residents and faculty; this data may suggest that didactic sessions are more than "nice to have"—they may be critical for learning, which is consistent with what we are learning about learning.[5] And active learning — case-based, interactive, with prework—is the gold standard for supporting retrieval and integration of knowledge. Program directors have also offered the hypothesis that, with the pandemic, residents have been reading less—that the culture of reading about cases before and after the patient has been seen has suffered. Finally, it is important to keep in mind the experience of many residents during the pandemic, from inadequate PPE, to rotations changing frequently and flux in home situations and day care. Essentially, distraction has been the rule rather than the exception in the residency learning environment of the past 3 years.

What should we do about this trend? Clearly it is important not to "blame the victims," the residents who have "leaned in" so many ways in response to the national emergency. Nor should we blame faculty, whose commitment to both patient care and to residency education has been nothing short of heroic in many settings. This is no time for "scarlet letters." At the same time, we believe it is important to insist on the importance of study and preparation. Has the availability of easy but superficial information on drug dosage and other elements of care inhibited deeper clinical understanding? Deep independent reading helps residents to learn critical thinking: to learn how to frame a clinical question, to assess information from multiple sources critically, to evaluate and prioritize available treatment options objectively, and eventually form a personalized treatment plan in shared decision making with the patient and family. In addition, to the extent that decreased scores are a marker for less effective clinical rotation curriculums, an attempt should be made to support replacement or substitution of

J Am Board Fam Med: first published as 10.3122/jabfm.2023.230092R0 on 7 April 2023. Downloaded from http://www.jabfm.org/ on 10 November 2026 by guest. Protected by copyright.

specific clinical rotations or experiences. The new residency standards give ample flexibility for this.

ABFM sees several practical next steps. From 2021 to 2022, 17% of Family Medicine residencies improved their average scores. How did they do that? ABFM will be surveying these programs and reporting on the results. We believe that our community of residency educators will have good ideas. Clearly, explicit "signposting" by program directors and faculty on the importance of reading and preparation is important. Anecdotally, some programs have taken on board preparation directly, in the form of an hour of mandatory board review every week for all residents, or giving residents with less than a 95% chance of passing on the Bayesian score predictor dedicated time each week for studying.

We also hope that ABFM tools will be useful as part of a study regime. The Continuous Knowledge Self Assessment (CKSA) tests the broad scope of care and gives instant feedback about why a right answer is correct, and why a specific answer is wrong along with references and a predicted score for the examination. The new ABFM National Journal Club allows review and questions about the clinical applicability of the 100 most important articles for family physicians from that year. The PDF of the article is a click away. Finally, Knowledge Self Assessments (KSAs) allow focused reviews of the most important diseases family physicians see as well as supporting a broad scope of care, including the care of children, the elderly and palliative care. KSAs are deliberately designed to be difficult—last year, less than 6% of diplomates passed a KSA on the first pass—but also as learning tools, allowing retaking and requiring 80% of questions correct before passing. Some residencies are beginning to use them as prework for rotations or to support group performance in didactic conferences. All these ABFM activities are free to all residents and available in the residents' ABFM portfolio.

A final question raised by some residents quietly has been, "is clinical knowledge really still important, in an age of Google and now ChatGPT and similar AI interfaces?" The answer is yes. It is important for our community to take this issue head-on. ABFM believes that what is in the family physician's "hard disk"— walking around knowledge—is critical to quality of care. Knowledge influences all aspects of the clinical encounter in real time—from the answers given to patients to the options given to the patient in shared decision making to guiding quality improvement. Importantly, whenever patients or patient representatives are asked, they are unanimous about wanting the highest knowledge possible in their personal physicians, even as they appreciate physicians taking the time to look things up. Exactly how we look up clinical questions also seems to be critical—it is clear from the ABFM's experience with family medicine certification longitudinal assessment (FMCLA) that a significant proportion of questions get *harder* when there is more time and access to information! We interpret this as ineffective looking up or lack of critical thinking about given information. ChatGPT and similar products may end up being useful, if they can get over the implicit bias that has been demonstrated in machine learning models[6] and begin to demonstrate the capacity to continuously keep up to date with the most important and practice-changing data, and develop capacity with math and images—and know better what they do not know.[7]

Of course, clinical knowledge is only 1 of the core competencies needed to provide excellent care. The others, as codified by the ABMS and ACGME 20 years ago, include communication and patient care, systems-based practice, problem-based learning and improvement, interpersonal and communication skills, and professionalism. All are important, and all are key features of family medicine residency education. But clinical knowledge remains important. Clinical practice is not just a matter of simple look-ups but also deeper understanding of principles, evidence, and their integration into all of what family physicians do. Reading and literacy are important, not only for residents but also their faculty and peers in practice. Medicine remains a learned profession.

To see this article online, please go to: http://jabfm.org/content/36/3/523.full.

References

1. Wang T, O'Neill T, Eden A, et al. Racial/ethnic group trajectory differences in exam performance among US family medicine residents. Fam Med 2022;54:184–92.
2. Middleton KV. The Longer-Term Impact of COVID-19 on K-12 Student learning and assessment. educational measurement: issues and practice 2020; 39:41–4.
3. Newton WP, Hoekzema G, Magill M, Hughes L. Dedicated time for education Is essential to the

J Am Board Fam Med: first published as 10.3122/jabfm.2023.230092R0 on 7 April 2023. Downloaded from http://www.jabfm.org/ on 10 November 2026 by guest. Protected by copyright.

residency learning environment. J Am Board Fam Med 2022;35:1035–7.
4. Zakrajsek T, Newton W. Promoting active learning in residency didactic sessions. Fam Med 2021;53:608–10.
5. Newton WP, Baxley EG, Price DW, et al. Advances in the cognitive science and their implications for ABFM knowledge assessment. J Am Board Fam Med 2022;35:878–81.
6. Chen IY, Pierson E, Rose S, Joshi S, Ferryman K, Ghassemi M. Ethical machine learning in health care. arXiv. Available at: http://arxiv.org/abs/2009.10576.20207.
7. Kung TH, Cheatham M,C, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. 2022; Available at: https://www.medrxiv.org/content/10.1101/2022.12.19.22283643v2.full-text.

J Am Board Fam Med: first published as 10.3122/jabfm.2023.230092R0 on 7 April 2023. Downloaded from http://www.jabfm.org/ on 10 November 2026 by guest. Protected by copyright.