

# COVID-19 Impact on Family Medicine Residents Exam Performance

Ting Wang, PhD<sup>a</sup>; Thomas R. O'Neill, PhD<sup>a</sup>; Lars E. Peterson, MD, PhD<sup>a,b</sup>; Warren P. Newton, MD, MPH<sup>a</sup>

## AUTHOR AFFILIATIONS:

<sup>a</sup>American Board of Family Medicine, Lexington, KY

<sup>b</sup>Department of Family and Community Medicine, College of Medicine, University of Kentucky, Lexington, KY

## CORRESPONDING AUTHOR:

Ting Wang, American Board of Family Medicine, Lexington, KY, [twang@theabfm.org](mailto:twang@theabfm.org)

**HOW TO CITE:** Wang T, O'Neill TR, Peterson LE, Newton WP. COVID-19 Impact on Family Medicine Residents Exam Performance. *Fam Med.* 2024;56(X):1–6. doi: [10.22454/FamMed.2024.719362](https://doi.org/10.22454/FamMed.2024.719362)

**PUBLISHED:** 30 January 2024

**KEYWORDS:** American Board of Family Medicine, InTraining Examination, residency training, COVID

© Society of Teachers of Family Medicine

## ABSTRACT

**Background and Objectives:** The COVID-19 pandemic began interrupting family medicine residency training in spring 2020. While a decline in scores on the American Board of Family Medicine In-Training Examination (ITE) has been observed, whether this decline has translated into the high-stakes Family Medicine Certification Examination (FMCE) is unclear. The goal of this study was to systematically assess the magnitude of COVID-19 impact on medical knowledge acquisition during residency, as measured by the ITE and FMCE.

**Methods:** A total of 19,101 initial certification candidates from 2017 to 2022 were included in this study. Annual ITE scores and FMCE scores were reported on the same scale (200–800) and served as the outcome measure. We conducted multilevel regression analysis to determine ITE score growth and FMCE scores compared to cohorts prior to COVID-19.

**Results:** During COVID-19, the increase in ITE scores from postgraduate year 2 (PGY-2) to PGY-3 was 25.5 points less, representing a 57.6% relative decrease; and from PGY-3 ITE to FMCE, it was 8.6 points less, a 12.7% relative decrease, compared with cohorts prior to COVID-19. FMCE scores were 6.6 points less during COVID-19, representing a 1.2% relative decline from the average FMCE score prior to COVID-19.

**Conclusions:** This study found nonsubstantive COVID-19 impact on FMCE scores, but a considerable knowledge acquisition decline during residency, especially during the PGY-2 to PGY-3 period. While COVID-19 impacted learning, our findings indicated that residencies were largely able to remediate knowledge deficits before residents took the FMCE.

## INTRODUCTION

During the COVID-19 pandemic (2020–2022), several national surveys revealed that family medicine residency training had experienced disruption.<sup>1–3</sup> Family medicine residencies are designed to train residents to develop the clinical skills to diagnose, manage, and treat a wide range of health conditions in various settings, including hospitals, outpatient clinics, and long-term care facilities.<sup>4</sup> However, most residents during 2020–2022 experienced changes to the content and structure of their education. For example, elective surgeries were canceled, and the number of outpatient visits decreased significantly. The traditional didactic and hands-on training styles between faculty and residents were impacted by efforts to reduce the risk of infection.<sup>2</sup> Moreover, many family medicine residents were pulled into non-curriculum-based clinical services to care for patients admitted with COVID-19, which depleted mental and physical resources they could have used for independent learning.<sup>5</sup>

Partly due to reduced learning opportunities, many professions observed decreased pass rates in their licensing examinations. For example, in nursing, the National Council Licensure Examination pass rate dropped more than 8 percentage points—from 88.2% in 2019 to 79.9% in 2022 for all first-time US-educated candidates.<sup>6</sup> The total pass rate for the Comprehensive Osteopathic Medical Licensing Examination of the United States dropped more than 3 points from 93.0% in 2019–2020 to 89.1% in 2021–2022 under the same passing standard.<sup>7</sup> While a decline in scores on the American Board of Family Medicine (ABFM) In-Training Examination (ITE) has been observed, how this decline has translated into the high-stakes Family Medicine Certification Examination (FMCE) is unclear.<sup>8</sup> In addition, we have not systematically quantified the learning loss. Previous research has shown that family medicine residents' average score increase was 30–40 points per year of residency across all subgroups prior to the COVID-19 pandemic.<sup>9</sup> The goal of this study was to assess the magnitude of COVID-19's impact on medical knowledge acquisition during residency, as measured by the ITE and FMCE.

## METHODS

### Participants

The study's participants were residents who graduated from a family medicine residency accredited by the Accreditation Council for Graduate Medical Education (ACGME) and took the FMCE between 2017 and 2022. To ensure comparability of educational experience, this study included only residents who were enrolled in a 3-year program, did not transfer to another residency, completed residency on schedule, took the ITE in each year of residency, and took the FMCE in the first administration (spring) of their postgraduate year 3 (PGY-3). This study was approved by the institutional review board of the American Academy of Family Physicians. All statistical analyses were performed in R version 4.0.2 (R Foundation).

### Demographic Data

We obtained gender, country of medical school, and medical degree demographics from ABFM administrative data sets. Self-reported race and ethnicity data are collected from the demographics section of the Initial Certification Questionnaire (ICQ), which is completed 3 to 4 months prior to the examination. Gender was categorized as either male or female. Country of medical school was dichotomized into US/Canada and international. Medical degrees included only MD and DO. To be consistent with the US Census Bureau, race and ethnicity were considered separately. Specifically, ethnicity was dichotomized as Hispanic or Latino and non-Hispanic; respondents self-selected the single race category that best described themselves: Asian, White, Black, American Indian/Alaska Native, Native Hawaiian/other Pacific Islander, or Other.

### Instruments

#### Family Medicine Certification Scale

The Family Medicine Certification Scale is the common scale that is used to describe examinee performance on several of ABFM's examinations, including the ITE and FMCE. Scores on this scale can range from 200 to 800 and are reported in increments of 10. The ABFM examinations that use this scale are built to common specifications according to the current ABFM certification examination blueprint.<sup>10</sup> Additionally, to facilitate direct comparisons, the difficulty of the questions and the estimated ability of the physicians are equated onto the same scale. Because an equated common scale was used, direct comparisons of scores are possible across exams and over time. In this study, the residents' scores were their scaled scores on the ITE and FMCE.

#### ITE

The ITE is a formative low-stakes, multiple-choice question examination intended to provide residents with the opportunity to take a test with the same look and feel as the FMCE. During the study period, the ITE consisted of 200 to 240 questions. The ITE is administered by the residency program.

### FMCE

Passing the FMCE is a requirement for earning ABFM certification; hence, it is a high-stakes examination. During the entire period of this study (2017–2022), the exam consisted of 260 scorable multiple-choice questions, and the minimum passing standard was a scaled score of 380. The FMCE is administered in a secured testing facility.

### Design

Table 1 displays the cohorts and examination calendar years for each cohort during residency. A cohort is defined as the year in which residents took their FMCE exam. Cohorts 2017–2019 are those that were prior to COVID-19, and none of their exams occurred during the COVID-19 pandemic. Cohorts 2020–2022 are those that had one to three exams impacted by COVID-19. The exam administrations that occurred under COVID-19 conditions are highlighted in Table 1. Specifically, cohort 2020 had only one exam (FMCE) impacted by COVID-19, whereas cohorts 2021 and 2022 had both ITE and FMCE exams impacted. This study period allowed for score trajectory analysis: PGY-3 for cohort 2021; PGY-2 and PGY-3 for cohort 2022.

**TABLE 1.** Calendar Exam Year From Cohort 2017 to Cohort 2022\*

Conditions	PGY-1	PGY-2	PGY-3	FMCE (Cohort)
COVID-19	2019	2020	2021	2022
COVID-19	2018	2019	2020	2021
COVID-19	2017	2018	2019	2020
Pre-COVID-19	2016	2017	2018	2019
Pre-COVID-19	2015	2016	2017	2018
Pre-COVID-19	2014	2015	2016	2017

\*Cohort is defined as the calendar year that residents are supposed to take the FMCE. Gray shading indicate exams that were impacted by COVID-19. Abbreviations: PGY, postgraduate year; FMCE, Family Medicine Certification Examination

### Analysis

As shown in Table 1, the impact of COVID-19 on exam performance should be manifested in the FMCE scores, specifically in the score trajectory during two growth periods: PGY-2 to PGY-3 and PGY-3 to FMCE. The growth period from PGY-1 to PGY-2 was excluded from this study because no cohorts were impacted by COVID-19 during that period. Because COVID-19 impacted cohorts at different exam instances, we conducted a multilevel regression analysis. We controlled for multiple exam scores from the same resident and residency programs clustering as random effects, and we controlled for residents' demographic characteristics including gender, medical degree, country of medical school, race, and ethnicity as fixed effects. Our main exposure variable was the interval between examinations occurring during the COVID-19 2020–2022 cohorts (1 as yes; 0 as no). Specifically, we coded exam levels PGY-1, PGY-2, PGY-3, and FMCE as 1, 2, 3, 4 categorical levels in case the trajectory was not linear, and we created an interaction

term between exam level and COVID-19 impacted exams. If the score growth during the two periods differed between COVID-19 impacted cohorts and prior to COVID-19 cohorts, we could conclude that COVID-19 had impact on the score growth, and the difference would inform us on the magnitude of the impact. Moreover, this coding scheme would also allow us to compare the FMCE score difference between the COVID-19 impacted cohorts and prior to COVID-19 cohorts. Finally, to control for the potential inflated Type I error caused by the large sample size in this study, we used effect size, denoted as  $\eta^2$ , to measure the magnitude of score growth across multiple comparisons. Conventionally, an effect size of  $\eta^2 \leq 0.01$  is interpreted as a small effect;  $\eta^2 \geq 0.06$ , but  $\eta^2 < 0.14$  as a medium effect, and  $\eta^2 \geq 0.14$  as a large effect.<sup>11</sup>

## RESULTS

Our analytic sample included 19,101 residents after excluding 3,724 residents with irregular progression through residency and an additional 12 residents with incomplete demographic information. The demographic information on the analytic sample is summarized in Table 2. The proportions of male and female physicians were about equal. Approximately 77% of the physicians held an MD degree, and the rest had obtained a DO degree. Two-thirds of the physicians were trained in US or Canadian medical schools and one-third in international medical schools. The proportions of race and ethnicity for each group were consistent with those found in other recent studies of residents.<sup>9</sup> Each cohort included approximately 3,000 examinees.

The exam performance of each cohort is depicted in Figure 1. The average scaled score (y-axis) of each exam (x-axis) is plotted for all cohorts. The dashed lines indicate the exam or exam trajectory that was impacted by COVID-19. The FMCE average scaled scores were similar across cohorts, but we found moderate declines in score growth during PGY-3 to FMCE for Cohort 2021 and almost no score growth during PGY-2 to PGY-3 for Cohort 2022. In multilevel regression controlling for demographic variables and clustering of scores within the resident and by program, we found that FMCE scores during COVID-19 decreased by only 6.6 points ( $P < .001$ ,  $\eta^2 < 0.01$ ), representing only 1.2% of the average FMCE score prior to COVID-19. This comparison is similar to analysis of variance but takes demographic variables and clustering into account. The pass rates were almost identical for cohorts that experienced COVID-19 during residency and those that did not (99.8% vs 99.9%,  $\chi^2 = 0.0006$ ,  $df = 1$ ,  $P = .980$ ,  $\eta^2 < 0.01$ ). However, the learning loss earlier in residency was relatively substantive, as shown in Table 3. The increase in ITE scores from PGY-2 to PGY-3 during COVID-19 was 25.5 points lower than prior years, representing a 57.6% relative decrease in knowledge gain. This learning decline also was observed in the PGY-3 to FMCE growth period, but the magnitude was smaller (8.6 scaled score, representing 12.7% of the score growth in previous cohorts).

Table 4 further demonstrates the impact of resident characteristics. We found no statistically significant difference between female and male physicians' FMCE scores, but physi-

**TABLE 2.** Characteristics of Family Medicine Residency Graduates from 2017 to 2022 (N=19,101)

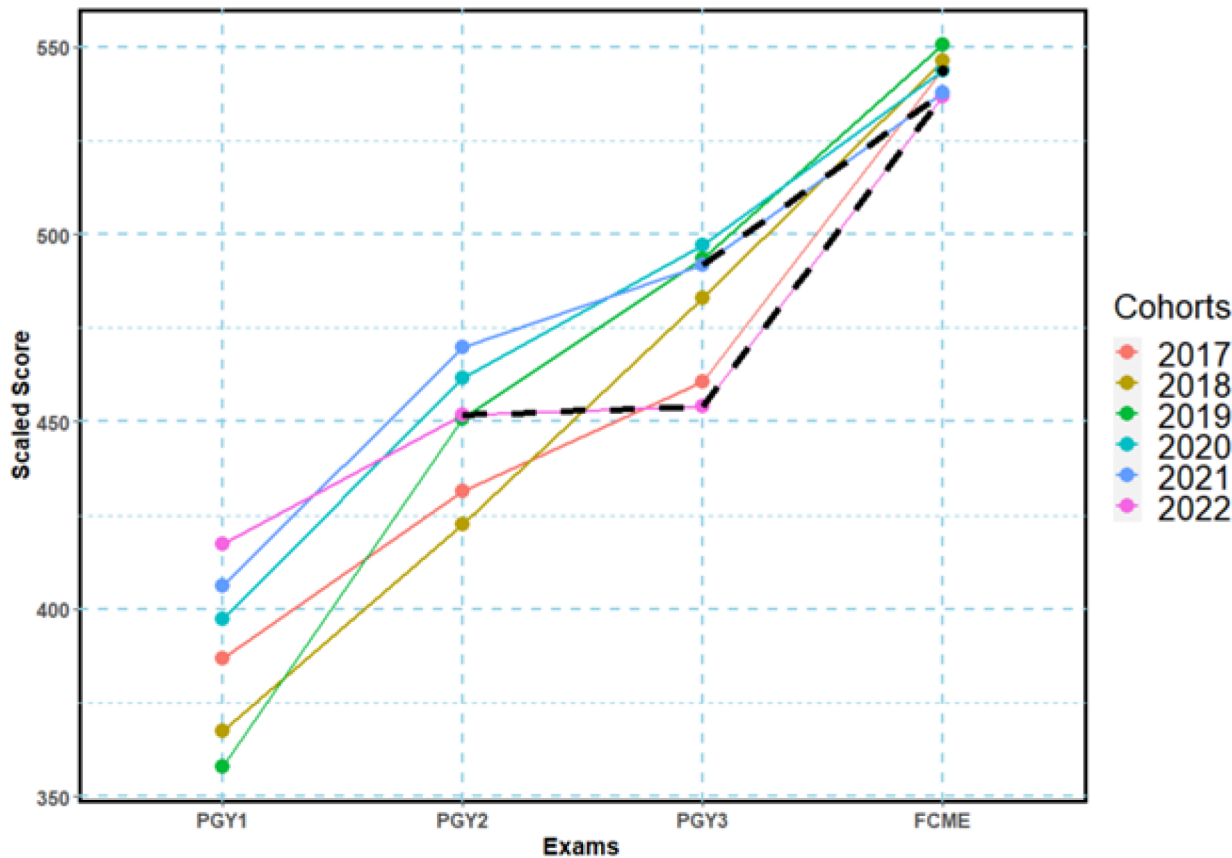
Variable	n (%)
<b>Gender</b>	
Male	8,810 (46.1)
Female	10,291 (53.9)
<b>Degree</b>	
MD	14,829 (77.6)
DO	4,272 (22.4)
<b>Medical training</b>	
US or Canada	13,177 (69.0)
International	59,24 (31.0)
<b>Race</b>	
Asian	4,296 (22.5)
Black or African American	1,472 (7.7)
White	11,608 (60.8)
Other	1,540 (8.1)
Native Hawaiian or Other Pacific Islander	60 (0.3)
American Indian or Alaska Native	125 (0.7)
<b>Ethnicity</b>	
Non-Hispanic	17,267 (90.4)
Hispanic or Latino	1,834 (9.6)
<b>Cohort</b>	
2017	2,922 (15.3)
2018	3,026 (15.8)
2019	3,000 (15.7)
2020	2,998 (15.7)
2021	3,492 (18.3)
2022	3,663 (19.2)

cians with an MD degree outperformed DO physicians. Similarly, US and Canadian medical school graduates scored higher on the FMCE than international medical school graduates. White physicians scored somewhat higher than physicians in the other groups. These results are consistent with previous literature.<sup>9</sup> Furthermore, the individual- and program-level clustering explained 47.8% and 8.6% of the total variance, indicating the necessity of using a multilevel model.

## DISCUSSION

Our analysis of more than 19,000 graduating residents' ABFM examination scores found nonsubstantive COVID-19 impact on FMCE scores but considerable declines in knowledge acquisition earlier in residency. This finding controlled for differences in resident characteristics and different learning environments of residencies, thus providing a more accurate and robust assessment of the impact of COVID-19 on residents' medical knowledge acquisition. Overall, our analysis indicated that residencies were largely able to remediate knowledge deficits due to COVID-19 prior to residents taking the high-stakes FMCE. This remediation was possible because COVID-19 cohorts (2020–2022) had higher PGY-1 scores than pre-COVID-19

FIGURE 1. Exam Score for ITE and FMCE Across Cohorts



The y-axis represents scaled score. The x-axis represents exams. Different colored lines indicate cohorts. The dashed black lines and black dots represent growth period and exams that were impacted by COVID-19.

Abbreviations: ITE, In-Training Examination; FMCE, Family Medicine Certification Examination; PGY, postgraduate year

TABLE 3. Changes in ABFM Examination Score Growth During COVID-19 Among Family Medicine Residency Graduates 2017 to 2022

	Growth in scores prior to COVID-19			Growth in scores during COVID-19			Absolute difference	Relative % change
	Estimate	P value	Effect size ( $\eta^2$ )	Estimate	P value	Effect size ( $\eta^2$ )		
PGY-2 to PGY-3	44.3	<.001	0.06	18.8	<.001	0.01	-25.5	-57.6
PGY-3 to FMCE	67.9	<.001	0.14	59.3	<.001	0.12	-8.6	-12.7

Abbreviations: ABFM, American Board of Family Medicine; PGY, postgraduate year; FMCE, Family Medicine Certification Examination

cohorts (2017–2019), as depicted in Figure 1. These elevated PGY-1 scores imply that the quality of their medical education remained largely unaffected, or at least less detrimentally impacted by the pandemic. Additionally, had they sustained a growth trajectory analogous to that of the pre-COVID-19 cohorts, they would have been expected to obtain higher FMCE scores. However, instead, they experienced two growth periods that were impacted by COVID-19: the transition from PGY-2 to PGY-3 for the 2022 cohort, and from PGY-3 to FMCE for both the 2021 and 2022 cohorts. For the 2022 cohort, residents’ scores showed almost no growth from PGY-2 to PGY-3, but their FMCE scores were compensated by higher PGY-1 scores and similar growth from PGY-3 to FMCE compared to pre-

COVID-19 cohorts. For the 2021 cohort, residents’ higher or comparable PGY-1 to PGY-3 scores, compared to pre-COVID-19 cohorts, compensated for the impact on their score growth from PGY-3 to FMCE, resulting in final FMCE scores that were still comparable to those of pre-COVID-19 cohorts. The comparability of the FMCE average scores with those of cohorts prior to the COVID-19 pandemic is noteworthy. The observed sustainability of FMCE scores has several plausible explanations. One explanation might be a potential delayed drop in the certification exam pass rate, which could be observed in the 2023 cohort that experienced their entire residency training during COVID-19. Second, the improvement in scores from PGY-3 to FMCE might be due to residency



**TABLE 4.** Adjusted Association Between Resident Characteristics and ABFM Examination Scores Controlling for Family Medicine Residency Graduates 2017 to 2022

Demographic variable	Estimate (95% CI)	P value
<b>Gender</b>		
Male	REF	
Female	0.49 (−1.10, 2.10)	.546
<b>Country of medical school</b>		
US or Canada	REF	
International	−23.3 (−21.0, −19.8)	<.001
<b>Medical degree</b>		
MD	REF	
DO	−23.7 (−25.9, −21.5)	<.001
<b>Ethnicity</b>		
Non-Hispanic	REF	
Hispanic	−16.6 (−19.5, −13.6)	<.001
<b>Race</b>		
White	REF	
American Indian or Alaska Native	−25.2 (−35.1, 15.4)	<.001
Asian	−18.2 (−20.3, −16.0)	<.001
Black or African American	−34.5 (−37.7, −31.4)	<.001
Native Hawaiian or Other Pacific Islander	−24.9 (−39.0, −10.7)	<.001
Other	−15.2 (−18.3, −12.2)	<.001

Abbreviations: ABFM, American Board of Family Medicine; REF, reference

programs' changing didactics in reaction to the drop in residents' ITE PGY-3 scores. The use of a single FMC-Scale and the Bayesian score predictor enable programs to accurately gauge the probability of future scores based on ITE results. Third, when resources are scarce, residency programs and residents might deprioritize or have less motivation to perform well on the ITE (low stakes), yet they continue to prioritize the high-stakes certification exam even under limited resource conditions. This strategy is reasonable given the circumstances.

However, the ITE score drop is significant. The ITE score drop persisted even after controlling for resident characteristics, including prior scores, and clustering of residents in residencies with different clinical and learning environments. The relative learning drop percentage in the PGY-2 to PGY-3 period was more than 50%. In other words, this learning drop occurred during COVID-19 regardless of residents' personal characteristics. Additional analysis further revealed that the learning drop occurred at a consistent rate across all race/ethnicity groups. The interaction between race/ethnicity and slope was not significant and is not shown in the main results. This consistency indicates that no differential effect learning drops were found with respect to race and ethnicity. This finding is consistent with our previous study, which showed that the learning trajectory is similar across different race and ethnicity groups during residency.<sup>9</sup>

Indeed, a significant number of first-year residents, entering between 2020 and 2022, were not adequately prepared for clinical duties, primarily due to restricted training and experience during this period.<sup>12</sup> This pattern has been noted consistently across various medical specialties throughout the nation. Program directors have consistently reported relatively low confidence in the readiness of new PGY-1 residents to assume specific tasks expected of them at the onset of their residency training.<sup>13</sup> This lack of experience may not be reflected in the final high-stakes certification exams, but may manifest in residents' competency. The former focuses only on medical cognitive decision-making in applied settings, whereas the latter includes not only medical knowledge, but also skills and attitudes essential for providing high-quality care, such as communication skills, teamwork skills, system-based practice, and professionalism. In the long-term, the disruption of COVID-19 may serve as a reminder for the profession to continue implementing competency-based medical education.<sup>14,15</sup> The COVID-19 pandemic is only one form of interruption. Given an interruption, if the goal is to provide a primary care workforce of similar competency, a redesign<sup>16,17</sup> of family medicine residency education to provide some additional learner-centered training might be necessary.

This study had several limitations. First, the results are applicable only to residents who went through residency training with regular progression. Although we could not compare scores for residents with irregular progression or missing scores, the percentage of regular versus irregular progression was relatively stable across cohorts (85%–90% vs 15%–10%). Second, medical knowledge is only one component of all the experience residents obtain from training, and other clinical experience was not measured in this study. Third, more learning metrics/tools need to be developed to measure residents' learning and encourage learning behavior during residency. Measurement is crucial for the practical execution of competency-based medical education across all residency programs nationally in the future.

## CONCLUSIONS

In summary, we found that disruptions in family medicine residency due to COVID-19 decreased the expected score growth more than half from PGY-2 to PGY-3, but the FMCE scores were not impacted substantively. While COVID-19 impacted learning, our findings indicate that residencies were largely able to remediate knowledge deficits before residents took the FMCE.

## REFERENCES

1. Awadallah NS, Czaja AS, Fainstad T. The impact of the COVID-19 pandemic on family medicine residency training. *Fam Pract*. 2021;38(suppl1):9–15.
2. Hogan SO, Holmboe ES. Effects of COVID-19 on residency and fellowship training: results of a national survey. *J Grad Med Educ*. 2022;14(3):359–364.
3. Chen SY, Lo HY, Hung SK. What is the impact of the COVID-19 pandemic on residency training: a systematic review and

- analysis. *BMC Med Educ.* 2021;21(1):618.
4. Carek PJ, Anim T, Conry C. Residency training in family medicine. *Fam Med.* 2017;49(4):275–281.
5. Messner E, Parascando JA, Gamma L, et al. Effects of early COVID-19 restrictions on resident well-being and burnout. *Fam Med.* 2022;54(9):708–712.
6. NCLEX pass rates drop again. *Am J Nurs.* 2023;123(4):14.
7. COMLEX-USA examination performance data. *National Board of Osteopathic Medical Examiners.* 2023. <https://www.nbome.org/assessments/comlex-usa/performance>.
8. Newton WP, Wang T, O'Neill TR. The decline in family medicine In-Training Examination scores: what we know and why it matters. *J Am Board Fam Med.* 2023;36(3):523–526.
9. Wang T, O'Neill TR, Eden AR, et al. Racial/ethnic group trajectory differences in exam performance among US family medicine residents. *Fam Med.* 2022;54(3):184–192.
10. Norris TE, Rovinelli RJ, Puffer JC, Price DW. From specialty-based to practice-based: a new blueprint for the American Board of Family Medicine cognitive examination. *J Am Board Fam Pract.* 2005;18(6):546–554.
11. Cohen J. *Statistical Power Analysis for the Behavioral Sciences (revised edition).* L. Erlbaum Associates; 1987. .
12. Schrager S. Residency training during a pandemic. *Fam Med.* 2022;54(9):681–682.
13. Howley L, Grbic D, Speicher MR, Roskovensky LB, Jayas A, Andriole DA. The resident readiness survey: a national process for program directors to provide standardized feedback to medical schools about their graduates. *J Grad Med Educ.* 2023;15(5):572–581.
14. Holmboe ES. The transformational path ahead: competency-based medical education in family medicine. *Fam Med.* 2021;53(7):583–589.
15. Van Melle E, Frank JR, Holmboe ES, Dagnone D, Stockley D, Sherbino J. International competency-based medical education collaborators. a core components framework for evaluating implementation of competency-based medical education programs. *Acad Med.* 2019;94(7):1,002–1,009.
16. Newton WP, Magill M. It takes a village to redesign residencies. *J Am Board Fam Med.* 2022;35(2):445–448.
17. Newton WP, Mitchell KB, Magill MK. The future of family medicine residency education: the specialty has spoken. *Ann Fam Med.* 2021;19(2):185–187.