JACEP OPEN WILEY

ORIGINAL RESEARCH

Education

Comparison of performance data between emergency medicine 1-3 and 1-4 program formats

Michael S. Beeson MD, MBA[1] | Melissa A. Barton MD[2] | Earl J. Reisdorff MD[2] [ID] |
Wallace A. Carter MD[3] | Marianne Gausche-Hill MD[4] | Diane L. Gorgas MD[5] |
Kevin B. Joldersma PhD[2] | Sally A. Santen MD, PhD[6]

[1]Department of Emergency Medicine, Summa Health, Akron, Ohio, USA

[2]American Board of Emergency Medicine, East Lansing, Michigan, USA

[3]Department of Emergency Medicine, Weill Cornell Medicine, New York, New York, USA

[4]Department of Emergency Medicine, Harbor-University of California Los Angeles, Los Angeles County Emergency Medical Services Agency, Los Angeles, California, USA

[5]Department of Emergency Medicine, The Ohio State Wexner Medical Center, Columbus, Ohio, USA

[6]Department of Emergency Medicine, University of Cincinnati College of Medicine, Cincinnati, Ohio and Department of Emergency Medicine, Virginia Commonwealth University, Richmond, Virginia, USA

Correspondence
Earl J. Reisdorff, MD, American Board of Emergency Medicine, East Lansing, 3000 Coolidge Rd., East Lansing, MI 48823, USA.
Email: ereisdorff@abem.org

Abstract

Objective: This study compares performance data from physicians completing 3-year versus 4-year emergency medicine residency training programs. Currently, there are 2 training formats and little is known about objective performance differences.

Methods: This was a retrospective cross-sectional analysis of emergency residents and physicians. Multiple analyses were conducted comparing physicians' performances, including Accreditation Council of Graduate Medical Education Milestones and American Board of Emergency Medicine In-training Examination (ITE), Qualifying Examination (QE), Oral Certification Examination (OCE), and program extensions from 3-year and 4-year residency programs. Some confounding variables were not or could not be considered, such as rationale for medical students to choose one format over another, as well as application and final match rates.

Results: Milestone scores are higher for emergency medicine 3 residents in 1-3 programs (3.51) versus emergency medicine 3 residents in 1-4 programs (3.07; $P < 0.001$, $d = 1.47$) and highest for emergency medicine 4 residents (3.67). There was no significant difference in program extension rates (emergency medicine 1-3, 8.1%; emergency medicine 1-4, 9.6%; $P = 0.05$, $\omega = 0.02$). ITE scores were higher for emergency medicine 1, 2, and 3 residents from 1-3 programs and emergency medicine 4 residents from 1-4 programs scored highest. Mean QE score was slightly higher for emergency 1-3 physicians (83.55 vs 83.00; $P < 0.01$, $d = 0.10$). QE pass rate was higher for emergency 1-3 physicians (93.1% vs 90.8%; $P < 0.001$, $\omega = 0.08$). Mean OCE score was slightly higher for emergency 1-4 physicians (5.67 vs 5.65; $P = 0.03$, $d = -0.07$) but did not reach a priori statistical significance ($\alpha < 0.01$). OCE pass rate was also slightly higher for emergency 1-4 physicians (96.9% vs 95.5%; $P = 0.06$, $\omega = -0.07$) but also non-significant.

Conclusions: These results suggest that although performance measures demonstrate small differences between physicians from emergency medicine 1-3 and 1-4 programs,

---

these differences are limited in their ability to make causal claims about performance on the basis of program format alone.

1 | INTRODUCTION

1.1 | Background

The length of training needed to ensure that a physician has the knowledge, skills, and abilities to provide independent, safe, high-quality emergency care has been a topic of discussion throughout the history of emergency medicine residency training. There has been an increase in the breadth and depth of emergency medicine practice since its foundation as an independent specialty. Emergency medicine is an unusual specialty with an established history of having a variable residency training duration.

Emergency medicine program requirements and accreditation were first established through the Liaison Residency Endorsement Committee (LREC) between 1975 and 1976, which was composed of members from the American College of Emergency Physicians, University Association for Emergency Medicine (now the Society for Academic Emergency Medicine), and the American Medical Association. At its inception, the LREC approved formats of emergency medicine residencies that were already in place, that is, postgraduate year (PGY) 1–3, PGY 2–3, and PGY 2–4. Several PGY 2–3 and PGY 2–4 programs were similar to fellowships and were designed to retrain physicians from other specialties. LREC continued in that role until 1981, when the Accreditation Council for Graduate Medical Education (ACGME) was formed.

In 1980, the Review Committee for Emergency Medicine was established under the auspices of ACGME, and the American Board of Emergency Medicine (ABEM) moved to require 36 months of graduate medical education, of which 24 months must have occurred in an ACGME-accredited emergency medicine program. The PGY 2–3 format was permitted until 1985, when ABEM stated that all 36 months of training must occur in an ACGME-accredited program. The PGY 2–4 format, popular with military programs, ended in 2011, with the belief that a prerequisite year was no longer necessary to enter a 3-year residency, as there were ACGME-accredited emergency medicine 1-3 programs. Emergency medicine 1-4 programs became the only acceptable 4-year training format in 1988. Starting the transition in 2015 and completed in 2020 under the single accreditation system, the ACGME serves as the nation's sole accreditor for both osteopathic and allopathic residencies and fellowships. This change, among others, increased the number of emergency medicine residency programs. As of December 2021, there were 276 residencies, of which 221 (80%) were 1-3 programs and 55 (20%) were 1-4 programs.[1]

Although competency-based medical education is gaining increasing consideration as an educational framework, the ACGME through individual specialty program requirements still specifies length of training.

1.2 | Importance

Little evidence exists that compares physician performance in the 2 residency formats. ABEM has an interest in physician performance "to ensure the highest standards in the specialty of Emergency Medicine."[2] Furthermore, any difference in physician performance that was associated with residency duration would also be of interest. A survey of emergency medicine residency program directors indicated that the preferred length of training was predicated largely on the duration of the program that the program directors trained in and currently led.[3] The investigators concluded that, "PD [program director] opinions may be too biased by their own experiences to provide objective input on the ideal LoT [length of training] for emergency medicine residency programs."[3]

1.3 | Goals of this investigation

The purpose of this study was to compare performance data from physicians completing 3-year versus 4-year emergency medicine residency training programs.

2 | METHODS

ABEM undertook a review of 7 different parameters to compare programs that included physician performance related to ACGME emergency medicine Milestone acquisition, standardized test performance, and program extensions for graduating residents.

2.1 | Study design and setting

This was a retrospective cross-sectional analysis of emergency residents and physicians. Multiple analyses were conducted comparing physicians' performances, including ACGME Milestones and ABEM In-training Examination (ITE) and certification Examinations (ie, Qualifying Examination [QE] and Oral Certification Examination [OCE]),

from 3-year and 4-year residency programs. This study received an exemption for human subject research from the Summa Health System Institutional Review Board, Akron, Ohio.

2.2 | Selection of participants

The sample was cross-sectional. Specific comparison analyses were performed for the following samples. ITE scores compared emergency medicine 1-3 and 1-4 residents' scores by training year for the 2021 ITE examination ($n = 8491$). Milestone scores compared program format on June 2021–reported Milestones ($n = 8649$). Mean QE score and the aggregate pass rate for all emergency medicine 1-3 graduates were compared with mean scores for all emergency medicine 1-4 graduates for 3 years (2018, 2019, and 2020; $n = 6,682$). Mean scaled OCE score and aggregate pass rate for all emergency medicine 1-3 graduates were compared with the mean scaled score for all emergency medicine 1-4 graduates for 3 years (2018, 2019, and 2020; $n = 4,215$). The frequency of program extensions was compared between emergency medicine 1-3 and emergency medicine 1-4 programs for 3 years (2018, 2019, and 2020; $n = 7028$) and in aggregate.

2.3 | Measurements

The ACGME Milestones serve as objective competency measures that should be achieved throughout training.[4] For the emergency medicine Milestones, the 6 ACGME competencies are delineated into 23 subcompetencies, with 227 Milestones across 5 proficiency levels.[5,6] Each subcompetency is scored on a 10-point scale, which consists of half-point increments ranging from 0.5 to 5.0. Level 1 reflects the competency expected of a medical school graduate entering residency. Level 4 is the recommended performance for a graduating emergency medicine resident, and level 5 is expected to be achieved after years of clinical practice. Residency programs report Milestones for all residents to the ACGME in December and June of each year. These data are then securely transferred to ABEM. Only the June Milestones were compared. This study used Milestones 1.0 because version 2.0 was released in 2021–2022.[7,8]

The ABEM ITE is a secure, computer-delivered examination administered annually in late February. The examination has significant supporting validity evidence for its use as an objective measure of medical knowledge and cognitive skill that predicts a physician's likelihood of passing the written ABEM QE.[9] Scores are statistically equated and reported as scaled scores (0–100).

The ABEM QE is a secure, computer-delivered examination administered once per year in late November or early December in high-security testing centers around the country. The examination is well researched as an assessment of emergency medicine knowledge[10–12] and its application in a written format. Scores are statistically equated and reported as scaled scores (0–100). The passing scaled score is 77 and is a criterion standard. This written assessment is followed by an oral examination in ABEM's certification process.

The Bottom Line

The outcomes of training compared favorably between emergency medicine 1-3 and 1-4 program formats. The results suggest that there are uncovered differences between the programs but that at graduation, the residents from these 2 cohorts show about the same results in their In-Training Examinations, Qualifying Examinations, Oral Certifying Examinations, and Milestones scores. However, there are many confounding variables that should be explored in future research on the differences between these 2 program formats.

The ABEM OCE is the specialty's initial certification test. Historically, the examination was an in-person examination consisting of 7 cases delivered by an ABEM-trained examiner. The examination has been studied thoroughly and has been found to have sufficient validity evidence and reliability support to make summative certification decisions.[13–18] During the COVID-19 pandemic, ABEM converted the administration of the OCE to a virtual platform. OCE results from the virtual oral examination format are excluded from this investigation. The OCE passing score is not a static number. The OCE passing score is adjusted by periodic standard setting, which is a psychometric best practice when the assessment content or format changes for a specialty. In addition, the OCE is statistically equated and reported as a scaled score (1.00–8.00). Equating is necessary to create fairness across examinations, thus compensating for exam-to-exam variations in test difficulty.

Residents who have medical, personal, or academic issues may extend their residency training past the standard academic year. These incidents are reported routinely to ABEM as part of its training verification and are included in the analysis. Although the causes for program extension vary (eg, parental leave and academic remediation), this study did not attempt to differentiate subgroups.

2.4 | Analysis

Multiple analyses were performed comparing available data on ACGME Milestones and ABEM ITE, QE, and OCE scores for residents from 3-year and 4-year programs. Preliminary analyses demonstrated sufficiently normal distributions to use parametric analyses. Data analyses, including mean, SD, Student t-test, chi-square test, and Cohen's effect sizes (d for means, ω for chi-square), were performed using R analysis software (R Core Team, version 3.6.2, Vienna, Austria). Effect sizes were interpreted as small (0.20), medium (0.50), and large (0.80).[19] On the basis of robust sample sizes for all analyses and to avoid capitalization on chance, the statistical significance was determined a priori to be $\alpha = 0.01$. No correction for multiple comparisons was added because the analyses consist of simple mean comparisons.

TABLE 1 2021 June Accreditation Council for Graduate Medical Education milestones scores by program format.

Emergency medicine year	Emergency medicine1-3 program		Emergency medicine1-4 program		Student t test (P value)[a]	Effect size (d)
	No. of residents	Mean score (95% CI)	No. of residents	Mean score (95% CI)		
1	2117	1.63 (1.61–1.65)	672	1.45 (1.42–1.48)	<0.001	0.45
2	2003	2.64 (2.62–2.66)	670	2.32 (2.29–2.35)	<0.001	0.85
3	1894	3.51 (3.50–3.53)	650	3.07 (3.05–3.09)	<0.001	1.47
4	–	–	643	3.67 (3.65–3.69)	NA[b]	NA

Abbreviations: CI, confidence interval; NA, not applicable.

[a]Statistical significance was determined a priori to be $\alpha < 0.01$.

[b]t test comparison of emergency medicine 3 in 1-3 programs and emergency medicine 4 residents showed a statistically significant difference; $P < 0.001$; Cohens $d = -0.56$ (moderate).

TABLE 2 Mean 2021 American Board of Emergency Medicine in-training examination scaled scores by program format.

Emergency medicine year	Emergency medicine 1-3 programs		Emergency medicine 1-4 programs		Student t test (P value)[a]	Effect size (d)
	No. of residents	Mean score (95% CI)	No. residents	Mean score (95% CI)		
1	2127	70.2 (69.9–70.5)	624	68.2 (67.6–68.8)	<0.001	0.28
2	1996	76.6 (76.3–76.9)	630	73.8 (73.2–74.4)	<0.001	0.38
3	1861	79.7 (79.4–80.0)	616	78.1 (77.5–78.7)	<0.001	0.22
4	NA	NA	637	80.3 (79.6–80.9)	NA[b]	NA

Abbreviations: CI, confidence interval; NA, not applicable.

[a]Statistical significance was determined a priori to be $\alpha < 0.01$.

[b]t test comparison of emergency medicine 3 in 1-3 programs and emergency medicine 4 residents showed no statistical difference; $P = 0.1$; Cohens $d = -0.08$ (negligible).

3 | RESULTS

3.1 | Milestone ratings

June Milestone ratings comparing each year by format showed that the Milestone ratings were universally higher in emergency medicine 1-3 programs as compared to 1-4 programs with an equivalent number of years of training. (Table 1) This difference in emergency medicine 1 was 1.63 in emergency medicine 1-3 format versus 1.45 in 1-4 ($P < 0.001$, $d = 0.45$). For emergency medicine 2 residents, the difference was 2.64 in emergency medicine 1-3 programs versus 2.32 for residents of 1-4 programs ($P < 0.001$, $d = 0.85$). Emergency medicine 3 residents in emergency medicine 1-3 programs had a mean of 3.51 versus 3.07 for residents in 1-4 programs ($P < 0.001$, $d = 1.47$). The highest ratings were received by emergency medicine 4 residents in 1-4 programs (3.67). The effect size is small at emergency medicine 1 but becomes large at emergency medicine 2 and 3 years of training.

3.2 | Examination scores

ITE scores followed a similar pattern, with higher scores for emergency medicine 1, 2, and 3 residents from 1-3 programs, and emergency medicine 4 residents from 1-4 programs scoring the highest (Table 2).

All mean differences were statistically significant and the effect sizes ranged from 0.22 to 0.38; all medium in magnitude. For board certification examinations (Table 3), QE scores were slightly higher for emergency 1-3 physicians compared with emergency 1-4 physicians (83.55 vs 83.00, respectively; $P < 0.001$, $d = 0.10$). The QE pass rate was higher for emergency 1-3 physicians than for emergency 1-4 physicians (93.1% vs 90.8%, respectively; $P < 0.001$, $\omega = 0.08$). In both cases of significant differences, the magnitude of the effect size is small. Mean OCE score was slightly higher for emergency 1-4 physicians than for 1-3 physicians (5.67 vs 5.65, respectively; $P = 0.03$, $d = -0.07$) but did not reach a priori statistical significance ($\alpha < 0.01$). Mean OCE pass rate was also slightly higher for emergency medicine 1-4 physicians (1-3, 95.5%; 1-4, 96.9%; $P = 0.06$, $\omega = -0.07$) but not statistically significant.

3.3 | Program extensions

Program extension rates did not vary between emergency medicine residency duration formats for years 2018 and 2019. In 2020, however, the program extension rate for emergency medicine 1-4 programs (11.6%) was statistically significantly higher than for 1-3 programs (7.1%; $P < 0.01$, $\omega = 0.01$). The effect size was extremely small at 0.01. The comparison of program extensions across all years studied did not reach a priori statistical significance ($\alpha < 0.01$) (Table 4).

BEESON ET AL.

JACEP OPEN
JOURNAL OF THE AMERICAN COLLEGE OF EMERGENCY PHYSICIANS Open Access
A GLOBAL JOURNAL OF EMERGENCY MEDICINE

WILEY | 5 of 8

TABLE 3 American Board of Emergency Medicine Certification Examinations by program format.

Mean qualifying examination scaled scores (2018–2020)				
Emergency medicine 1-3 programs		**Emergency medicine 1-4 programs**		
No. of physicians	**Mean score (95% CI)**	**No. of residents**	**Mean score (95% CI)**	**P value,[a] effect size (d)**
4763	83.55 (83.4–83.7)	1919	83.00 (82.8–83.2)	<0.001, 0.10
Qualifying examination pass rates (2018–2020)				
Passed, n (%)	**Failed, n (%)**	**Passed, n (%)**	**Failed, n (%)**	**P value,[a] effect size (ω)**
4435 (93.1)	328 (6.9)	1743 (90.8)	176 (9.2)	<0.001, 0.08
Oral certification examination scaled scores (2018–2020)				
No. of residents	**Mean score (95% CI)**	**No. of residents**	**Mean score (95% CI)**	**P value,[a] effect Size (d)**
3095	5.65 (5.64–5.66)	1120	5.67 (5.65–5.68)	0.03, −0.07
Oral certification examination pass rates (2018–2020)				
Passed, n (%)	**Failed, n (%)**	**Passed, n (%)**	**Failed, n (%)**	**P value,[a] effect size (ω)**
2956 (95.5)	139 (4.5)	1085 (96.9)	35 (3.1)	0.06, −0.07

Abbreviation: CI, confidence interval.
[a]Statistical significance was determined a priori to be $\alpha < 0.01$.

TABLE 4 Rates of program extensions (2018–2020).

Year	Emergency medicine 1-3 programs		Emergency medicine 1-4 programs		Chi-square (P value), (effect size;(ω))[a]
	On time completion, n (%)	**Delayed completion, n (%)**	**On time completion, n (%)**	**Delayed completion, n (%)**	
2018	1372 (91.0)	136 (9.0)	617 (90.3)	66 (9.7)	0.63, (0.01)
2019	1499 (91.6)	138 (8.4)	642 (92.5)	52 (7.6)	0.45, (0.01)
2020	1694 (92.9)	129 (7.1)	604 (88.4)	79 (11.6)	<0.001, (0.01)
Total	4565 (91.9)	403 (8.1)	1863 (90.4)	197 (9.6)	0.05, (0.02)

[a]Statistical significance was determined a priori to be $\alpha < 0.01$.

4 | LIMITATIONS

First, many of the metrics presented in this study focused on performance in standardized testing situations. These are, inherently, point-in-time measurements that have validity for a variety of purposes but are generally associated with other board certification outcomes and quality measures. For tests that have predictive validity, test performance on 1 standardized test is associated with similar performance on other standardized tests. However, the QE and OCE have not been analyzed to determine the strength of correlation with pre-medical and medical examinations.

Second, although Milestones data showed trends that have been consistent since their first use, Milestones data also have limited validity support for predicting the ultimate quality of the clinically active physician. In addition, there are known variations in Milestone ratings at the program level, with, at times, idiosyncratic scoring. It is unlikely that these idiosyncrasies carry through to an aggregate level, but it is possible that the Milestones may covary by program format.

Third, program format selection by the physicians was not random and might have a biasing impact on study results and outcomes. It is not clear why applicants to residency may choose one format over another. Program selection could have influenced wellness considerations (other than program extension data), indebtedness, academic fellowship preparation, scholarly productivity, or perceived future employment opportunities. Moreover, it is possible that medical students with less academic proficiency could preferentially select emergency medicine 1-4 programs. We did not compare the United States Medical Licensing Examination scores or other academic performance criteria to determine if that selection bias was occurring. Any impact on performance by format selection remains an area for further investigation.

Fourth, the different analyses involved discontinued cohorts. Specifically, the data from each outcome (Milestones, ABEM exam scores, and ABEM records of program extensions) were cross-sectional cohorts for time-specific periods and therefore did not represent longitudinal measurement. This segregation of cohort analyses limits the ability to form a composite model of resident performance based on training format.

6 of 8 | **WILEY** JACEP OPEN
JOURNAL OF THE AMERICAN COLLEGE OF EMERGENCY PHYSICIANS OPEN
A GLOBAL JOURNAL OF EMERGENCY MEDICINE   Open Access

BEESON ET AL.

Each analysis should be regarded as a unique finding with limited to no generalizability to any other analysis.

This study does not make any attempt to explore whether program format differences have future success in continued board activities. It also does not explore whether program format has any subsequent link to differences in clinical outcomes.

There is considerable current interest among all specialties in shifting away from time-defined residency training to a competency-based medical education framework. This study does not help to argue one framework over another.

5 | DISCUSSION

The merits of emergency medicine 1-3 versus 1-4 program duration have been debated since the beginning of accredited training programs.[3,20–22] Various rationales for both have been provided, including the increased core content of emergency medicine over time; student indebtedness increasing because of longer training formats; academic development associated with additional training; and most recently, workforce issues within emergency medicine.[23,24] This study evaluated objective measures of performance differences using Milestone acquisition, standardized test performance, and program extensions.

Despite the long-standing existence of residencies of varied durations and formats in emergency medicine, there is a paucity of data on objective performance outcome differences between graduates of the 2 types of emergency medicine residency programs. Performance outcome differences may be affected by factors that led to program type selection. Hopson et al. described emergency medicine program directors' views of program length; however, their study was limited by program directors' subjective preferences regarding program length.[3] Our results help to fill the gap in the literature and to advance the discussion of the merits of different residency lengths in emergency medicine.

This study used ACGME Milestones data and ABEM test performance data. These results do not provide sufficient evidence to make a confident determination of the superiority of one training duration compared with the other. An objective discussion of program format differences must continue to be a community-wide discussion, including whether a single format is necessary or desirable. Other objective measures may prove to be more relevant, such as cost data, admission/discharge data, morbidity/mortality data, and so forth.

5.1 | Milestone acquisition

Milestone scores were modestly higher for emergency medicine 2 and 3 residents in emergency medicine 1-3 programs compared with those for emergency medicine 2 and 3 residents in emergency medicine 1-4 programs. This does not necessarily mean that residents in emergency medicine 1-3 programs achieve Milestone subcompetencies faster compared with residents in emergency medicine 1-4 programs. Emer-

gency medicine 1-4 programs may intentionally design their curricula so that there is a more gradual acquisition of subcompetencies. This may reflect the organic structural differences that an additional year for emergency medicine 1-4 programs have. Milestone determination is independent of ABEM-designed standardized tests and reported directly to the ACGME.

Emergency medicine 4 residents achieved a mean Milestone score of 3.67, versus 3.51 for emergency medicine 3 residents in emergency medicine 1-3 programs. This result suggests that physicians have greater proficiency in Milestone subcompetencies at the time of graduation from emergency medicine 1-4 programs. Despite this, neither population achieved level 4 in the aggregate, which, for individual Milestones in general, is the recommended performance for a graduating emergency medicine resident. The clinical significance of this greater Milestone score difference is unknown. Milestone scores are reported by each program, following general scoring anchors provided by the ACGME as guidelines. This may introduce some subjectivity into scoring, with Milestone levels spread out over a greater time frame in emergency medicine 1-4 programs. Likewise, the addition of an additional year may allow emergency medicine 1-4 programs to spread Milestone acquisition over a longer time frame rather than in a compressed emergency medicine 1-3 format. One challenge in understanding the highest scores for emergency medicine 4 residents is there is no comparison with emergency medicine 3 graduates in their first year of practice. Such a comparison would potentially provide a more direct postgraduate-year comparison. Milestone acquisition at graduation may or may not inform discussions about the value of an additional year of emergency medicine residency training.

5.2 | Standardized test performance

Standardized test performance data of physicians from different training formats vary. The data suggest progression in performance in both training formats. Specifically, ITE scores demonstrated a small performance advantage for residents in emergency medicine 1-3 programs when comparing matched levels of training. This finding is not surprising, given that the core content of EM is likely distributed differently over the different time periods (ie, residents in emergency medicine 1-4 programs presumably have 4 years to acquire medical knowledge of emergency medicine, instead of 3 years). There may be intrinsic differences in the pool of residents in each residency format that may account for this minor difference, such as 1 format attracting different qualities in applicants. Some confounding variables were not or could not be considered, such as rationale for medical students to choose one format over another, as well as application and final match rates. Emergency medicine 4 residents had a mean score of 80.3 versus 79.7 for emergency medicine 3 residents in 1-3 programs, demonstrating similar ITE scores at the time of graduation regardless of program format.

For the QE, emergency medicine 1-3 graduates scored modestly better compared with 1-4 graduates (83.55 vs 83.00, respectively;

$P < 0.001$, $d = 0.10$). The clinical significance of this performance difference is unknown and the effect size is very small. The overall pass rate for the QE favored graduates from emergency medicine 1-3 programs (93.1% vs 90.8%; $P < 0.001$, $\omega = 0.08$), and although significant, is a small effect size. This 2.3% difference has unknown clinical significance, although this represents the percentage of individuals who have not gained passage into the next step of the ABEM Board Certification pathway: the OCE.

There were minimal differences in performance scores on the OCE. The mean overall pass rate favored graduates from emergency medicine 1-4 programs to a small extent, but the difference was not statistically significant (96.9% vs 95.5%; $P = 0.06$). Because the differences are non-significant, the effect sizes are irrelevant.

5.3 | Program extensions

Program extensions occur for many reasons, including academic difficulty, family leave, personal illness, family-related events, and other life events.[25,26] Although there were relatively more program extensions in emergency medicine 1-4 programs across all years studied, this difference was not statistically significant. This difference might be expected, given the longer length of the program and the contribution of maternity leave to program extension. In 2020 there was an increase in delayed completion in emergency medicine 1-4 programs compared to 1-3 programs.

6 | CONCLUSION

Little is known about whether a program's duration might influence a medical student's choice of program. One study by Laskey and Cydulka listed 18 different factors affecting residency selection, none of which was program duration.[27] The most important factors influencing selection were (1) reputation of the institution, (2) hospital facilities, and (3) reputation of the program director. As the specialty continues to engage in this discussion, additional considerations should include the validity of training format based on the breadth of the core content to be mastered, costs to trainees, and differences in institutional support.

Graduates of emergency medicine 1-3 programs performed statistically significantly better on the ABEM QE in terms of pass rates and scores; physicians from emergency medicine 1-4 programs scored statistically significantly higher on the OCE, but their pass rate was not statistically significantly higher. In all cases, the effect sizes of scores or pass rate differences would be categorized as "trivial" using Cohen's classification system.[18] These results suggest that although performance measures demonstrate small differences between physicians from emergency medicine 1-3 and 1-4 programs, these differences are limited in their ability to make causal claims about program format. These differences can help to inform discussions about program format, curriculum design and delivery, expected performance, and test preparation advising to benefit residents in all training formats. Finally, the discussion of time-based formats for emergency medicine residency programs will continue to evolve as competency-based education models develop.

CONFLICT OF INTEREST STATEMENT

Drs. Carter, Gorgas, and Gausche-Hill are on the Board of Directors of the American Board of Emergency Medicine (ABEM); Dr. Beeson is a former director for ABEM; Dr. Santen is an oral examiner volunteer for ABEM and consultant for the American Medical Association; and Drs. Joldersma, Barton, and Reisdorff are employees of ABEM.

ORCID

Earl J. Reisdorff MD https://orcid.org/0000-0003-3553-446X

REFERENCES

1. Nelson LS, Calderon Y, Ankel FK, et al. American Board of Emergency Medicine report on residency and fellowship training information (2021-2022). *Ann Emerg Med.* 2022;80(1):74-83. doi:10.1016/j.annemergmed.2022.05.014
2. American Board of Emergency Medicine (ABEM). "About ABEM". Available online at: https://abem.org/public/about-abem/[Last accessed 9/8/2022]
3. Hopson L, Regan L, Gisondi MA, et al. Program director opinion on the ideal length of residency training in emergency medicine. *Acad Emerg Med.* 2016;23(7):823-827. doi:10.1111/acem.12968
4. Korte RC, Beeson MS, Russ CM, et al. The emergency medicine milestones: a validation study. *Acad Emerg Med.* 2013;20(7):730-735. doi:10.1111/acem.12166
5. Beeson MS, Holmboe ES, Korte EC, et al. Initial validity analysis of the emergency medicine milestones. *Acad Emerg Med.* 2015;22(7):838-844.
6. Beeson MS, Carter WA, Christopher TA, et al. The development of the emergency medicine milestones. *Acad Emerg Med.* 2013;20(7):724-729.
7. Cooney RR, Murano T, Ring H, et al. The Emergency Medicine Milestones 2.0: setting the stage for 2025 and beyond. *AEM Educ Train.* 2021;5(3):e10640. doi:10.1002/aet2.10640
8. Accreditation Council for Graduate Medical Education. Emergency Medicine Milestones. Accreditation Council for Graduate Medical Education. 2021. Available online at: [Last accessed 7/29/2022]. https://www.acgme.org/globalassets/pdfs/milestones/emergencymedicinemilestones.pdf
9. American Board of Emergency Medicine. *In-training Examination Score Reports. Frequently Asked questions (FAQS).* American Board of Emergency Medicine. 2023. Availableonline at: [Last accessed 7/29/2022].

10. Caffery T, Fredette J, Musso MW, et al. Predicting American Board of Emergency Medicine qualifying examination passage using United States Medical Licensing Examination step scores. *Ochsner J.* 2018;18(3):204-208. doi:10.31486/toj.17.0101

11. Counselman FL, Kowalenko T, Marco CA, et al. The relationship between ACGME duty hour requirements and performance on the American Board of Emergency Medicine qualifying examination. *J Grad Med Educ.* 2016;8(4):558-562.

12. Marco CA, Counselman FL, Korte RC, et al. Delaying the American Board of Emergency Medicine qualifying examination is associated with poorer performance. *Acad Emerg Med.* 2014;21(6):688-693. Erratum in *Acad Emerg Med.* 2014;21(7):841.

13. Downing SM. The assessment of clinical competence on the emergency medicine specialty certification examination: validity of clinically relevant multiple-choice items. *Ann Emerg Med.* 1980;9(11):554-556.

14. Downing SM, Maatsch JL. The effects of clinically relevant multiple-choice items. *Annu Conf Res Med.* 1979;18:124-128.

15. Kowalenko T, Heller BN, Strauss RW, et al. Initial validity analysis of the American Board of Emergency Medicine enhanced oral examination. *Acad Emerg Med.* 2017;24(1):125-129.

16. Bianchi L, Gallagher EJ, Korte RC, et al. Interexaminer agreement on the American Board of Emergency Medicine oral certification examination. *Ann Emerg Med.* 2003;41(6):859-864.

17. Maatsch JL. Linking theories of competence to the construct validity of performance assessment tests. In: Willis SL, Dubin SS, eds. *Maintaining Professional Competence: Approaches to Career Enhancement, Vitality, and Success Throughout a Work Life.* Jossey-Bass Publishers; 1990:95-124..

18. Maatsch JL, Munger BS, Podgorny G. On the reliability and validity of the board examination in emergency medicine. In: Wolcott BA, Rund DA, eds. *Emergency Medicine Annual: Nineteen Eighty-Two.* Appleton-Century-Crofts; 1982:183-222.

19. Cohen J. A power primer. *Psychol Bull.* 1992;112(1):155-159. doi:10.1037//0033-2909.112.1.155

20. Ross TM, Wolfe RE, Murano T, et al. Three- vs. four-year emergency medicine training programs. *J Emerg Med.* 2019;57(5):e161-e165. doi:10.1016/j.jemermed.2019.07.013

21. Three-year verses [sic] four year residency. Accessed July 29, 2022. https://www.saem.org/docs/default-source/saem-documents/students/3vs4yearresidency.pdf?sfvrsn=4

22. Kellogg A. 3 vs. 4 year residency – which is better? EM Advisor. Accessed July 29, 2022. https://emadvisor.blogspot.com/2015/07/3-vs-4-year-residency-which-is-better.html

23. ACEP Board of Directors. Workforce considerations: ACEP's commitment to you and emergency medicine. ACEPNow. Accessed July 29, 2022. https://www.acepnow.com/article/workforce-considerations-aceps-commitment-to-you-and-emergency-medicine/?singlepage=1

24. Marco CA, Courtney DM, Ling LJ, et al. The emergency medicine physician workforce: projections for 2030. *Ann Emerg Med.* 2021;78(6):726-737. doi:10.1016/j.annemergmed.2021.05.029

25. Baldwin DC Jr, Rowley BD, Daugherty SR, et al. Withdrawal and extended leave during residency training: results of a national survey. *Acad Med.* 1995;70(12):1117-1124.

26. McDeavitt JT, Appelbaum NP, Raddatz MM, et al. Taking leave during residency: types of absences and subsequent delays and variations in physical medicine and rehabilitation medical board pass rates. *Am J Phys Med Rehabil.* 2022;101(7 Suppl 1):S30-S34. doi:10.1097/PHM.0000000000002004

27. Laskey S, Cydulka RK. Applicant considerations associated with selection of an emergency medicine residency program. *Acad Emerg Med.* 2009;16(4):355-359.

AUTHOR BIOGRAPHY

Michael S. Beeson, MD, MBA, is the Program Director for the US Acute Care Solutions Emergency Medicine Residency at Summa Health in Akron, OH, where he also is Vice-Chair of Education in the Department of Emergency Medicine.