UC Irvine

Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health

Title

Understanding the Decline in Emergency Medicine Qualifying Examination Pass Rates

Permalink

https://escholarship.org/uc/item/81v3n0t9

Journal

Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health, 27(2.1)

ISSN

1936-900X

Authors

Weeker, Neelou
Srica, Nickolas
Coughlin, Ryan
et al.

Publication Date

2026-03-25

DOI

10.5811/westjem.64059

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

eScholarship.org Powered by the California Digital Library
University of California

Tukey's post hoc test. RCAS eSLOE scores with and without grades were compared using a T-test. This project was deemed exempt by the IRB.

Results: 917 SLOEs from 454 applicants were analyzed. There was a statistically significant difference in mean eSLOE score between scoring methods ($F(2,1359)=7.93$, $p<0.001$). Post-hoc analysis found that the mean eSLOE score was significantly different between our internal score and the standard RCAS score including grades ($p<0.001$, CI=2.06-9.37) and the RCAS score with grades excluded ($p=0.004$, CI=1.31-8.62). Grades were unavailable for 63 applicants (13.9%). The RCAS eSLOE score with grades excluded resulted in increased scores for 214 (54.7%) applicants by an average of 6.5 (SD=4.6, $p<0.001$) and in decreased scores for 177 applicants (45.3%) by an average of 5.8 (SD=4.4, $p<0.001$).

Conclusion: The RCAS eSLOE score is calculated differently if the applicant received a pass/fail clerkship grade. Given the absence of standardization in EM clerkship grading, we question the utility of including the clerkship grade when calculating eSLOE scores due to added statistical variability which limits comparisons between applicants.

Table 1. Analysis results.

Scoring Tools		Mean Difference (x1-x2)	Std. Error	Confidence Interval		p-value
(x1)	(x2)			Lower	Upper	
Internal Score	ResidencyCAS Standard Score (including clerkship grade)	5.71*	1.10	p<0.001	9.37	p<0.001
Internal Score	ResidencyCAS Score with Clerkship Grades Excluded	4.96*	1.10	p<0.001	8.62	p<0.001
ResidencyCAS Standard Score (including clerkship grade)	ResidencyCAS Score with Clerkship Grades Excluded	0.75	1.10	p=0.88	4.41	p=0.88

Applicant Categories	n. (%)	Standard ResidencyCAS Score (including clerkship grades)		ResidencyCAS Score with Grades excluded		Average of Differences	SD of Differences	t-test
		M	SD	M	SD			
Applicants with increased eSLOE scores	214 (54.7%)	64.3	23.6	70.8	21.0	6.5	4.6	20.6*
Applicants with decreased eSLOE Scores	180 (46%)	52.4	22.9	46.7	24.4	-5.8	4.4	-17.5*

*Statistically significant with p<0.001

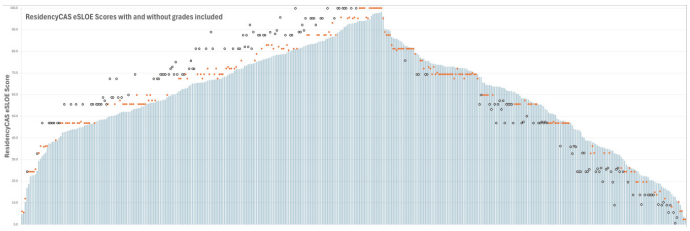


Figure 1. Applicant eSLOE scores with and without grades included. Blue bars represent standard eSLOE scores (weighted 33% each for clerkship grade, guidance, and expected rank list position). Orange circles represent the recalculated eSLOE score with grades excluded (50% weight for guidance and expected rank list position). White circles indicate statistically significant score increases or decreases when grades were excluded.

3 Do Year-To-Year Changes in In-Service Training Exam Performance Predict First-Attempt Success on Written Board Certification?

Brian Walsh, Fred Fiesseler

Background: In-service training examinations (ITE) are used annually to gauge resident progress, but the predictive value of changes in ITE performance for successfully passing the ABEM Qualifying Board Exam remains under-explored. We sought to determine whether year-to-year ITE percentile changes / improvements (deltaITE) forecast first-pass success on specialty written boards.

Methods: Retrospective cohort of all residency graduates (2015–2024). ITE percentiles were recorded in PGY-1, PGY-2, and PGY-3. deltaITE1-2 and deltaITE2-3 were computed as percentile point gains. Primary outcome was passing the ABEM Qualifying Board Exam on the first attempt as reported by the residents to the program director. Logistic regression modeled odds of first attempt pass success by deltaITE thresholds, adjusted for initial ITE percentile. ROC analysis evaluated deltaITE-based prediction. We further analyzed a subgroup of at-risk residents who had a low baseline ITE defined as less than the 30th percentile.

Results: 86 total residents were included in the analysis. Mean deltaITE1-2 was +19.4 (SD16.1); deltaITE2-3 was +12.8 (SD10.3). Each +10-point gain in deltaITE1-2 raised odds of first-pass success by 2.9-fold (OR 2.91, 95% CI: 1.8–4.7, $p<0.001$). DeltaITE2-3 of +15 points independently predicted success (OR 5.6, 95% CI: 2.3–13.8, $p<0.001$). Combined deltaITE model AUC = 0.91 (95% CI: 0.85–0.97). In the low-baseline subgroup (n=28), sustained deltaITE > +10/year yielded 93% first-pass rate versus 43% if deltaITE <= 0 ($p<0.001$). DeltaITE alone explained 41% of variance in first-pass success.

Conclusion: Year-to-year ITE improvement is a robust, independent predictor of first-attempt board success, outperforming static scores. Monitoring deltaITE enables early identification of at-risk residents, supporting targeted intervention to maximize first-pass rates.

4 Understanding the Decline in Emergency Medicine Qualifying Examination Pass Rates

Neelou Weeker, Nickolas Srica, Ryan Coughlin, Robert Sobehart, Kaila Pomeranz, Michael Menowsky, Tara Cassidy-Smith

Background: The American Board of Emergency Medicine (ABEM) Qualifying Examination (QE) is a key step in physician certification. In 2024, the first-time pass rate dropped to 82%, the lowest in recent history, raising concerns about resident

preparedness and program effectiveness.

Objectives: Identify program-level factors associated with QE performance.

Explore national trends in residency program approaches to exam preparedness.

Methods: This cross-sectional, anonymous survey study targeted U.S. EM residency program directors (PDs). A national working group of EM educators designed the survey, which was distributed via Qualtrics to 280 PDs. Descriptive statistics and chi-square goodness-of-fit tests were used to evaluate response distributions, with $p < 0.05$ considered significant. Qualitative data were analyzed using inductive coding and descriptive analysis.

Results: Out of 280 surveys sent, we received a total of 128 (45.7%) responses. Several variables were collected to assess predictors of ABEM QE failure. The ITE percentile was the strongest individual predictor of ABEM QE failure, with an H-statistic of 35.36 ($p < 0.000004$). Geographic region was associated with a higher incidence of failures ($H = 11.23$, $p = 0.01$). Programs offering fewer structured educational hours were also associated with higher QE failure rates ($H = 14.85$, $p = 0.021$). No statistically significant differences in QE failures were observed based on program type (academic vs. community), program length, or reported trends in ITE performance. The qualitative analysis of open-ended responses revealed three major themes: resident study habits, decreased EM competitiveness, and the rigor of undergraduate medical education.

Conclusions: The recent decline in ABEM QE pass rates may reflect broader systemic pressures within EM training rather than a transient testing anomaly. This study highlights the need for ongoing programmatic reflection and national dialogue.

5 Signals of Inclusion: Prevalence and Patterns of DEI Statements on EM Residency Websites

Abagayle Bierowski, Brittany Tian, Sulabh Neupane, Fernando Nombera-Bueno, Erin Hoag, Casey Morrone, Kathleen Cruz, Jiten Patel, Kelly Kehm, Peter Tomasell

Background: Diversity, equity, and inclusion (DEI) are increasingly recognized as essential components of resident education, workforce development, and institutional climate in graduate medical education. Because residency websites are often the first source of information about program culture, clear communication of program values is important; however, the extent to which EM programs include DEI content (and what that content entails) is unknown.

Objective: To evaluate the prevalence of DEI statements on EM residency program websites and examine whether program characteristics were associated with the presence of DEI content. A secondary aim was to characterize common themes within publicly posted DEI statements.

Methods: A cross-sectional analysis of all available ACGME accredited emergency medicine (EM) residency program websites (N=283) was conducted in 2025 to assess the presence and content of DEI statements. Chi square tests assessed whether DEI information was associated with program size, age, length, and region; thematic analysis identified recurrent content domains within DEI statements.

Results: Of the included 283 programs, most (n=213, 75.3%) did not include a DEI statement.

Four-year programs were more likely than 3-year programs to include a DEI statement (36.7% vs 22.2%; $\chi^2(1,N=283)=4.58$, $p=0.03$). Larger (≥11 annual positions) and older (est. 2006 or earlier) programs demonstrated higher inclusion of DEI language compared with smaller [31.9% vs 17.6%; $\chi^2(1,n=283)=7.78$, $p=0.005$] and newer [31.6% vs 18.4%; $\chi^2(1,N=283)=6.66$, $p=0.01$] programs. Regional differences were also observed, with programs in the Northeast (33.3%) and West (30.8%) more frequently including DEI statements than those in the South (13.5%) [$\chi^2(3,N=283)=10.54$, $p=0.01$]. Among those with DEI statements (n=70, 24.7%), thematic analysis identified recurrent domains (Table 1).

Conclusions: DEI content on EM residency websites is uncommon and varies significantly by program characteristics, highlighting gaps in transparency and opportunities for programs to better communicate their DEI priorities to applicants seeking programs aligned with their values.

Theme	n (%)
Commitment or mission-oriented language	36 (51.4%)
Health equity or social justice emphasis	32 (45.7%)
Education or training initiatives	26 (37.1%)
LGBTQ+ or gender-inclusive language	21 (30.0%)
Support for URiM or historically excluded groups	16 (22.9%)
References to underserved or marginalized communities	11 (15.7%)
Non-discrimination policy statements	5 (7.1%)
Description of formal DEI structures (e.g., committees, offices)	3 (4.3%)

**Some websites included more than one theme, so percentages do not total to 100.*

"Best of the Best" Innovation Abstracts

1 Realistic Dual-Setting Mass Casualty Incident Simulation to Enhance Triage and Definitive Care Skills

Scott Russo, Molly Basilio, Cosimo Laterza, Michael Berkenbush, Michael Brown

Background: Mass casualty incidents (MCI) require rapid triage, coordinated teamwork, and high-stakes decision-making that traditional instruction cannot replicate. Although