

## 16 A Virtual Reality Approach to Standardizing Mass-Casualty Training across Residency Programs

Cosimo Laterza, Stephen Powell, Michael Berkenbush, Nicholas Cozzi, Matthew Steenberg, Nicholas Kman

**Background:** Mass-casualty incident (MCI) training is limited by logistical and financial barriers. The ACGME requires education in disaster preparedness, MCI triage, simulation-based training, and systems-based practice. Virtual reality (VR) offers a scalable method that standardizes instruction across programs.

**Objectives:** To evaluate accuracy, efficiency, and consistency in a VR MCI using SALT triage across geographically distinct EM residencies. We hypothesized that VR would provide uniform assessment across sites and demonstrate reliable performance metrics with expected relationships between accuracy and error patterns.

**Methods:** We conducted a prospective observational multi-site pilot across four academic EM residency programs in fall 2024. Forty-six residents participated via convenience sampling (n=19,12,9,6), none were excluded. PGY1–3 residents completed a Stop-the-Bleed and SALT refresher prior to the session. Participants then triaged and managed 14 simulated patients in a standardized VR simulation. One trained proctor oversaw all sessions. Simulator log data were reviewed by two statisticians. Descriptive statistics were reported as means(SD). Pearson correlations assessed associations, and one-way ANOVA compared sites.

**Results:** Mean triage accuracy was 71.8%(11.7) with no site differences(p=0.18). Time to triage was 566s(159) and hemorrhage control time 426s(145), without site differences(p=0.76,p=0.97). Accuracy correlated negatively with total errors(r=-0.80,p<0.001), over-triage(r=-0.50,p<0.001), and under-triage(r=-0.60,p<0.001). Hemorrhage control time showed a trend-level correlation with total errors(r=0.29,p=0.078).

**Conclusions:** VR MCI training produces consistent performance metrics across EM programs in different states and aligns with ACGME disaster training requirements. Expected accuracy–error correlations support construct validity. Limitations include small sample size, one platform, and no follow-up. These findings support VR MCI simulation as a scalable method to meet disaster education requirements and a practical alternative to resource-intensive MCI drills.

## 17 A Comparison of Two Prediction Models for the American Board of Emergency Medicine Qualifying Exam

Kestrel Reopelle, Szymon Kutyla, Delaney Bates, Timothy Friedmann

**Background:** Recently, the American Board of

Emergency Medicine Qualifying Exam (ABEM QE) pass rate has declined, forcing post graduate educators to re-evaluate their preparatory curricula as well as the reliability of common preparatory resources and indicators. The ABEM-administered In-Training Exam (ITE) and commercially available question banks (Qbanks) are used for formative assessment by residencies and board eligible EM physicians to guide studying and to predict QE outcomes.

**Objectives:** This study aimed to determine whether a popular test prep Qbank’s QE pass prediction model differs significantly from the ABEM ITE’s prediction model. We hypothesized that there would be no difference between the two models for predicted probability of passing the ABEM QE.

**Methods:** This was a retrospective observational cohort study conducted at an urban academic residency. Qbank dashboard data for EM residents was collected from 2023-2025, correlated to ITE score reports from the same academic years, and de-identified for the study team. Other data collected included resident PGY and number of Qbank questions completed in the given academic year. Residents who did not take the ITE or who did not complete enough Qbank questions to have a reported probability of passing in a given year were excluded.

**Results:** 177 residents were included in this study and 15 were excluded. The average probability of passing the ABEM QE was 89.2% according to the commercial Qbank’s model and 82.4% according to ABEM’s model (a statistically significant difference, p<0.001)(figure 1). The significant difference between models persisted for residents with higher Qbank usage (figure 2).

**Conclusions:** The Qbank’s model has a more generous outcome prediction than ABEM’s, except at the lowest levels

	PGY1	PGY2	PGY3	PGY4	All Residents
ITE/ABEM reported mean likelihood of passing the ABEM QE:	86.0% chance	77.3% chance	82.4% chance	85.1% chance	82.4% chance
Commercial Qbank reported mean likelihood of passing the ABEM QE:	88.4% chance	90.3% chance	88.5% chance	89.5% chance	89.2% chance
Comparison of above mean values (two-tail paired t-test):	p=0.314	p<0.001	p=0.003	p=0.01	p<0.001

Figure 1: Comparison of mean probability of passing the ABEM QE by both models, sorted by PGY cohort and total study population

Comparison to mean # of Qbank questions completed:	>1 std dev below	Within 1 std dev below	Within 1 std dev above	>1 std dev above
# of Questions completed over the academic year:	8 - 135 questions	149 - 870 questions	910 - 1616 questions	1674 - 3935 questions
ITE/ABEM reported mean likelihood of passing the ABEM QE:	79.4% chance	80.7% chance	85.9% chance	85.6% chance
Commercial Qbank reported mean likelihood of passing the ABEM QE:	86.0% chance	88.11% chance	92.0% chance	90.9% chance
Comparison of above mean values (two tailed paired t-test):	p=0.07	p<0.001	p=0.002	p=0.009

Figure 2: Comparison of mean probability of passing the ABEM QE between both models, organized by number of questions completed in the Qbank