

Implementation of a Holistic Review Process Eliminates Non-Comparable Metrics and Bias in General Surgery Residency Interview Invitations:

Purpose

Use of traditional scoring metrics for residency recruitment creates racial and gender bias. In addition, widespread use of pass/fail grading has led to non-comparable data. To adjust to these challenges, we developed a holistic review (HR) rubric for scoring residency applicants for interview selection.

Objective

To examine the effect of a holistic review process in a general surgery residency program

Methods

Due to high application volume, initial screening filters, including academic failures, legal challenges, and a USMLE score threshold, were applied. HR process in 6 categories were then applied (volunteerism, research, leadership, recommendation letters, institution/state connection and background/adversity). Academic honor society membership and grades/scores were not included in HR. 34 reviewers were trained in implicit bias and the HR process and evaluated a median of 20 applications each. Interview offers were made based on overall HR score only. Statistical analysis using chi-squared test and multivariable logistic regression model evaluated HR performance in 2022 compared to various metrics in previous years (2015-2020).

Data from 2021 was excluded from analysis due to development and piloting of HR process.

Results

After initial screening, a total of 2507 applicants were narrowed down to 586 (23%) for HR. 52% were female and 17% identified as underrepresented in medicine (URM). Median STEP 1 and 2CK scores were 241 (range 220-271) and 253 (range 217-283), respectively. Median number of peer reviewed publications was 3 (range 0-31).

Based on overall HR score, 20% (118/586) of applicants were selected for interview. Median HR score was 11 (range 4-19). There was a fourfold higher coefficient of variation of HR scores (22.3; 95% CI 21.0–23.7) compared to USMLE scores (5.1; 95% CI 4.8-5.3), resulting in greater spread and distinction among applicants. There were no significant differences in HR score between gender ($p=.60$) or URM vs non-URM ($p=.08$). There were no significant differences in STEP1 ($p=.60$) and 2CK ($p=.30$) scores between those who were offered interview or not. On multivariable analysis, USMLE scores (OR 1.01; 95% CI 0.98-1.03), URM status (OR 1.71 95% CI 0.98-2.92) and gender (OR 0.94, 95% CI 0.60-1.45) did not predict interview selection (all $p>.05$). Compared to previous years (2015-2020) there was a significant increase in percentage of URM interviewed after HR implementation (12.9% vs 23.1%, $p=.016$).

Conclusion

The holistic review process eliminates the use of non-comparable metrics and bias for surgical applicant interview invitation and increases the percentage of underrepresented in medicine applicants invited to interview.