Title: Association between Race/Ethnicity/Gender and Evaluations of Resident Performance

Objective: To discover themes underlying the Underrepresented in Medicine (URM) resident experience and investigate any differences in resident performance evaluations based on URM status and gender.

Study Design: Grounded theory tenets and retrospective cross-sectional study

Materials and Methods: Qualitative data was gathered in the form of focus groups involving current URM residents in the Brigham and Women’s Hospital/Massachusetts General Hospital Integrated OB/GYN Residency Program. Quantitative data was gathered in the form of online summative performance evaluations for individual residents by faculty from academic years 2009-2010 through 2014-2015. Evaluations for individual competencies assessed were defined as above, below and at the mean. Qualitative concerns regarding resident performance were defined as yes or no. Chi-squared and Fisher exact tests were used to compare the proportion of URM and non-URM residents scoring below the mean for all competencies. A p-value of <0.05 was considered statistically significant.

Results:

Focus Groups: Seventy-five percent of URM residents (nine individuals) participated in one of two focus groups. Four major themes emerged from the narrative – discrimination, contrast in diversity in the residency vs. faculty, pressure to conform, and differing expectations. Overt acts of racism were rare.

Resident Evaluations: A total of 66 residents were evaluated during the study period; 56 were female (82%) and 19 were URM (34%). Of URM residents, 12 were black (63% of URM residents, 18% of total residents). A total of 197 summative evaluations were generated, comprised of two different types of forms – historical (2009-2012) and current (2012-present) – reflecting a change in the method of evaluation over time. When both types of evaluation forms were combined, a statistically significantly larger proportion of URM residents scored below the mean for the medical knowledge competency as compared to non-URM residents (56.6% vs. 40.9% respectively, p =0.041). There was no significant difference between the two groups for other competencies. A sub-analysis of results of the current evaluation forms revealed a significantly larger proportion of URM residents scoring below the mean for the following competencies: patient care (66.7% vs. 30.4%, p =0.003), medical knowledge (77.1% vs. 35.3%, p <0.001), practice based learning (71.4% vs. 40%, p =0.002), systems based practice (65.7% vs. 40%, p =0.01), communication (60% vs. 37.7%, p =0.025) and transitions of care (71.4% vs. 44.7%, p = 0.008). No differences were observed in the professionalism and teaching competencies. Overall, evaluating faculty expressed concerns about a larger proportion of URM vs. non-URM residents (48.6% vs. 10.6%, p <0.001). No differences were observed in evaluation data when stratified by gender.

Conclusions: Themes emerging from URM focus group data mirror those expressed in the literature. In the current evaluation system, URM residents appear to do more poorly on summative evaluations in nearly all competencies compared with their non-URM counterparts. Unconscious bias may play a role in these observed differences.

References: