The Journal of Education in Perioperative Medicine

Original Research

Exploring How Gender and Race Are Related to the Alignment of Anesthesiology Residents' Self-Assessments and Faculty Evaluation of Clinical Competencies

Fei Chen, PhD
Emily Teeter, MD, FASA
Adriana Ruby Gaona, BS
Robert Isaak, DO, FASA

Introduction

A resident physician who is capable of accurately self-assessing their competencies, which include both strengths and weaknesses, can potentially identify areas in need of improvement. This facet of improvement can lead to enhancing a resident's medical knowledge, professional development, and subsequent career development.[1] However, research, unfortunately, suggests that, in general, residents are unable to accurately assess their own performance as revealed by the discrepancies between resident self-assessment and faculty evaluations.[1-4] According to the Dunning-Kruger effect, individuals who perform poorly in certain tasks are also more likely to lack self-awareness of their errors and overestimate their own performance on self-assessment.[5,6]

In addition to the general association of level of competency and self-assessment, demographic and background characteristics of trainees, such as gender and race, may play a role in their level of confidence in their abilities. Studies of demographic disparity in graduate medical trainee self-assessments were mostly done in surgery.[7] The existing literature presents mixed results when examining gender-based differences in self-assessment. Some studies demonstrate significant disparities between male and female resident self-assessment, highlighting that female residents rate themselves lower than their male counterparts.[8-10] These studies also reported that female residents tend to have larger gaps between their self-assessments and faculty evaluations, suggesting a general lack of confidence compared with male residents. In contrast, other studies contradict these findings, indicating no significant gender-based differences in self-assessment.[11-13] Given the mixed findings, it is imperative to further examine the role of gender in influencing resident self-assessment.

Furthermore, there is limited research investigating race-based disparities in competency-based assessments. A study by Boatright et al[14] examined racial and ethnic differences in competency assessments of 9026 internal medicine residents, finding significant disparities in faculty evaluations across different racial and ethnic groups and suggesting bias in assessment. Specifically, however, their study did not investigate race-based discrepancies between self-assessment and Clinical Competency Committee (CCC) assigned scores. Whereas self-assessments do not directly influence resident promotion, they play an important role in guiding resident development. Conversely, assessing the accuracy of self-assessments is contingent upon objective and unbiased faculty evaluations. Thus, it is essential to examine factors, such as race and gender, that may influence both faculty assessments and resident self-assessments as well as the alignment between them.

This study aimed to examine the trend of CCC assigned scores, resident's self-assessment scores, and the alignment between the 2 scores as functions of gender and race using the Accreditation Council for Graduate Medical Education (ACGME) anesthesiology milestone assessment.[15] Specifically, we addressed 3 questions: (a) Are there any gender or race differences in terms of CCC assigned scores? (b) Are there any gender or race differences in terms of self-assessment scores? (c) Are there any gender or race difference in terms of the alignment between self-assessment and CCC assigned scores? To our knowledge, this study is the first to investigate both gender- and race-based differences among resident's self-assessments, and this contributes to the literature on gender- and race-based biases in assessment.

Materials and Methods

This retrospective observational study used the ACGME milestone 1.0 competency scores assigned by the CCC and clinical anesthesiology resident self-assessment from a single program over 6 academic years (December 2015 through June 2021). The CCC conducted semiannual assessments of residents, resulting in 6 assessments over the 3-year clinical anesthesiology training period. The CCC assessment followed a criteria-referenced approach, incorporating the milestone competency descriptor rubrics along with our institution-developed key performance indicator. A detailed description of the

continued on next page

continued from previous page

process has been published previously.[16]

Residents completed a total of 5 semiannual self-assessments beginning in December of the first clinical anesthesiology year and ending in December of the third clinical anesthesiology year, resulting in 5 self-assessment results. The self-assessment was completed independently by each resident. Approximately 1 month before the CCC meeting, residents received an email invitation to complete their self-assessment on each of the milestone competencies, using the same scale as the CCC. They were informed that the self-assessment was mandatory. The purpose of gathering the self-assessments was to provide comparison data to the CCC milestone assessments and formative feedback with the resident during the semiannual 1-on-1 meeting with a residency program director. This approach ensured that the self-assessment was integrated into the broader feedback process.

The outcomes included CCC assigned milestone scores reported to ACGME (CCC scores), self-assessment scores (SA scores), and the assessment alignment scores (alignment scores). The alignment score was measured by the difference between self-assessment and CCC scores (SA scores minus CCC scores).

Anesthesiology ACGME milestone assessment spans 6 core competencies: patient care (PC), medical knowledge (MK), systems-based practice (SBP), practice-based learning and improvement (PBLI), professionalism (PROF), and interpersonal and communication skills (ICS).[15] Assessment includes 25 subcompetencies scored on a scale from 0 (*has not yet achieved level 1/novice*) to 5 (*aspirational/expert*) with .5 increments. Results are reported as total score and core competence score for each of the 6 core competencies. The total score was obtained by adding all subcompetency scores, and the core competency score was obtained by adding the subcompetency scores within each core competency.

Gender and race data self-reported from the Electronic Residency Application Service. All residents identified as either female or male for sex, and this was treated as synonymous as gender in this study. The study included residents who self-identified as Asian, black or African American (hereafter, black), Hispanic, Latino, or of Spanish origin (hereafter, Hispanic), white and other race and ethnicity (hereafter, other). The other category includes American Indian or Alaska native, Middle Eastern, multiple races, other without specification, and those who preferred not to answer. Due to the large proportion of white residents and much smaller sample sizes in the other race categories, we combined all nonwhite race categories into a single nonwhite group for comparison.

Statistical Analysis

Descriptive statistics, including sample size and percentages, were summarized for the demographic characteristics of the sample. The mean and standard deviation of CCC scores, SA scores, and alignment scores for the total score and core milestone competencies were calculated. The significance of the magnitude of alignment score, measured by difference between SA and CCC scores, was assessed using paired *t* tests. To examine the relationship of interest regarding the roles of gender and race in 3 outcome sets (CCC scores, SA scores, and alignment scores), each set included 6 core competency scores and a total score as measures, a total of 27 mixed-effects analysis of variance (ANOVA) models were performed. Each model included resident gender (female or male), race (white or nonwhite), and their interaction term as predictors. Time was included as a repeated measure. The Tukey honest significant difference test was applied for post hoc multiple comparisons. Data were analyzed using SAS 9.4. A p value of less than .05 was considered statistically significant. Power analysis for ANOVA repeated-measures, between-factor analysis showed that a total sample size of 104 is required to achieve the desired power of .9 with a effect size of .25 and an alpha level of .05. The study was reviewed by the Office of Human Research Ethics of the University of North Carolina and was determined to be exempt from further review (IRB #23-0921).

Results

The final sample included 359 matched observations from 117 residents, totaling 17 773 subcompetency ratings (SA: n = 8800, CCC: n = 8973). On average, each resident completed 3 SA out of the possible 5 SA opportunities throughout the clinical training years. Gender and race distribution of the residents are summarized in Table 1. The mean and standard deviation of SA scores, CCC scores, and alignment scores for the total score and core milestone competencies are presented in Table 2.

The Role of Gender and Race in CCC Scores

The mixed-effects ANOVA model did not reveal any significant effects of gender, race, or their interaction on CCC scores. No significant effects were found for gender on any of the core competencies or total scores. No significant effects were found for race on any of the core competencies or total scores. The interaction between gender and race did not significantly impact CCC scores on any of the core competencies or total scores.

The Role of Gender and Race in SA Scores

For SA scores, significant interaction between gender and race was observed in PROF ($F_{(1,112)} = 5.15$, $p = .025$; see Figure 1). Specifically, female white residents tended to rate themselves lower than female nonwhite residents (mean [95% confidence interval or CI] = 14.26 [13.23, 15.30] versus 16.37 [15.15, 17.61], $p = .049$) and male white residents (mean [95% CI] = 14.26 [13.23, 15.30] versus 16.15 [15.41, 16.89], $p = .020$) on PROF. Other than PROF, we did not find any significant gender or race effect on SA scores for any other competencies or total scores.

The Role of Gender and Race in the Alignment Scores

With regard to the alignment scores, paired *t* tests found that residents tended to overrate their milestone competencies relative to the CCC placement (Table 2). In addition, the mixed-effects ANOVA models found improvement in alignment over time on total score and all competencies (p range: < .0001 to .002) except ICS ($F_{4,112} = 2.06$, $p = .091$). Residents consistently rated themselves higher than the CCC, and the score difference remained at approximately the same level on ICS at different time points.

Furthermore, a significant effect of race was observed in alignment scores of MK

continued on next page

continued from previous page

($F_{1,112}$ = 5.04, p = .027) such that nonwhite residents overrated themselves more than white residents relative to CCC scores on MK for most of the time ($Difference_{non\text{-}white - white}$[95% CI] = .22 [.03, .40]; see Figure 2). No other significant effects of gender or race were observed in the alignment scores of other competencies.

Discussion

In this study, we examined the roles of gender and race in assessing resident clinical competency, using 3 outcome measures, namely, CCC scores, SA scores, and alignment between the two. These outcomes were evaluated within the context of the anesthesiology ACGME milestone competency framework. Our major finding showed that residents generally overestimated their competencies compared with CCC scores with notable improvement in alignment scores over time except for interpersonal and communication skills. White residents' self-assessment was more aligned with CCC's evaluations of their medical knowledge than the nonwhite residents. Additionally, a significant gender-race interaction in the self-assessment of professionalism was revealed with female white residents rating themselves lower than their peers.

In addressing our first research question, we did not find any significant gender or race differences in the CCC scores. This finding is consistent with a recent study that found no gender-based differences in entrustment ratings of anesthesiology residents regardless of whether the assessments were completed by a male or female faculty member.[17]

As to our second research question, the results of our analysis indicate no gender- and race-based differences in self-assessment ratings of all competencies other than professionalism. A significant interaction between gender and race was observed in self-assessment of professionalism such that female white residents generally rated themselves lower than female nonwhite residents and male white residents. Given that no significant impact of gender and race was revealed in CCC scores, such significant interaction between gender and race in influencing self-assessment scores may suggest that female white residents feel less confident than nonwhite residents and male white residents who have undergone similar training. Previous study on ACGME milestones found that female orthopedic surgery residents give themselves lower scores on average than male colleagues and faculty evaluators, especially on patient care and medical knowledge milestones, but the role of race was not investigated.[9] Future studies may be warranted regarding the influence of race and other background characteristics of the residents in investigating trends and patterns of residents' competency self-assessment.

Our third question examined the alignment of resident self-assessment and CCC assessment. Our results suggest an overall tendency of residents being overconfident about their clinical competency, which is consistent with the findings from many previous resident competency assessment studies[3,4,9,12,18] although there was also evidence to suggest a different pattern.[2,10,19] Furthermore, the alignment between the self and CCC assessment improved over time. These findings are consistent with the Dunning-Kruger effect such that less competent individuals tend to overestimate their ability.[5] One explanation of such improvement in alignment may be a pure statistical regression to mean artifact rather than a consequence of lack of metacognition or self-awareness.[6,20] Yet there could also be a learned effect to account for the improved alignment over time. Specifically, junior residents may not understand the milestones as well as senior residents, and the more aligned assessment results may be due to better understanding of what and how they are being evaluated.

Gender and race did not appear to be significant predictors of alignment scores for most of the competencies. Unlike the disadvantages for female residents in terms of competency assessment reported in some previous studies,[8-10] our study found both female and male residents tend to overrate their milestone clinical competencies using CCC scores as the reference. It may be worth noting that such gender-related discrepancy was mostly reported in surgical education literature, whereas no significant gender-based difference findings were found in other specialties, such as family medicine, pathology, and anethesiology,[11-13] and this may indicate a specialty difference. Additionally, we found a significant effect of race-based difference for medical knowledge assessment alignment with white residents' ratings more aligned with CCC scores than the nonwhite residents. In our program, unlike the other milestones, CCC assessment of medical knowledge primarily relies on objective metrics, such as performance on in-training exams and the BASIC exams. Because we did not find any racial differences in the CCC scores for medical knowledge in this study, the racial disparities in the alignment scores may suggest differences in confidence in medical knowledge. Targeted educational interventions may be considered to address these specific competencies that show marked discrepancies. For example, junior residents or nonwhite residents may require more education to improve their understanding of the milestone assessment criteria.

Our findings could help restructure how feedback is given regarding milestone self-assessment or, at a higher level, underconfidence or overconfidence with residency performance. Likewise, our findings may help programs determine which residents are possibly experiencing internalized bias about their performance on competencies, enabling targeted interventions to address these concerns. Nonetheless, our findings must be interpreted with caution given the limitation in the sample from a single US anesthesiology residency program from a specific time frame. Although residents were informed that the self-assessment was mandatory, a notable proportion of assessments remained incomplete. In some cases, residents on leave did not complete the self-assessment. However, we were unable to track all the reasons for the missing data. Because there is no penalty for incompletion, we ultimately rely on residents to complete the self-assessment, which could introduce the potential for self-selection bias. Thus, although we did not identify any gender or race differences in CCC scores in our sample, it is still worth investigating possible biases exhibited by faculty evaluators from a more diverse population as next steps to further enhance

continued on next page

continued from previous page

resident's confidence to score themselves and promote equitable training environments. The significant effect of gender and race on self-assessment that we identified could be specific to our program. Because this is the first investigation on the role of race in competency assessment alignment, we are not able to review our results in the context of broader literature. Therefore, continued research is warranted regarding gender and racial biases in resident competency–based assessment, including faculty evaluation and resident self-assessment. Additionally, the improvement in alignment between CCC and self-assessment over time revealed in this study could be attributed to the ceiling effect of the milestone assessment scale. Furthermore, other factors such as nonmedical degrees, previous work experience, implicit bias, confidence, and metacognition ability could also be accounted for in future studies to better understand the relationship between personal characteristics and self-assessment ability.[12,20]

Our study investigated the relatively underexplored area of the roles gender and race play in competency-based assessments within anesthesiology graduate medical education. We found that residents generally overestimated their competencies relative to CCC, and alignment between self and CCC assessment improved over time except in interpersonal and communication skills. Notably, female white residents rated their professionalism lower than their peers, indicating a possible confidence disparity. Our findings suggest that further research is warranted to explore the impact of personal characteristics on competency assessment and to develop targeted interventions for improving assessment and reducing potential biases.

Acknowledgments

Part of the study results were presented as a poster presentation at the American Board of Medical Specialties Conference 2024 in Chicago, Illinois, in September 2024. The authors thank Seidah Congleton for assisting with the search of literature in preparing an early version of the abstract.

References

1. Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA*. 2006;296(9):1094-102.
2. Fleming M, Vautour D, McMullen M, et al. Examining the accuracy of residents' self-assessments and faculty assessment behaviours in anesthesiology. *Can Med Educ J.* 2021;12(4):17-26.
3. Yildirim M, Saglam AI, Ugurlu C, et al. Performance of residents during laparoscopic cholecystectomy: is self-assessment reliable? *Surg Laparosc Endosc Percutan Tech.* 2021;31(4):414-20.
4. Goldflam K, Bod J, Della-Giustina D, Tsyrulnik A. Emergency medicine residents consistently rate themselves higher than attending assessments on ACGME milestones. *West J Emerg Med.* 2015;16(6):931-5.
5. Mazor M, Fleming SM. The Dunning-Kruger effect revisited. *Nat Hum Behav.* 2021;5(6):677-8.
6. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol.* 1999;77(6):1121-34.
7. Helliwell LA, Hyland CJ, Gonte MR, et al. Bias in surgical residency evaluations: a scoping review. *J Surg Educ.* 2023;80(7):922-47.
8. Watson RS, Borgert AJ, O′Heron CT, et al. A multicenter prospective comparison of the accreditation council for graduate medical education milestones: clinical competency committee vs. resident self-assessment. *J Surg Educ.* 2017;74(6):e8-14.
9. Brady JM, Bray A, Kim P, et al. Female residents give themselves lower scores than male colleagues and faculty evaluators on ACGME milestones. *J Surg Educ.* 2021;78(4):1305-11.
10. Lyle B, Borgert AJ, Kallies KJ, Jarman BT. Do attending surgeons and residents see eye to eye? an evaluation of the accreditation council for graduate medical education milestones in general surgery residency. *J Surg Educ.* 2016;73(6):e54-8.
11. Athy S, Talmon G, Samson K, Martin K, Nelson K. Faculty versus resident self-assessment using pathology milestones: how aligned are we? *Acad Pathol.* 2021;8:23742895211060526.
12. Xu X, Ruan X, Yu C, et al. Differences between resident self-assessments and faculty- assessments on anesthesiology milestones and associated factors: a cross-sectional study. *BMC Med Educ.* 2024;24(1):551.
13. Bhardwaj N, Porterfield L, Kuo YF, Campbell KM. Resident self-assessment and clinical competency committee evaluations in family medicine: a single-institution study. *Fam Med.* 2023;55(5):333-8.
14. Boatright D, Anderson N, Kim JG, et al. Racial and ethnic differences in internal medicine residency assessments. *JAMA Netw Open.* 2022;5(12):e2247649.
15. The anesthesiology milestone project. *J Grad Med Educ.* 2014;6(1 Suppl 1):15-28.
16. Chen F, Arora H, Martinelli SM. Use of key performance indicators to improve milestone assessment in semi-annual clinical competency committee meetings. *J Educ Perioper Med.* 2017;19(4):E611.
17. Matava CT, Alam F, Kealey A, et al.. The influence of resident and faculty gender on assessments in anesthesia competency-based medical education. *Can J Anaesth.* 2023;70(6):978-87.
18. Lipsett PA, Harris I, Downing S. Resident self-other assessor agreement: influence of assessor, competency, and performance level. *Arch Surg.* 2011;146(8):901-6.
19. Ross FJ, Metro DG, Beaman ST, et al. A first look at the Accreditation Council for Graduate Medical Education anesthesiology milestones: implementation of self-evaluation in a large residency program. *J Clin Anesth.* 2016;32:17-24.
20. McIntosh RD, Moore AB, Liu Y, Della Sala S. Skill and self-knowledge: empirical refutation of the dual-burden account of the Dunning-Kruger effect. *R Soc Open Sci.* 2022;9(12):191727.

continued on next page

continued from previous page

The authors are at the University of North Carolina at Chapel Hill in Chapel Hill, NC. ***Fei Chen*** *is an Associate Professor in the Department of Anesthesiology;* ***Emily Teeter*** *and* ***Robert Isaak*** *are Professors in the Department of Anesthesiology;* ***Adriana Ruby Gaona*** *is a Research Assistant in the TEACHER Lab in the Department of Anesthesiology.*

Corresponding author: *Fei Chen, PhD, Department of Anesthesiology, University of North Carolina at Chapel Hill, N2198, CB7010, UNC Hospitals, Chapel Hill, NC 27599-7010. Telephone: (919) 966-5136*

Email address: *Fei Chen: fei_chen@med.unc.edu*

Financial disclosure: *The Excellence in Anesthesiology at Chapel Hill Education Research (TEACHER) lab is funded by the Department of Anesthesiology, the University of North Carolina at Chapel Hill.*

Abstract

Background: Accurate self-assessment is critical for self-directed learning and clinical competency development. Identifying factors that influence resident's competency-based assessments is imperative to address potential disparities and foster an equitable training environment. However, studies on the relationship between demographic characteristics, such as gender and race, and self-assessment are scarce. This study aims to examine the alignment between residents' self-assessment and faculty evaluation of clinical competencies and investigate the gender- or race-related discrepancies in assessment.

Methods: This retrospective study analyzed Accreditation Council for Graduate Medical Education milestone scores and self-assessments from clinical anesthesiology residents at a single site over 6 academic years (December 2015 through June 2021). Semiannual Clinical Competency Committee (CCC) assessments and resident self-assessments were compared to measure assessment alignment. Data were analyzed using mixed-effects analysis of variance and Tukey test.

Results: The sample included 17 773 subcompetency ratings from 117 residents with no significant gender or race effects on CCC scores. Self-assessment scores showed a significant gender-race interaction in professionalism milestones (p = .025) with female white residents rating themselves lower than their female nonwhite (mean [95% confidence interval or CI] =14.26 [13.23, 15.30] versus 16.37 [15.15, 17.61], p = .049) and male white peers (mean [95% CI] = 14.26 [13.23, 15.30] versus 16.15 [15.41, 16.89], p = .020). Residents generally overestimated their competencies compared to CCC scores (p range: < .0001 to .702) with notable improvement in assessment alignment over time (p range: < .0001 to .002) except for interpersonal and communication skills (p = .091). White residents' medical knowledge assessment alignment was better than the nonwhite residents (mean difference, 95% CI = .22 [.03, .40], p = .027).

Conclusions: Our study investigated the underexplored area of the roles gender and race play in residents' competency assessments. The findings suggest that further research is warranted to explore the impact of personal characteristics on competency assessment and to develop targeted interventions for improving competency assessment and reducing potential biases.

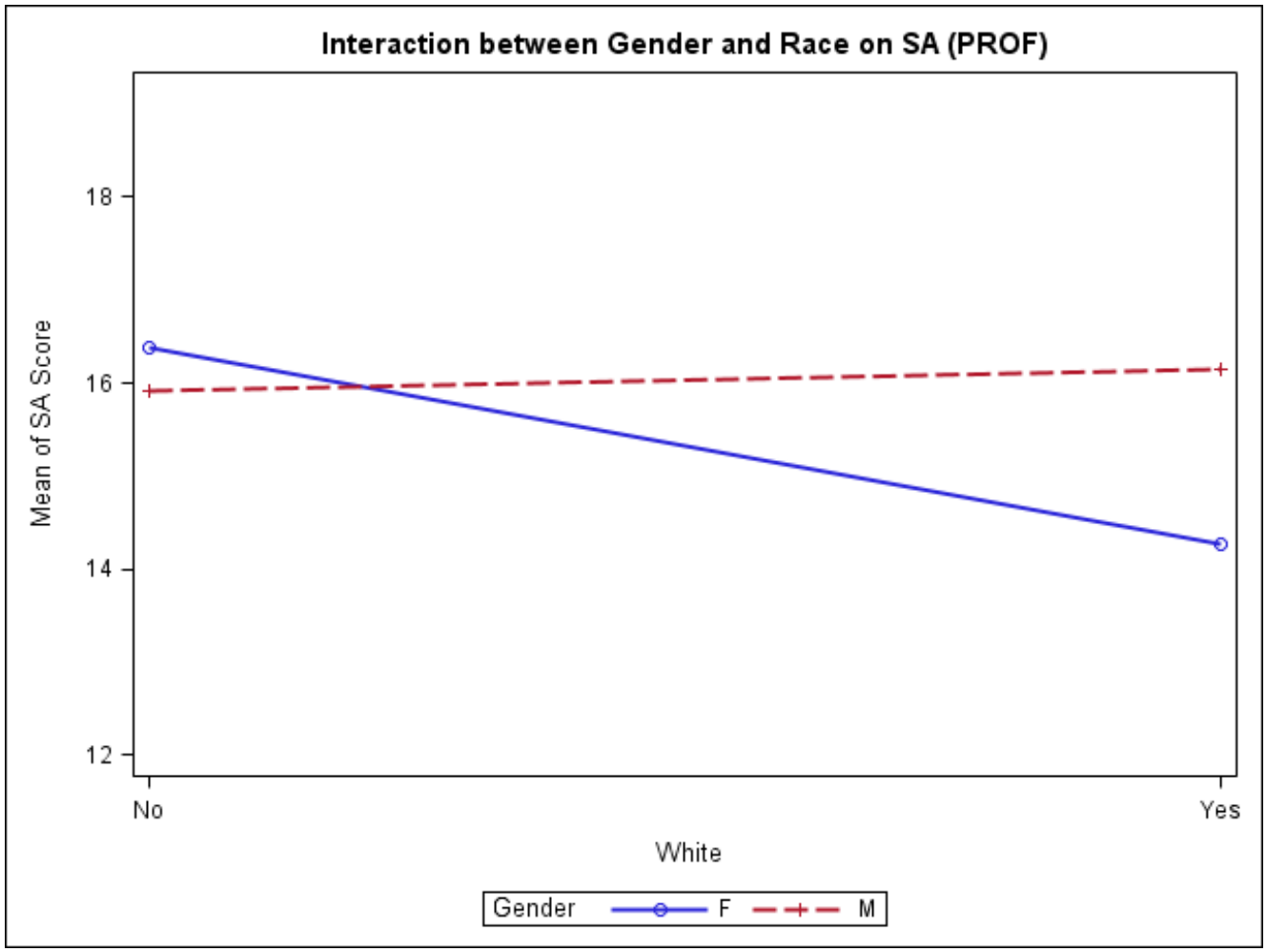
Keywords: Anesthesiology, competency-based assessment, graduate medical education, gender, race, faculty assessment, self-assessment

continued on next page

continued from previous page

Figures

Figure 1. *Interaction between gender and race on resident self-assessment (SA) on professionalism milestone competency (PROF). Legend: SA = self-assessment; PROF = professionalism milestone competency. Gender: F = female, M = male. White = race being white (yes/no). Mean of SA score: Calculated by adding the 5 professionalism subcompetency scores. Each subcompetency was rated on a scale from 0 (has not yet achieved level 1/novice) to 5 (aspirational/expert) with .5 increments.*

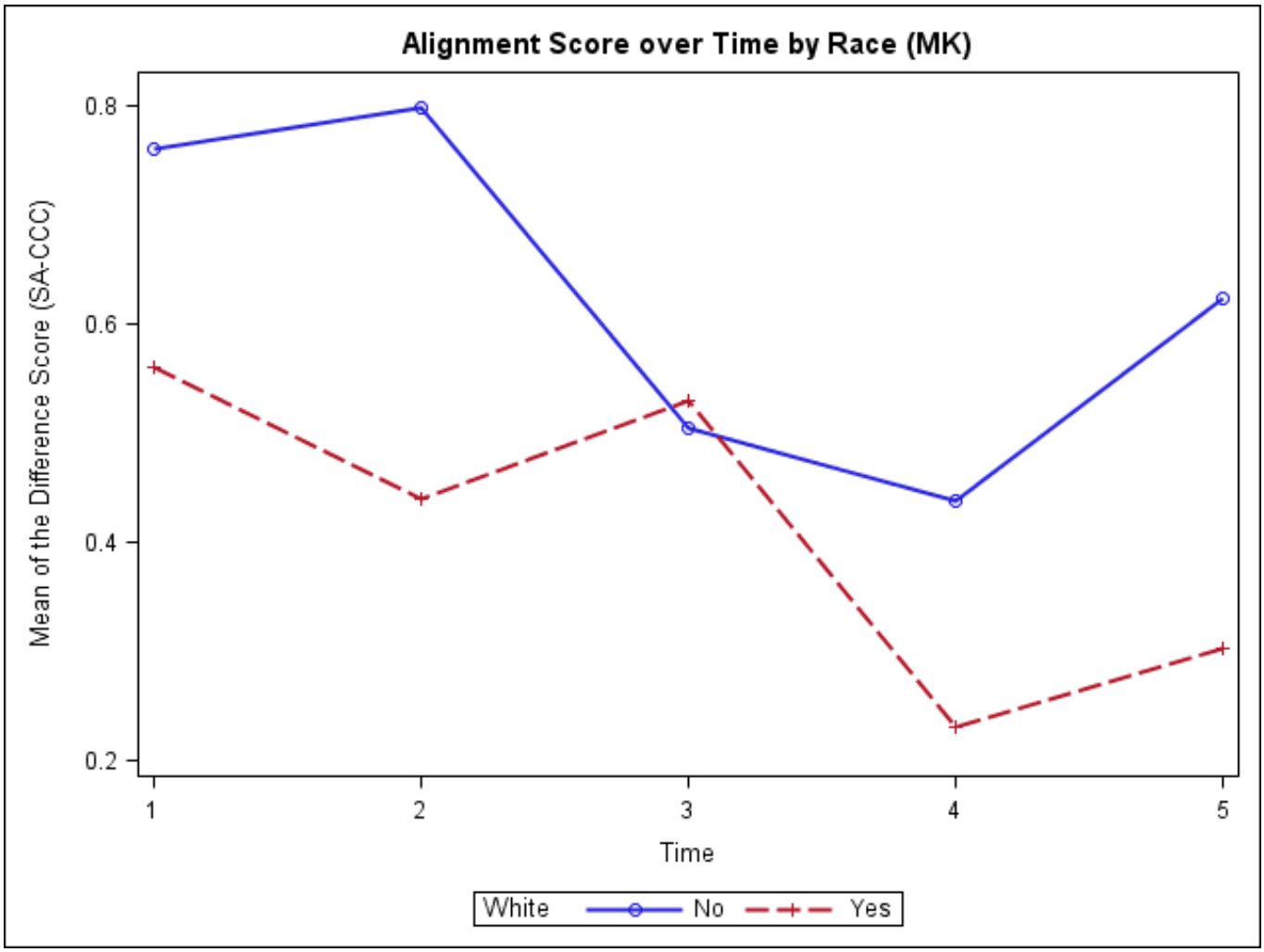


continued on next page

continued from previous page

Figures *continued*

Figure 2. *Assessment alignment score on medical knowledge milestone competency (MK) over time by race. Legend: MK = medical knowledge. White = race being white (yes/no). Mean of the difference score (SA - CCC): Calculated by subtracting CCC score from SA score for MK. MK consists of a single competency, which was rated on a scale from 0 (has not yet achieved level 1/novice) to 5 (aspirational/expert) with .5 increments. Time = 1: December (midyear) of CA1, Time = 2: June (year-end) of CA1, Time = 3: December of CA2, Time = 4: June of CA2, Time = 5: December of CA3.*



continued on next page

continued from previous page

Tables

Table 1. *Characteristics of Residents (N = 117)*

Characteristics	N	(%)
Gender		
Female	46	(39.3)
Male	71	(60.7)
Race		
Asian	18	(15.4)
Black or African American	7	(6.0)
Hispanic, Latino, or of Spanish Origin	7	(6.0)
Other[a]	7	(6.0)
White	78	(66.7)

[a] Includes American Indian or Alaska Native (n = 2), Middle Eastern (n = 2), more than one (n = 1), and other with no specification (n = 2).

continued on next page

continued from previous page

Tables *continued*

Table 2. *Comparison Between Self and Faculty Assessments on Anesthesiology Milestones by Competency and Time*

Competency	Time	SA		CCC		Alignment		*P* value[a]
		Mean	SD	Mean	SD	Mean	SD	
PC	1	22.028	5.94	14.464	1.135	7.725	5.933	< .0001
	2	27.043	5.362	19.799	1.626	7.243	5.429	< .0001
	3	29.813	5.030	25.39	1.507	4.396	5.351	<.0001
	4	33.174	4.876	29.604	1.071	3.558	4.97	< .0001
	5	37.029	4.397	34.93	1.584	2.086	4.612	.0003
MK	1	1.937	0.621	1.261	.357	.676	.61	< .0001
	2	2.621	0.699	1.951	.358	.664	.706	<.0001
	3	2.938	0.605	2.390	.336	.549	.582	< .0001
	4	3.268	0.645	2.951	.377	.319	.612	< .0001
	5	3.650	0.567	3.225	.412	.429	.677	< .0001
SBP	1	4.507	1.596	3.021	.132	1.486	1.606	< .0001
	2	5.371	1.348	3.993	.351	1.379	1.314	< .0001
	3	5.757	1.144	5.329	.435	.424	1.093	.0016
	4	6.630	1.059	6.535	.506	.101	1.083	.4393
	5	7.300	1.001	7.338	.412	-.043	.932	.7015
PBL	1	9.465	2.661	7.056	.532	2.408	2.649	< .0001
	2	11.05	2.557	9.181	.539	1.864	2.522	< .0001
	3	12.083	2.044	11.171	.641	.889	1.941	.0002
	4	13.572	2.024	13.313	.507	.268	1.973	.263
	5	14.829	1.841	15.063	.77	-.243	1.916	.2927
PROF	1	13.458	3.983	7.817	.581	5.641	3.995	< .0001
	2	14.921	3.278	10.056	.674	4.857	3.203	< .0001
	3	15.229	3.139	12.534	.733	2.667	3.087	< .0001
	4	17.261	2.787	14.938	.65	2.319	2.733	< .0001
	5	18.40	2.819	17.408	1.073	.993	2.794	.0041
ICS	1	7.556	2.444	3.042	.203	4.514	2.436	< .0001
	2	8.550	2.068	4.049	.267	4.500	2.036	< .0001
	3	9.104	2.037	5.116	.328	3.986	2.010	< .0001
	4	10.174	1.778	6.063	.365	4.101	1.748	< .0001
	5	10.843	1.684	7.176	.471	3.664	1.654	< .0001
Total	1	58.951	16.136	36.710	1.889	22.754	15.889	< .0001
	2	69.557	14.346	49.028	2.688	20.507	14.034	< .0001
	3	74.924	12.924	61.932	2.745	12.910	12.760	< .0001
	4	84.080	12.202	73.403	2.160	10.667	12.005	< .0001
	5	92.050	11.459	85.141	3.556	6.886	11.597	< .0001

Time = 1: December (midyear) of CA1, Time = 2: June (year-end) of CA1, Time = 3: December of CA2, Time = 4: June of CA2, Time = 5: December of CA3. SA = self-assessment, CCC = CCC assigned score, Alignment = score difference between SA and CCC scores (SA minus CCC). The total score was obtained by adding all subcompetency scores, and the core competency score was obtained by adding the subcompetency scores within each core competency.

[a] Based on paired *t* test comparing SA and CCC.