Implicit Bias in Step 1 Practice Question Banks

Paul Kent, MD[1], Asantewaa Ture, MD[2], Jeremy Chapman, MD[3], and James Gerhart, PhD[4]

[1]Rush University Medical Center
[2]Northwestern Memorial Hospital
[3]Children's Hospital of Wisconsin
[4]Central Michigan University

Introduction

As a scientific notion, race has been defined as subdivisions of the human species that differ from one another phenotypically, on the basis of ancestral geographic origins, or that differ in the frequency of certain genes[3]. Further, "the small amount of real genetic differences among [humans] (0.01%), and the difficulties of recognizing the racial identity of individuals through their genes" doubts the biological reality of race[3]. With that being said, why is race then included in the clinical setting when describing patients or used in an academic setting to teach medical students? Crudely stated, in medicine, race is used as a "convenient proxy for more important factors, like muscle mass, enzyme level, [and] genetic traits," that physicians do not have time to investigate in a clinical setting[4]. Race in medicine leaves patients of color vulnerable to harmful biases and stereotypes. A study conducted in 2003 confirmed that physicians were less likely to prescribe opioids to African American and Latino patients for isolated long bone fractures despite rating patients' pain as the same severity as White patients[5]. Several factors were listed for this disparity including physicians' perceptions of patients varying by race, communication difficulties based on socioeconomic status, and physicians having more negative perceptions of minority patients attributing to their beliefs that minority patients were seeking out opioids for abuse or sale[5]. The use of race in medical education may perpetuate or increase biases among medical students. Question banks are considered significantly important since "current learning models...theorize that experience is key to effective learning"[7]. In previous studies, it has been determined that students who completed a greater number of practice questions had higher Step 1 scores[7]. Braun et al analyzed the use of race/ethnicity in a question bank for Step 1 to determine "whether the use of race/ethnicity was merely descriptive or was central to any part of the question, and whether the question associated race/ethnicity with genetic difference"[8]. Our study expands on their findings by examining three commercially available question banks. As previously stated, in question banks race is used as "hint" for the correct answer. For example, in Braun's study, "sickle cell disease (SCD), sarcoidosis, and G6PD (glucose-6-phosphate dehydrogenase) together made up 50% of all cases mentioning Black/African American..."[8]. Use of race as a risk factor reinforces the "use of race as a simplistic signifier of illness, which treats it as an easily visualized diagnostic tool and obscures its complex role in illness"[6]. Since implicit negative stereotypes shape the way in which physicians evaluate and interact with minority patients, this type of "racial profiling" in medicine ultimately leads to health disparities. Race medicine is bad medicine, poor science, and a false interpretation of humanity.

Figure 1: Social Stigma Score Tool & Examples

EXAMPLE OF SCORING

#4174
"An infant is born at term to a 27-year-old **Caucasian** female. The prenatal course was **uncomplicated**.." → +2

#2476
"An **African-American** boy is...born to an 18-year-old woman who **did not receive prenatal care** or take **prenatal vitamins**..." → -2

Methodology

Data Source: The authors conducted a blinded study between 2019 – 2020 of the three most commonly used commercially available Step 1 question banks. A novel scoring system, Social Stigma Scoring Tool, was used to calculate the Social Stigma Score (SSS) for patient vignettes (Figure 1). SSS is the summation of positive and negative characteristics. The primary outcome was differences in mean SSS based on race. Secondary outcomes were differences in mean SSS by question bank and differences in frequency of negative and positive characteristics based on race.

Inclusion and exclusion criteria: All questions irrespective of whether race and/or ethnicity is included in the question stem were analyzed. Race/ethnicity was categorized based on U.S. Census definitions. For example, questions that mention European descent were included in the White/Caucasian category, African descent included in Black/African American category, and Asian descent assigned to the Asian category. For questions where only nationality was mentioned, assumptions were not made about the race of the patient. (Figure 2)

Figure 2: Inclusion and Exclusion
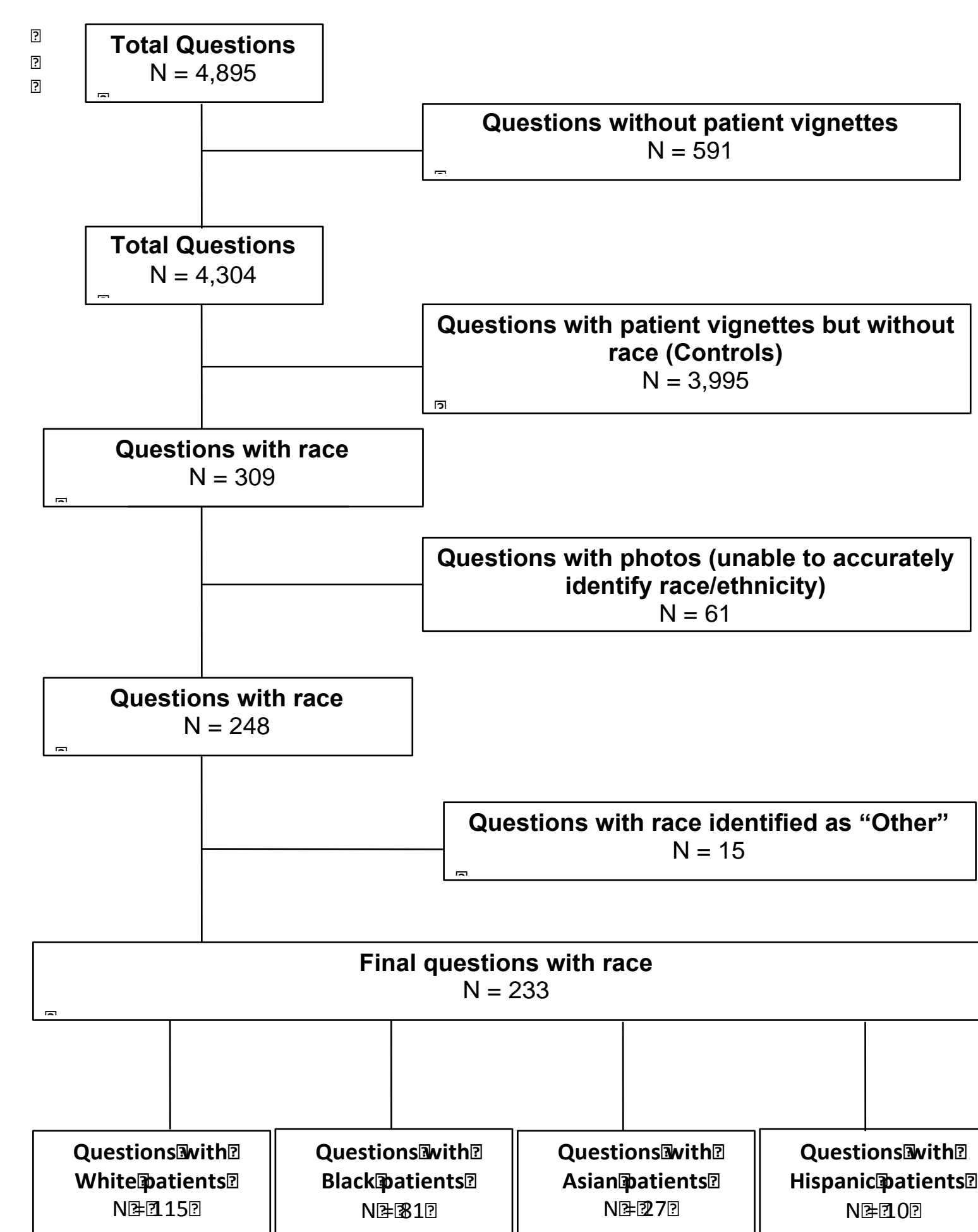
Analysis: For our analysis, we scored positive and negative characteristics associated with each patient. Each positive characteristic was given a score of +1, and each negative characteristic was given a score of -1. For example, stating that the patient eats a healthy diet is scored as +1, whereas eating a poor diet is scored as -1. (See Figure 1) For characteristics with definitive parameters, such as alcohol use, accepted guidelines were used. Each patient was then assigned a "Social Stigma Score" (SSS), equal to the sum of positive and negative points. Since the purpose of this study is to determine if the mention of race in question banks reinforces and/or introduces new biases/expectations, we chose to focus on socially stigmatized traits, such as personal responsibility, intelligence, and lifestyle choices. T-tests were used to compare the averages of SSS with p < 0.05 representing statistical significance. A one way analysis of variance (ANOVA) was used to determine if race had a significant impact on SSS.

Results

4,228 questions were analyzed: 3,995 did not include race, and 233 (5.5%) included unambiguous race: 49.4% White, 34.8% Black, 11.6% Asian, and 4.3% Hispanic. Vignettes with Black patients received the lowest mean SSS (-0.36, 95% CI -0.57 to -0.14). Vignettes with Asian patients received the highest mean SSS (0.26, 95% CI -0.27 to 0.79). By race and gender, Black men and Black women had the lowest mean SSS: (-0.44) and (-0.29), respectively. Hispanic men and Asian women had the highest mean SSS: (0.33) and (0.80), respectively. A one-way ANOVA showed that the effect of race on SSS was significant, $F_{(3, 229)} = 2.75$, $p = 0.044$.

Figure 3: Social Stigma Based on Race
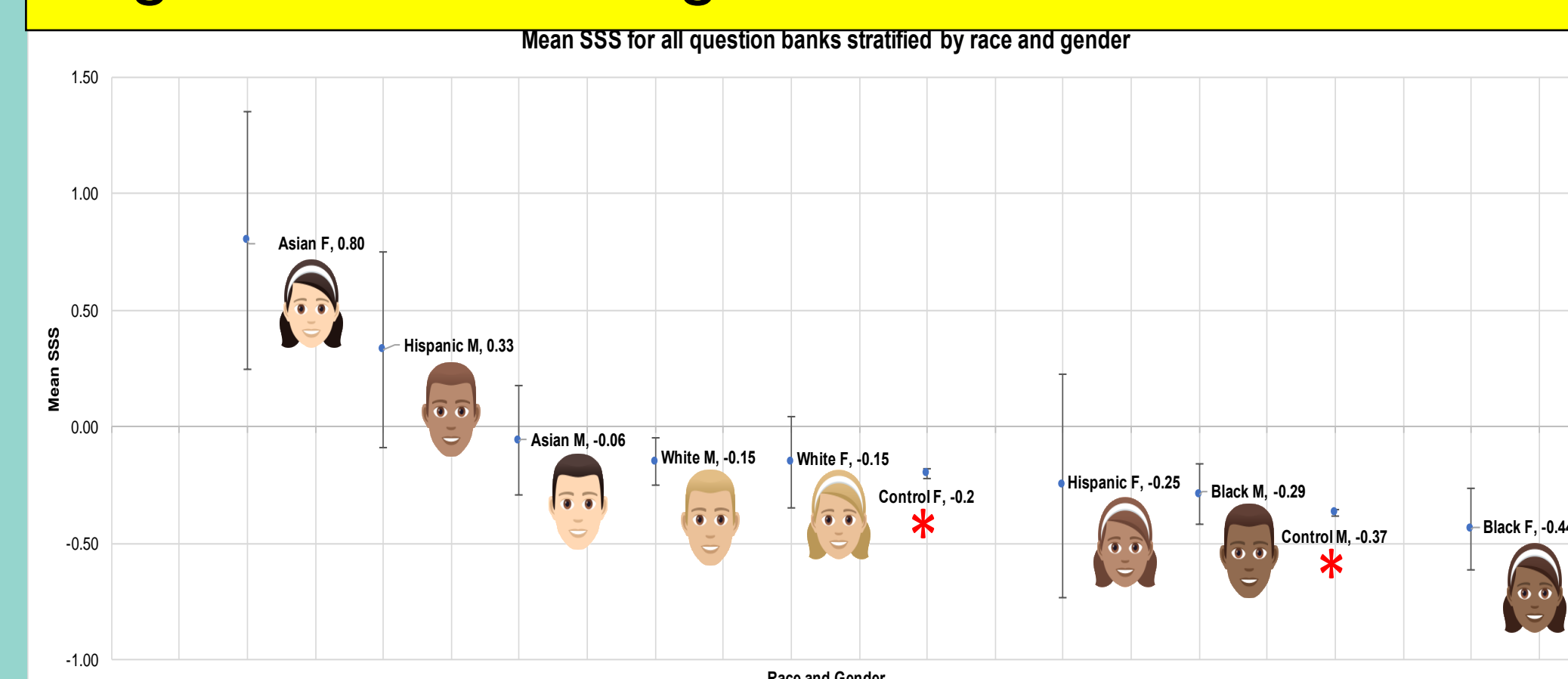
Positive characteristics included **uncomplicated pregnancies** for White and Hispanic patients and **medication adherence** for Black, Hispanic, and Asian patients. **Negative characteristic**s included **lack of physical activity** only mentioned in vignettes with White patients and **unprotected sex** only mentioned in vignettes with Asian patients. **Homelessness, altered mental status/forgetfulness, and IV drug** use were only mentioned in vignettes with Black patients. Figure 4. We also found increasingly negative portrayal of black man as they got older compared to controls. (Figure 5)

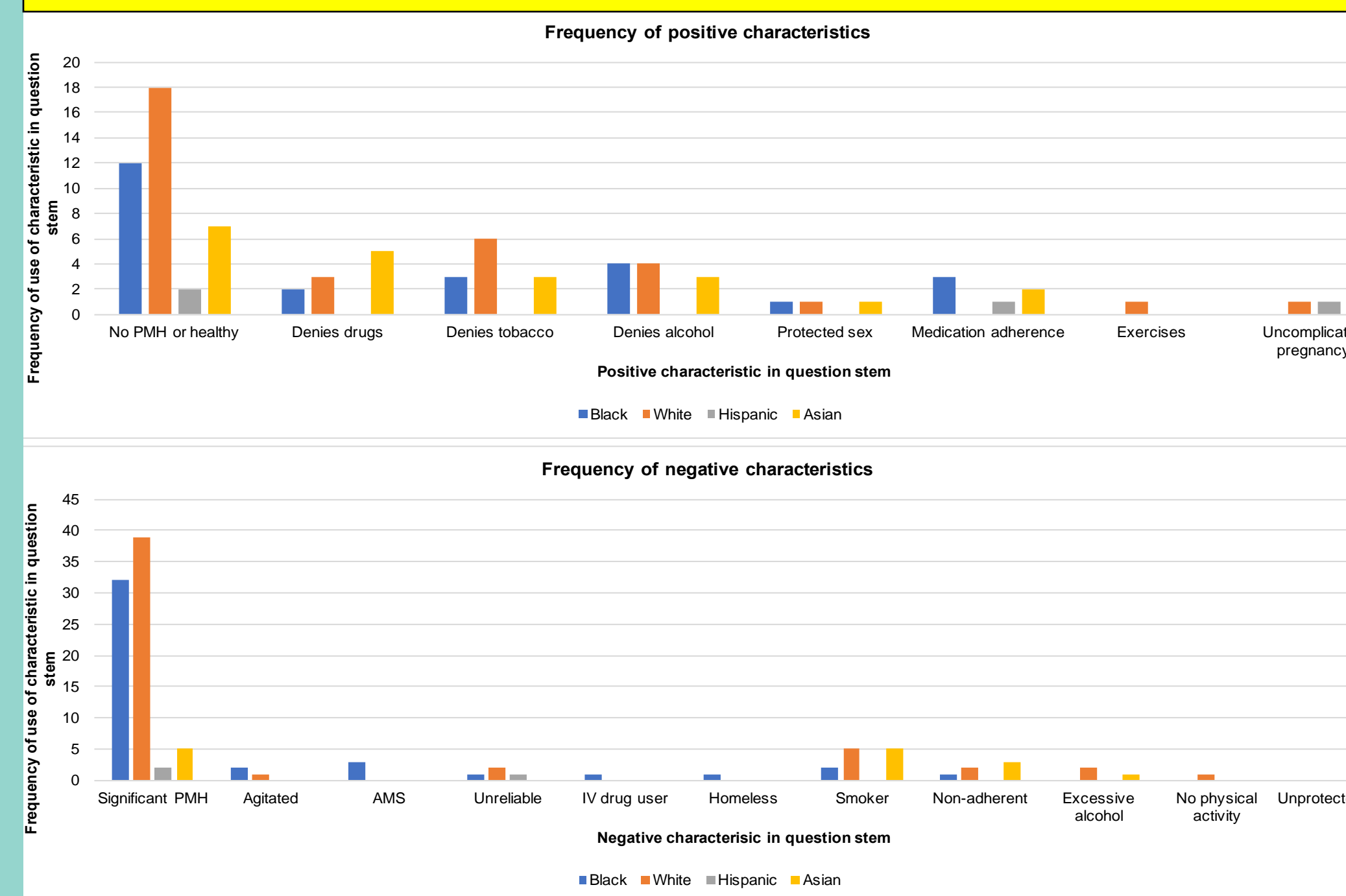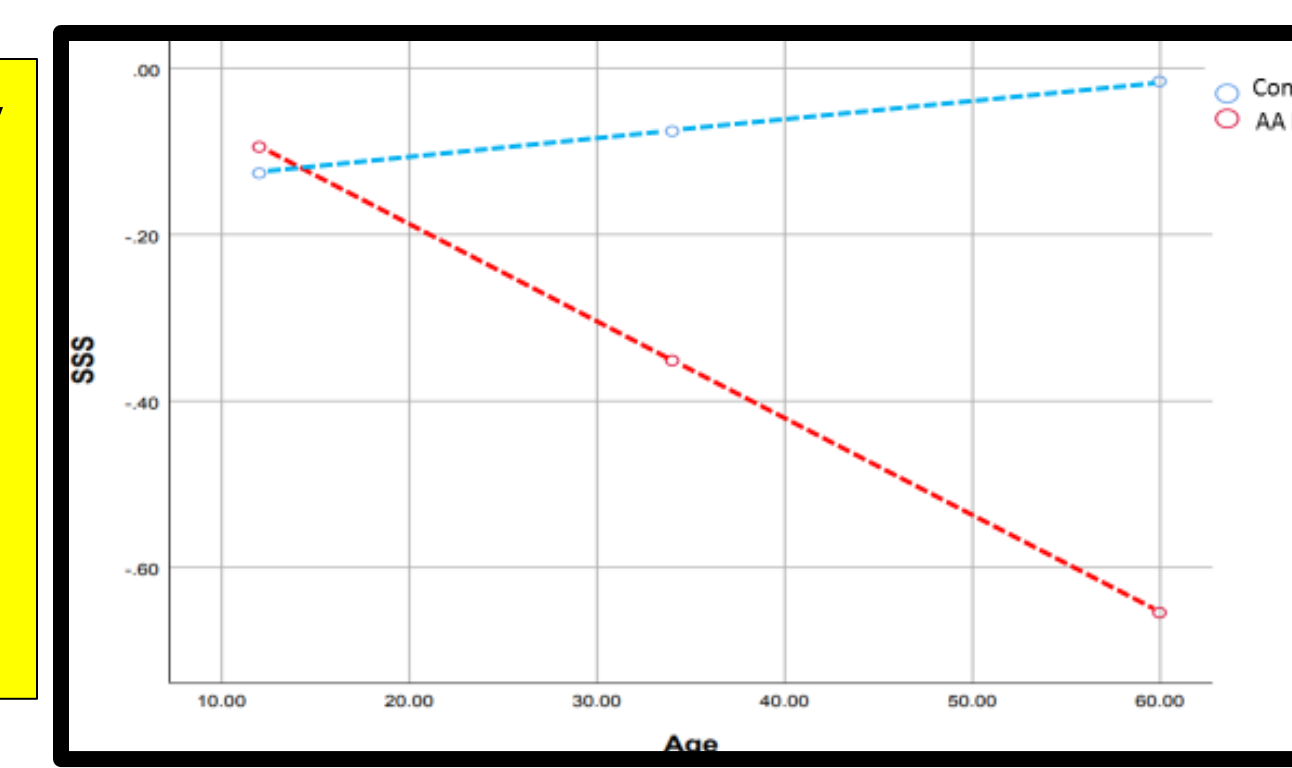### Figure 4: Frequency of Positive / Negative Characteristics Based on Race

Fig. 5: Increasingly Negative Stigmatization of Black Men with Age Compared to Controls

Conclusion

Last summer, on the heels of the George Floyd murder, and on the 155th anniversary of Juneteenth, 8 weeks from new medical students starting, Vyas et al in the NEJM warned us against the "...risk [of] baking inequity into the system", especially "when these tools influence high-stakes decisions."[15] Similarly, medical students spend their first 2 years using practice question banks as tools to prepare for the high-stakes USMLE Step 1 test. Our group studied over 4,000 question vignettes from the 3 most commonly-used Step 1 preparation question banks. Many question contain behavioral descriptions which are the focus of this study, as they have the potential to perpetuate/reinforce implicit biases and/or negative stigmata about certain demographic groups.

We found there is a difference in portrayal for patients based on race and gender. All races, except for Black (the most negative) and Asian (the most positive) were underrepresented compared to the 2018 U.S. Census.[13] While the largest number of questions that included race referred to White/Caucasian patients (49% v 77% in the US population), African-Americans represented 35% of questions, almost 3 times the 13% of U.S. census. Asians represented 12% of questions, but 6% of U.S. population, and Hispanic 4% of questions, but 18% of U.S. population[13]. When comparing SSS by race and gender, African-American men and women scored the lowest while Asian women and White women scored the highest. Thus the most negative portrayals (those of Black patients) was also tested, and reinforced, at almost 3 times the frequency of the national frequency.

These results replicate the findings of our 2017 study of UWorld Step 1 question bank and our 2018 study of a single question bank unblinded, which revealed that African-American patients were portrayed more negatively than White patients.

As question banks are designed to prepare students for licensing exams, these findings indicate a wider deployment of using race as a risk factor. Moreover, populations who experience the most negative cultural stereotypes also have the greatest healthcare inequalities[6]. For example, the incidence of severe maternal morbidity between 2012-2015 was significantly higher among deliveries to women in every racial/ethnic minority category compared to deliveries of White women. Further, severe maternal morbidity occurred in 231 and 139 per 10,000 delivery hospitalizations among African-American and White women, respectively (p = 0.001)[14]. While our study cannot provide direct evidence of what students learn from question banks, it is known that race shapes public beliefs. Therefore, at best, the mention of race reinforces existing biases, and at worst, introduces new biases to student-doctors. These biases have the potential to become part of larger systemic racism, which ultimately increases health disparities between white and minority patients.

NEXT STEPS:
Our research group is scoring blinded questions from Step 2, and 3 question banks and board review question banks to assess if the 'baked in' bias we found for Step 1 preparation tests is found there as well. We are also specifically targeting questions and teaching around sickle cell anemia, the most racialized disease in medicine.

Acknowledgements

References

1. Bennett, N. Hidden Curriculum in Continuing Medical Education. *Journal of Continuing Education in the Health Professions.* 2004 Summer;24(3):145-52.
2. Hafferty, F. The Role of the Hidden Curriculum in "On Doctoring" Courses. *AMA Journal of Ethics.* 2015 Feb 1;17(2):130-9.
3. Smedley A. Race as Biology Is Fiction, Racism as a Social Problem Is Real. *American Psychologist.* 2005 Jan;60(1):16-26.
4. Roberts, Dorothy. "The Problem with Race-Based Medicine." TEDMED, Nov 2015.
5. Tamayo-Sarver, J. Racial and Ethnic Disparities in Emergency Department Analgesic Prescription. *American Journal of Public Health.* 2003 Dec;93(12):2067-73.
6. Tsai, J. Race Matters? Examining and Rethinking Race Portrayal in Preclinical Medical Education. *Academic Medicine.* 2016 Jul;91(7):916-20.
7. Kumar A. Preparing to take the USMLE Step 1: a survey on medical students' self-reported study habits. *Postgraduate Medical Journal.* 2015 May;91(1075):257-61.
8. Ripp, K, Braun L. Race/Ethnicity in Medical Education: An Analysis of a Question Bank for Step 1 of the United States Medical Licensing Examination. *Teaching and Learning in Medicine.* 2017 Apr-Jun;29(2):115-122
9. Smedley BD. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. *National Academies Press;* 2003.
10. Stone, J. Non-conscious bias in medical decision making: what can be done to reduce it? *Medical Education.* 2011 Aug;45(8):768-76.
11. Haider, AH. Unconscious Race and Social Class Bias Among Acute Care Surgical Clinicians and Clinical Treatment Decisions. *JAMA Surg.* 2015;150(5):457-464.
12. Haider, AH, Scott VK. Racial disparities in surgical care and outcomes in the United States: a comprehensive review of patient, provider, and systemic factors. *J AM Coll Surg.* 2013; 216(3): 482-92.e12.
13. U.S. Census Bureau QuickFacts. 2013; Retrieved from https://www.census.gov/quickfacts/fact/table/US/IPE120217.
14. Admon, LK. Racial and Ethnic Disparities in the Incidence of Severe Maternal Morbidity in the United States, 2012-2015. *Obstet Gynecol.* 2018 Nov;132(5):1158-1166.
15. Darshali A. Vyas, et al, "Hidden in Plain Sight — Reconsidering the Use of Race." Aug 27, 2020. N Engl J Med 2020; 383:874-882