ORIGINAL RESEARCH

Racial/Gender Biases in Student Clinical Decision-Making: a Mixed-Method Study of Medical School Attributes Associated with Lower Incidence of Biases

Robert L. Williams, MD, MPH¹, Cirila Estela Vasquez, PhD², Christina M. Getrich, PhD³, Miria Kano, PhD⁴, Blake Boursaw, MS⁵, Crystal Krabbenhoft, BS¹, and Andrew L. Sussman, PhD, MCRP¹

¹Department of Family and Community Medicine; MSC09-5060, University of New Mexico, Albuquerque, NM, USA; ²Morehouse School of Medicine, Atlanta, GA, USA; ³Department of Anthropology, University of Maryland, College Park, MD, USA; ⁴Cancer Research and Treatment Center, University of New Mexico, Albuquerque, NM, USA; ⁵College of Nursing, University of New Mexico, Albuquerque, NM, USA.

BACKGROUND: Accumulating evidence suggests that clinician racial/gender decision-making biases in some instances contribute to health disparities. Previous work has produced evidence of such biases in medical students.

OBJECTIVE: To identify contextual attributes in medical schools associated on average with low levels of racial/gender clinical decision-making biases.

DESIGN: A mixed-method design using comparison case studies of 15 medical schools selected based on results of a previous survey of student decision-making bias: 7 schools whose students collectively had, and 8 schools whose students had not shown evidence of such biases.

PARTICIPANTS: Purposively sampled faculty, staff, underrepresented minority medical students, and clinical-level medical students at each school.

MAIN MEASURES: Quantitative descriptive data and qualitative interview and focus group data assessing 32 school attributes theorized in the literature to be associated with formation of decision-making and biases. We used a mixed-method analytic design with standard qualitative analysis and fuzzy set qualitative comparative analysis.

KEY RESULTS: Across the 15 schools, a total of 104 faculty, administrators and staff and 21 students participated in individual interviews, and 196 students participated in 29 focus groups. While no single attribute or group of attributes distinguished the two clusters of schools, analysis showed some contextual attributes were seen more commonly in schools whose students had not demonstrated biases: longitudinal reflective small group sessions; non-accusatory approach to training in diversity; longitudinal, integrated diversity curriculum; admissions priorities and action steps toward a diverse student body; and school service orientation to the community.

CONCLUSIONS: We identified several potentially modifiable elements of the training environment that are more

common in schools whose students do not show evidence of racial and gender biases.

KEY WORDS: students; medical; decision-making; healthcare disparities.

J Gen Intern Med 33(12):2056–64

DOI: 10.1007/s11606-018-4543-2

© Society of General Internal Medicine 2018

INTRODUCTION

Race, gender, and socioeconomic status remain key determinants of the quality of health care a person receives as well as the health outcome experienced.^{1–4} The causes of such disparities are multifactorial and include life circumstances; access to health care; and challenges of language, communication, and trust barriers.^{1,3} A growing body of research also demonstrates that stereotyping and bias on the part of practicing clinicians can contribute to disparities.^{5–21}

In earlier work, we explored the origins of these biases by testing for their presence in senior medical students using clinical vignettes.²² We examined the effects of varying patient race, gender, and socioeconomic status (SES) on student recommendations for further diagnostic or therapeutic steps. The results, while complex, showed that race, gender, and socioeconomic status had influence on student clinical decision-making. Importantly, we also found variations across the 84 participating schools in the extent to which each school's students, collectively, showed evidence suggestive of decisional biases. Students at some schools on average showed no apparent influence of patient race and gender on their recommendations, while students at other schools appeared to be so influenced.

The observed variation among schools in evidence of student racial and gender decision-making biases raises the question of whether and how medical schools can act to reduce or eliminate these biases in their students. Although the Liaison Committee on Medical Education (LCME) has required schools to include cultural competency training in their curricula as a possible step to reducing health disparities, the

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-018-4543-2>) contains supplementary material, which is available to authorized users.

Received November 3, 2017

Revised April 2, 2018

Accepted June 5, 2018

Published online July 11, 2018

evidence base linking the recommended content to student outcomes is unclear.²³ We are not aware of other systematic educational approaches—outside of that falling under the rubric of cultural competency training—that have been recommended to medical schools to assure their students make unbiased clinical decisions. Furthermore, there is no evidence of effects of specific cultural competency training approaches on improving health outcomes that can guide educators.²⁴

We undertook to study what elements of the medical school context or training might be associated on average with low levels of racial or gender biases in student clinical decision-making. By searching for differences in school environment and student training between schools whose students did or did not show evidence of decisional biases on the survey, our research goal was to identify actionable, feasible steps that schools could take to eliminate racial, gender, and SES biases in medical student clinical decision-making.

METHODS

Design and Overview. We conducted a mixed-method study with 15 case studies of medical schools.

Sample. In our previous study, we administered to 4603 senior medical students at 84 medical schools an Internet-based survey that included a set of cardiac care clinical vignettes.²² We asked students to select one of two clinically equivalent diagnostic or therapeutic recommendations. We varied the vignette patients' race, gender, and SES across the sample while holding other clinical aspects fixed, and observed the impacts on student recommendations of these changes in patient attributes. We then analyzed the percentage of students recommending procedural services within each subgroup of patients' attributes.

To identify specific schools for case studies in the current study, we calculated for each of the 84 medical schools the overall percentage of vignettes for which its students recommended procedural services, without regard to patient attributes. We then calculated percentages of procedural recommendations for each patient racial/gender subgroup (i.e., Black female, White male, etc.). Next, we calculated the mean absolute difference between the overall percentage vignette procedural recommendation for each school and the percentage of procedural recommendations for each patient racial/gender subgroup studied.

We selected schools for case studies from either end of a distribution of schools by mean difference in recommendations due to race/gender (range of mean absolute % difference in procedural recommendations: 0.8–14.1%). Sampled schools were those with mean differences approaching zero ("no evidence of bias cluster") and those whose mean differences were highest ("evidence of bias" cluster). Secondary analyses of the earlier survey data had shown some variation

by geographic region and by self-categorization as private or public, so we refined our school selection to represent variation in these characteristics.²²

Data Collection. Case study quantitative and qualitative data collection was guided by factors theorized or shown in the literature to be influential in development of unbiased student decision-making (Table 1). At each school, we gathered a set of quantitative measures (e.g., census data for surrounding community, medical student racial and gender composition). We also conducted qualitative data collection: key informant interviews of 6–10 purposively selected faculty (e.g., Dean of Undergraduate Medical Education, Cultural Competency course director), 1–3 purposively selected staff (e.g., staff director of admissions), and 1–3 purposively selected clinical-level medical students (e.g., underrepresented minority students identified by faculty as likely to be helpfully informative); and 2 focus groups, each of 6–10 clinical level medical students first responding to school-wide recruitment emails. Interviews and focus groups followed semi-structured guides (online Appendix 1) designed to capture aspects of the institutional and training context relevant both to the theorized factors and to the participants' areas of experience. Interviews and focus groups were transcribed with names removed. All participants received gift cards for their participation.

Data Analysis. We conducted two complementary data analytic processes with qualitative data: (1) standard qualitative analysis and (2) fuzzy set qualitative comparative analysis. Quantitative data were integrated into the fuzzy set analysis (see below) and provided one element contributing to interpreting the data in the standard qualitative analysis.

Standard Qualitative Data Analysis. The standard qualitative analysis followed an iterative grounded hermeneutic editing approach,²⁵ using as a foundation published concepts thought to be influential on student clinical decision-making biases. Four team members (one family physician, three medical anthropologists) participated in review of transcripts for the first three case studies to refine data collection processes. The anthropologists and a medical sociologist then independently reviewed the transcripts of one case study, each developing a coding structure. After meeting to resolve differences in codes, the process was repeated for two additional case studies to develop a final coding structure. Subsequent case studies were coded using this structure with periodic dual coding to confirm reliability of coding. The coding team met regularly as additional case studies were reviewed to refine thematic structures that were further modified in discussion with the full study team. After completion of 14 case studies, a prefinal thematic structure was reviewed at length. One additional case study was used to confirm/disconfirm the prefinal structure.

Table 1 Attributes of Schools Examined in Each Case Study

Attribute	Conceptual linkage to formation of unbiased clinical decision-making	Thematic grouping for fuzzy set qualitative comparative analysis
1. Student experiences prior to entering medical school	Experiential and peer influence: greater and more diverse experience broadens student perspectives	Student Body
2. Student orientation toward community service	Peer and values influence: peer attitudes toward needs of others promote acceptance	Student Body
3. Stigmatizing of underrepresented minority students	Peer influence: stigmatizing promotes concept of "other," devaluing diversity	Student Body
4. Class size	Peer influence: larger class size exposes students to greater spectrum of experiences and opinion	Student Body
5. Presence of diversity goals in formal school statement of mission	Values influence: value attributed to diversity promotes acceptance	School Priorities
6. Class racial, gender, and socioeconomic status diversity—percent of underrepresented minority students among graduates	Peer influence: exposure to diversity in peers broadens student perspectives	School Priorities
7. Class racial, gender, and socioeconomic status diversity—percent of ALANA (African, Latino, Asian, Native American) students among graduates	Peer influence: exposure to diversity in peers broadens student perspectives	School Priorities
8. Class racial, gender, and socioeconomic status diversity—percent of female students among graduates	Peer influence: exposure to diversity in peers broadens student perspectives	School Priorities
9. Presence of admissions priorities and/or preferences aimed at diversity in student body—stated goals and programs to increase diversity of student body, racial/ethnic/SES diversity preferences in admissions	Structural influence: school creates mechanisms to increase diversity, secondarily influencing student attitudes	School Priorities
10. Transparency to students and applicants of admissions preferences aimed at diversity in student body	Perceptions influence: student perceptions of fairness of policy influence attitudes toward diversity	School Priorities
11. External pressures on school to increase diversity in student body and faculty	Structural influence: school creates mechanisms to increase diversity, secondarily influencing student attitudes	School Priorities
12. School orientation toward community service	Values influence: value attributed to needs of others promotes acceptance	School Priorities
13. Student exposure to a diverse patient population	Experiential influence: exposure to diversity broadens student perspectives	Learning Environment
14. Faculty diversity—percent underrepresented minority	Experiential influence: exposure to diversity in role models broadens student perspectives	Learning Environment
15. Broad political context of state and region school is located in, with regard to historical and current value in diversity	Values influence: broad cultural value of acceptance of diversity influences student attitudes	Learning Environment
16. Community racial and ethnic diversity	Experiential influence: exposure to diversity broadens student perspectives	Learning Environment
17. Early introduction to diverse patients	Experiential influence: exposure to diversity broadens student perspectives	Formal Training
18. Use of racially and ethnically diverse standardized patients	Experiential influence: exposure to diversity broadens student perspectives	Formal Training
19. Student confidence in skills for communicating with diverse patients	Perceptions influence: student confidence in communicating with diverse patient increases willingness to accept diversity	Formal Training
20. Formal curriculum on diversity, cultural competence—longitudinal integration	Knowledge influence, peer influence, perceptions influence: increased awareness of problems in communications, peer responses, perceptions of personal contributions influence student acceptance	Formal Training
21. Formal curriculum on diversity, cultural competence—presence, structure, tone of messaging	Knowledge influence, peer influence, perceptions influence: increased awareness of problems in communications, peer responses, perceptions of personal contributions influence student acceptance	Formal Training
22. Formal curriculum on diversity, cultural competence—student perspectives	Knowledge influence, peer influence, perceptions influence: increased awareness of problems in communications, peer responses, perceptions of personal contributions influence student acceptance	Formal Training
23. Formal curriculum on diversity, cultural competence—students taught specific skills in intercultural communication	Perceptions influence: student confidence in communicating with diverse patient increases willingness to accept diversity	Formal Training
24. Formal curricular structures to maximize interaction among students	Peer influence: exposure to diversity in and feedback from peers broadens student perspectives	Formal Training
25. Role modeling of appropriate, non-biased interactions by faculty	Important others influence: value attributed to diversity and non-biased interactions by important others promotes acceptance	Informal Training
26. Degree of impact of hierarchical structures on communications	Important others influence: acceptance of diverse opinions by important others promotes student acceptance of diverse opinions	Informal Training
27. Role modeling of appropriate, non-biased interactions by community teaching clinicians	Important others influence: value attributed to diversity and non-biased interactions by important others promotes acceptance	Informal Training

(continued on next page)

Table 1. (continued)

Attribute	Conceptual linkage to formation of unbiased clinical decision-making	Thematic grouping for fuzzy set qualitative comparative analysis
28. Negative role modeling of inappropriate or biased interactions by community teaching clinicians	Important others influence, values influence: examples of biased interactions in role models clarifies student values for non-biased communications	Informal Training
29. Student perceptions of pressure to conform to school mission	Perceptions influence: perception of inappropriate pressure by important others may decrease student acceptance of opinions of others	Informal Training
30. Student perceptions of pressure to conform to school mission, inconsistent with student desires	Perceptions influence: perception of inappropriate pressure by important others may decrease student acceptance of opinions of others	Informal Training
31. Student perceptions of pressure by faculty in personal and career choices	Perceptions influence: perception of inappropriate pressure by important others may decrease student acceptance of opinions of others	Informal Training
32. Reflective, small group activities in curriculum	Peer influence: exposure to feedback from peers broadens student perspectives	Opportunity for Reflection

Fuzzy Set Qualitative Comparative Analysis. Fuzzy set qualitative comparative analysis (FSQCA) is an analytic approach based on set theory that can be used to identify multiple pathways that lead to a common outcome. It can be most useful in complex environments to identify circumstances or factors that are “necessary” or “sufficient”—either singularly or in combination—to produce a specified outcome. Because our early qualitative data analysis made clear that there was no single or group of school attributes that would invariably lead to absence of decisional bias, we used FSQCA to help identify multiple pathways to that outcome. Online Appendix 2 presents added detail about our approach to FSQCA, summarized as follows.

We began with our list of 32 attributes linked to theorized influences on development of biases in medical students (Table 1). Following usual FSQCA approaches, and based on both quantitative and qualitative data collected in the site visits, the analytic team met together to first define the observed characteristics of both a high (i.e., “fully in the set”) and a low (i.e., “fully out of the set”) level of each attribute theorized to be related to non-biased decision-making. The team then scored each school from 1 to 5 for each attribute, with a score of 5 representing a high level (“fully in the set”) of that characteristic.

The team meeting together also calculated a weighting factor for each attribute based on our judgment of the reliability of data representing that attribute, and the degree of relevance of that attribute to theoretical influences on formation of decision-making. Next, following the usual FSQCA protocol, we grouped the 32 attributes into six mutually exclusive categories based on shared themes: Student Body, School Priorities, Learning Environment, Formal Training, Informal Training, and Opportunity for Reflection (Table 1). Category scores for each school were calculated as the weighted mean score of all attributes in that category and represented the degree to which that school was “in” or “out” of that category of attributes theorized to produce non-biased decision-making.

Because we were primarily interested in conditions that would result in low decisional biases, we performed a “sufficiency first” analysis, identifying categories and attributes with high scores for being “sufficient” to result in being in the “no evidence of bias” cluster. As a second step, we then evaluated possible complex configurations considering scores for both “sufficient” and “necessary.”

Integration of Qualitative Analysis and FSQCA. The analytic processes were carried out in parallel, allowing ongoing data collection and development of analytic structures to be informed by both approaches. Iterative examination of the data helped refine both analyses. The final result of both analyses were brought together to create the conclusions and recommendations, through a series of team meetings focused on comparing and integrating the findings from both approaches.

Review and Approval Process. The study protocol was reviewed and approved by the University of New Mexico Institutional Review Board. We also obtained approval to conduct the case study from the Institutional Review Board at each participating school. All participants gave informed consent for participation.

RESULTS

Sample. Fifteen of 17 medical schools approached agreed to participate in the case studies, including seven schools whose students collectively appeared to show some evidence of race/gender bias on the survey (“evidence of bias” cluster), and eight schools whose students did not appear to show that bias (“no evidence of bias” cluster). The mean response rate to the earlier survey among these 15 schools was 52% (range 30–

78), higher than the mean for the full set of 84 schools that participated in the survey, which was 40%. These schools were balanced geographically and by self-described public/private status within each cluster. Within the 15 schools, we conducted a total of 29 focus groups with 196 clinical-level medical students and individually interviewed 104 faculty, administrators and staff, and 21 underrepresented minority students.

Qualitative Analysis. We found no single attribute or set of attributes that could universally distinguish between the two clusters of schools. Instead, we found several attributes or factors that occurred more commonly in the “no evidence of bias” cluster. The more of these factors present in a school, the greater the likelihood of that school being in the “no evidence of bias” cluster of schools. Figure 1 presents these factors, grouped into three overall categories: External factors, Institutional factors, and Training Environment factors.

External Factors. While all schools were obliged to meet the LCME standard for providing training in cultural competency, we found this standard was more likely to be highlighted and described as important at schools in the “no evidence of bias” cluster.

Institutional Factors. While most schools spoke of their interest in recruiting a diverse student body, we found that schools in the “no evidence of bias” cluster were more likely to emphasize this. In some cases, these schools created admissions processes and selection criteria designed to truly prioritize student diversity. Likewise, those in the “no evidence of bias” cluster were somewhat more likely to operate educational and recruitment pipeline programs in underrepresented communities.

Training Environment. Several characteristics in both the formal curriculum and informal training were more common in the schools in the “no evidence of bias” group. These schools were more likely to have longitudinal student small groups providing an opportunity for reflective discussion. In most cases, these groups had stable membership throughout the pre-clinical training period. Schools in this “no evidence of bias” cluster were also more likely to approach training on topics such as cultural competency and clinician bias as professionalism issues. This approach was non-accusatory and less confrontational, perhaps appropriately described as more positive, than the alternative, more deficit-based and judgmental approach seen more often in the “evidence of bias” cluster of schools (emphasizing each individual’s biases and need to correct a lack of understanding). The “no evidence of bias”

Fig. 1 Multiple pathways for encouraging equitable decision-making in medical students.

group of schools were also more likely to integrate and reinforce discussion of cultural competency topics throughout the 4 years of training rather than present this subject in the format of a dedicated, time-limited block of the curriculum.

In the informal aspect of training, we more commonly heard examples (particularly from students) of less hierarchical relations with residents and faculty and across professions at schools in the “no evidence of bias” cluster. We also more often heard in this cluster examples of rapid and transparent responses by school administrators to instances of racial/gender bias that occurred within the institutional setting. Finally, in these schools, we found a greater likelihood of wide support among students and faculty for diversity of the student body integrated throughout the curriculum. In contrast, we were more likely to hear subtle statements of resentment of perceived admissions or training advantages given to under-represented student groups in the “evidence of bias” group of schools.

Fuzzy Set Qualitative Comparative Analysis. Table 2 shows the scores for “sufficient” and for “necessary” for each attribute and each category. In all instances, higher scores represent greater association with the “no evidence of bias” cluster (a score of 0.8 is the conventional FSQCA threshold score for association). Opportunity for Reflection and Formal Training were the two individual categories with the highest sufficient scores. Table 3 presents the sufficient and necessary scores of selected combinations of categories with Opportunity for Reflection and Formal Training. Any combination that included high Opportunity for Reflection had sufficient scores greater than 0.9. Thus, the presence in a school of Opportunity for Reflection represented one path to inclusion in the “no evidence of bias” cluster. Without Opportunity for Reflection, when high Formal Training was combined with either high Learning Environment or high School Priorities, it increased sufficient scores (though not to the level reached when Opportunity for Reflection was present). These combinations represented a second path to inclusion in the “no evidence of bias” cluster.

Examining the specific attributes within the three categories in this second path (Formal Training with Learning Environment or School Priorities), Table 2 shows that Existence of Integrated Diversity Curriculum, Skills for Diverse Groups, Formal Structures for Student Interaction, Faculty Diversity, Broad Political Context, and External Pressures on School were highest scoring, and therefore most important.

We examined data from individual case study schools and found them to be consistent with the results shown in Tables 2 and 3. Of the eight schools in the “no evidence of bias” cluster, six had Opportunity for Reflection category scores of greater than 0.5 with the two remaining schools each scoring at 0.5 or more in at least one of the categories of School Priorities, Learning Environment, or Formal Training. On the other hand, of the seven schools in the “evidence of bias” cluster, six

Table 2 Association of Categories of Medical School Attributes and of Specific Attributes with Unbiased Student Clinical Decision-Making

Thematic category Attribute	Sufficiency	Necessity
Student Body	0.777	0.807
Prior experience of students with diverse communities	0.859	0.702
Student orientation toward community service	0.758	0.685
Absence of stigma toward URM students	0.654	0.820
Class size	0.839	0.636
School Priorities	0.805	0.647
Diversity goals in school mission	0.792	0.612
Diversity in graduates: URM	0.774	0.339
Diversity in graduates: ALANA	0.769	0.605
Diversity in graduates: gender	0.638	0.875
Admissions priorities promoting diversity	0.750	0.667
Opacity/transparency admissions diversity preferences	0.765	0.402
External pressures on school to foster diversity	0.804	0.493
School community service orientation	0.696	0.751
Learning Environment	0.784	0.630
Exposure to diverse patients	0.687	0.702
Faculty diversity	0.898	0.445
Broad state/regional political context regarding diversity	0.831	0.764
Diverse community composition	0.672	0.745
Formal Training	0.848	0.775
Early introduction to diverse patients	0.730	0.745
Use of diverse SP for interactional skills	0.769	0.549
Student confidence in communicating with diverse patients	0.619	0.848
Existence of longitudinal, integrated diversity curriculum	0.891	0.553
Diversity curriculum: tone of presentation	0.728	0.860
Diversity curriculum: student perspective	0.712	0.851
Concrete skills training for diverse groups	0.872	0.598
Formal structures for cross-group student interactions	0.869	0.722
Informal Training	0.693	0.896
Positive role modeling by faculty	0.737	0.817
Hierarchy/approachability of faculty and residents	0.702	0.819
Positive role modeling by community clinicians	0.705	0.719
Negative role modeling by community clinicians	0.631	0.571
Pressure to conform to school service mission	0.601	0.946
Pressure conform school mission, contrary to student desires	0.601	0.955
Pressure to conform to faculty values	0.650	0.901
Opportunity for Reflection	0.874	0.663
Longitudinal small group reflective curricular component	0.874	0.663

Sufficiency = likelihood that attribute is sufficient to be associated with schools whose students show unbiased decision-making. Necessity = likelihood that attribute is necessary to be associated with schools whose students show unbiased decision-making. Higher values indicate stronger association; values of 0.80 or higher considered threshold to show association

URM underrepresented minority; ALANA African, Latino(a), Asian, and Native American; SP standardized patient

scored at or below 0.5 in Opportunity for Reflection with the one exception scoring below 0.5 on both the Faculty/Staff Diversity and Existence of Integrated Diversity Curriculum attributes.

Table 3 Association of Combinations of Categories of Medical School Attributes with Unbiased Student Clinical Decision-Making

Categories	Sufficiency	Necessity
High Opportunity for Reflection	0.874	0.663
With high Informal Training	0.935	0.383
With high Student Body	0.932	0.365
With high Learning Environment	0.927	0.340
With high School Priorities	0.926	0.334
With high Formal Training	0.925	0.618
High Formal Training	0.848	0.775
With high School Priorities	0.884	0.396
With high Learning Environment	0.855	0.431
With high Informal Training	0.807	0.456
With high Student Body	0.806	0.424

Sufficiency = likelihood that category or combination of categories is sufficient to be associated with schools whose students show unbiased decision-making. Necessity = likelihood that category or combination of categories is necessary to be associated with schools whose students show unbiased decision-making. Higher values indicate stronger association; values of 0.80 or higher considered threshold to show association

To summarize, we found in FSQCA that Opportunity for Reflection was sufficient—but not necessary—to be associated with one path to the “no evidence of bias” cluster. Combinations involving selected attributes in the Formal Training, Learning Environment, and School Priorities categories showed high sufficiency, representing a second path to inclusion in this “no evidence of bias” cluster.

Online Appendix 3 provides a few quotes illustrating data from schools in both clusters.

DISCUSSION

Through a set of case studies, we found aspects of the training environment that were more often present in medical schools whose students collectively did not show evidence of racial and gender biases in clinical decision-making in a prior survey. By focusing on potentially modifiable elements of the training environment, we have identified those aspects that could be feasibly adopted by schools wishing to reduce the possibility of these biases in their students. For example, our standard qualitative analysis found a greater likelihood of the presence of pipeline programs for underrepresented groups and of longitudinal small group reflective sessions in schools whose students did not demonstrate biases (figure). Our fuzzy set analysis found these “no evidence of bias” schools more commonly provided concrete skills for working with diverse groups and had longitudinal, integrated diversity curricula (Table 2).

At first glance, it might appear that our two approaches to data analysis (standard qualitative and fuzzy set qualitative comparative) resulted in different answers to the question of what steps medical educators might undertake to reduce decision-making biases. On closer inspection, however, it is apparent that these different approaches, while providing two

different windows on the problem, in fact have a high level of concordance.

Both the qualitative analysis and the FSQCA suggest that longitudinal small group reflective opportunities, found in many, but not all of our “no evidence of bias” cluster, may by itself be sufficient to reduce the likelihood of racial and gender bias in student clinical decision-making. The FSQCA also notes that having longitudinal small group reflective sessions is not necessary to be associated with the “no evidence of bias” group of schools. Alternatively, as suggested by overlapping findings in the qualitative analysis and the FSQCA, having several other attributes can be another pathway to being in the “no evidence of bias” cluster. These attributes include those grouped in Formal Training, School Priorities, and Learning Environment in our FSQCA. In the qualitative analysis, the overlapping and potentially actionable characteristics were identified as the following:

- A professionalism-focused, non-accusatory approach to training in diversity
- A longitudinal, integrated diversity curriculum
- Admissions priorities and action steps toward ensuring a diverse student body
- School service orientation to the community

Our team initially anticipated that the content of instruction in subjects such as cultural competency would be the most important variable differentiating the two clusters of schools. Remarkably, however, it appears that processes and environment of instruction were much more important. Creating training environments that (1) reflect diversity, (2) emphasize the importance of equitable care by longitudinal and non-judgmental exposure to concepts of providing this care, (3) model this importance through commitment to community service, (4) demonstrate responsiveness to incidents suggestive of bias when they occur, and (5) provide small group reflective opportunities for discussion all appear to be more important than the specific content of instruction. Indeed, these elements of the training environment are consistent with social-cognitive psychology concepts of steps necessary to reduce biases.²⁶

There is as yet little other evidence in the literature to guide educators seeking to reduce student decision-making biases. While some medical schools have incorporated the Implicit Association Test²⁷ as a means to demonstrate to students the existence of biases, evidence of resulting long-term change in student behavior or decision-making or of patient outcomes is lacking. In a recent study investigating elements of medical school training environments as contributors to students' readiness to provide equitable care, an association was found between White student perceptions of their medical school and individual learning orientation (i.e., personal emphasis on learning goals viewing challenges as opportunities to improve skills vs. performance goals with a focus on performance evaluation) and their perceived skills and self-efficacy to care for minority patients.²⁸ The greater the degree of

perceived learning orientation (as contrasted with a performance orientation emphasizing evaluation), the greater was the sense of self-efficacy. These findings can be seen as comparable to ours that a non-judgmental, professionalism approach to intergroup communication training is associated with less evidence of student biases. Interestingly, in further multivariate analysis, learning orientation was not associated with changes in a measure of implicit bias over time.²⁹

Limitations

In considering the findings of this study, several limitations should be kept in mind. First, the selection of schools for case studies was based on a single year, cross-sectional survey of students. It is possible that an ascertainment bias could have resulted from a school's being included in one of the clusters due to random variation in the mean student survey responses. However, the use of multiple sources of data from a school and multiple schools in the analysis would appear to lessen the chance that such a bias would impact the overall findings.

Similarly, the use of multiple sources of data from multiple perspectives at each school makes a Hawthorne effect (falsely enhancing the description of curricular elements) unlikely. Second, it is possible that a systematic error in sampling for interviews and focus groups could result in misrepresentative data for a school. The multiplicity of data sources for each school would seem to diminish, but not exclude this possibility. Third, it is possible because of the social stigma of racial and gender biases that our data collection methods were not sufficiently sensitive to uncover important distinguishing variables, or that our participants may have withheld some information that would be important in ways that were not apparent to our team. Despite these potential limitations, however, the multiplicity and triangulation of data sources and the parallel analyses provide robustness and strength to the findings. Finally, we cannot suggest that our findings represent best practices in efforts to reduce clinical decision-making biases, only that they represent common findings among the schools we studied.

Implications for Policy and Practice

Several of the attributes of training environments we found to be associated with lack of biases in student decision-making could be implemented without the need for policy changes. For example, many schools now incorporate learning communities, a type of longitudinal reflective small group such as we found to be associated with schools with unbiased student decision-making. Similarly, rapid and transparent leadership responses to instances of bias would not require policy change. On the other hand, integration of diversity topics longitudinally throughout the curriculum could be accomplished without policy change, but greater specificity about LCME requirements on diversity training to emphasize the importance of longitudinal training would likely enhance this change. Creating admissions processes that successfully diversify the student body may be the most challenging policy change for

schools operating under legal and regulatory restrictions. However, we found some schools are able to diversify by, for instance, using family of origin economic indicators as selection factors.

Conclusions

Medical schools have a unique responsibility in addressing health disparities. In creating physicians to care for all members of our society, they must aim to graduate students as free of clinical decision-making biases as possible. Our study suggests that there are potentially modifiable elements of the training environment that are more common in schools whose students do not show evidence of racial and gender biases.

Acknowledgements: The authors wish to thank Lee Green, MD, MPH, for consultation with the fuzzy set qualitative comparative analysis, and Denise Ruybal for untiring and unflappable administrative support.

Corresponding Author: Robert L. Williams, MD, MPH; Department of Family and Community Medicine; MSC09-5060 University of New Mexico, Albuquerque, NM, USA (e-mail: rlwilliams@salud.unm.edu).

Funders Research reported in this paper was supported by the National Institute on Minority Health and Health Disparities of the National Institutes of Health under Award Numbers R01MD006073 and P20MD004811.

Compliance with Ethical Standards:

Prior Presentations: The findings of this study have been presented in part or in whole at the following conferences: (1) 13th Annual AAMC Health Workforce Research Conference, Arlington, VA, May 2017; (2) Society of General Internal Medicine Annual Meeting, Washington DC, April 2017; (3) Society of Teachers of Family Medicine Conference on Medical Student Education, Anaheim, CA, February 2017; (4) 43rd Annual meeting of North American Primary Care Research Group, Cancun, Mexico, November 2015.

Conflict of Interest: The authors declare that they do not have a conflict of interest.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

1. **Smedley BD, Stith AY, Nelson AR, eds.** Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington, DC: National Academies Press, 2003.
2. Agency for Healthcare Research and Quality. 2016 National Healthcare Quality and Disparities Report. Available at: <http://www.ahrq.gov/research/findings/nhqrdr/nhqrdr16/index.html>. Accessed April 24, 2018.
3. **Fiscella K, Sanders MR.** Racial and ethnic disparities in the quality of health care. *Ann Rev Public Health* 2016; 37:375–94.
4. Centers for Medicare and Medicaid Services, Office of Minority Health and RAND Corporation. Racial and ethnic disparities by gender in health care in Medicare Advantage. Available at: <https://www.cms.gov/About-CMS/Agency-Information/OMH/Downloads/Health-Disparities-Racial-and-Ethnic-Disparities-by-Gender-National-Report.pdf>. Accessed April 24, 2018.
5. **Hall WJ, Chapman MV, Lee KM, et al.** Implicit racial/ethnic bias among health care professionals and its influence on health care outcomes: a systematic review. *Am J Public Health*. 2015;105:e60–76.

6. **Dovidio JF, Fiske ST.** Under the radar: how unexamined biases in decision-making processes in clinical interactions can contribute to health care disparities. *Am J Public Health.* 2012;102: 945-952.
7. **van Ryn M.** Research on the provider contribution to race/ethnicity disparities in medical care. *Med Care.* 2002;40:1140-51.
8. **Fincher C, Williams JE, MacLean V, Allison JJ, Kiefe CI, Canto J.** Racial disparities in coronary heart disease: a sociological view of the medical literature on physician bias. *Ethn Dis.* 2004;14:360-371.
9. **Blair IV, Steiner JF, Fairclough DL, et al.** Clinicians' implicit ethnic/racial bias predicts patients' perceptions of care among Black but not Latino patients. *Ann Fam Med.* 2013;11:43-52.
10. **Maserejian NN, Link CL, Lutfey KL, Marceau LD, McKinlay JB.** Disparities in physicians' interpretations of heart disease symptoms by patient gender: results of a video vignette factorial experiment. *J Womens Health (Larchmt).* 2009;18:1661-7.
11. **Green AR, Carney DR, Pallin DJ, et al.** Implicit bias among physicians and its prediction of thrombolysis decisions for Black and White patients. *J Gen Intern Med.* 2007;22:1231-1238.
12. **Cooper LA, Roter DL, Carson KA, et al.** The associations of clinicians' implicit attitudes about race with medical visit communication and patient ratings of interpersonal care. *Am J Public Health.* 2012;102:979-987.
13. **Sabin JA, Greenwald AG.** The influence of implicit bias on treatment recommendations for 4 common pediatric conditions: pain, urinary tract infection, attention deficit hyperactivity disorder, and asthma. *Am J Public Health.* 2012;102:988-995.
14. **Paradies Y, Truong M, Priest N.** A systematic review of the extent and measurement of healthcare provider racism. *J Gen Intern Med.* 2013;29:364-87.
15. **Chapman EN, Kaatz A, Carnes M.** Physicians and implicit bias: how doctors may unwittingly perpetuate health care disparities. *J Gen Intern Med.* 2013;28:1504-10.
16. **Schulman KA, Berlin JA, Harless W, et al.** The effect of race and sex on physicians' recommendations for cardiac catheterization. *N Engl J Med.* 1999; 340:618-25.
17. **Van Ryn M, Burke J.** The effect of patient race and socio-economic status on physicians' perceptions of patients. *Soc Sci Med.* 2000;50:813-28.
18. **Todd KH, Samaroo N, Hoffman JR.** Ethnicity as a risk factor for inadequate emergency department analgesia. *JAMA.* 1993;269:1537-9.
19. **Rathore SS, Lenert LA, Weinfurt KP, et al.** The effects of patient sex and race on medical students' ratings of quality of life. *Am J Med.* 2000;108:561-6.
20. **Haider AH, Sexton J, Sriram N, et al.** Association of unconscious race and social class bias with vignette-based clinical assessments by medical students. *JAMA.* 2011;306:942-51.
21. **Chiaromonte GR, Friend R.** Medical students' and residents' gender bias in the diagnosis, treatment, and interpretation of coronary heart disease symptoms. *Health Psychol.* 2006;25:255-66.
22. **Williams RL, Romney C, Kano M, et al.** Racial, gender, and socioeconomic status bias in senior medical student clinical decision-making: a national survey. *J Gen Intern Med.* 2015;30:758-67.
23. **Liaison Committee on Medical Education.** Functions and Structure of a medical school: standards for accreditation of medical education programs leading to the MD degree. Standard 7.6 Cultural competence and health care Disparities. Available at: <http://lcme.org/publications/>. Accessed April 24, 2018.
24. **Association of American Medical Colleges.** Assessing Change: Evaluating Cultural Competence Education and Training. Available at: <https://www.aamc.org/download/427350/data/assessingchange.pdf>. Accessed April 24, 2018.
25. **Addison RB.** A grounded hermeneutic editing approach. In: Crabtree BF, Miller WL, editors. *Doing Qualitative Research.* 2nd Ed. Thousand Hills, CA: Sage; 1999:145-61.
26. **Burgess D, van Ryn M, Dovidio J, Saha S.** Reducing racial bias among health care providers: lessons from social-cognitive psychology. *J Gen Intern Med.* 2007;22:882-7.
27. **Project Implicit.** Implicit Association Test. Available at: <https://implicit.harvard.edu/implicit/takeatest.html>. Accessed April 24, 2018.
28. **Burgess DJ, Burke SE, Cunningham BA, et al.** Medical students' learning orientation regarding interracial interactions affects preparedness to care for minority patients: a report from Medical Student CHANGES. *BMC Med Educ.* 2016;29:254.
29. **van Ryn M, Hardeman R, Phelan SM, et al.** Medical school experiences associated with change in implicit racial bias among 3547 students: a medical student CHANGES study report. *J Gen Intern Med.* 2015;30:1748-5.