

1 **Title:** Diversity During Recruitment At An Internal Medicine Residency Program

2

3 **Short Title:** Diversity

4

5 **Authors:** Sarwan Kumar MD, Assistant Professor, Internal Medicine

6 Deepak Gupta MD, Clinical Assistant Professor, Anesthesiology

7 **Affiliation:** Wayne State University, Detroit, Michigan

8 United States

9

10 **Financial Support:** None; **Conflicts of Interests:** None

11 **Corresponding Author:**

12 Dr Deepak Gupta

13 Clinical Assistant Professor, Anesthesiology

14 Detroit Medical Center/Wayne State University

15 Box No 162,

16 3990 John R, Detroit, MI 48201, United States

17 Ph: 1-313-745-7233; Fax: 1-313-993-3889

18 Email: [dgupta@med.wayne.edu](mailto:dgupta@med.wayne.edu)

19 Word Count: Abstract: 224; Manuscript (excluding references): 1564

20

21 Acknowledgements

22

23 The authors are thankful to Dr Ronald Thomas, Children's Research Center of Michigan, Wayne

24 State University, for overseeing the statistical analysis for this project.

25

26 Abstract

27

28 Background: The right problem for graduate medical education (GME) program directors is  
29 whether diversity in their GME programs is as good as diversity in feeder entities to their GME  
30 programs. Generally, the feeder entities to GME residency programs are their affiliated medical  
31 schools. However, the specific feeder entities to GME residency programs are the unfiltered  
32 applicants' pool who apply to these programs through Electronic Residency Application  
33 Service® (ERAS®).

34 Objectives: To analyze associations in diversity among the GME applicants, the GME  
35 interviewees and the GME residents within an internal medicine residency program assuming  
36 that unfiltered applicants' pool is the specific feeder entity to the analyzed GME program.

37 Methods: We analyzed associations in age-group, gender, ethnicity and race diversity among the  
38 GME applicants, the GME interviewees and the GME residents within an internal medicine  
39 residency program for ERAS® 2018-2020 seasons to decipher Cramer's V as association  
40 coefficients ("diversity scores").

41 Results: The only significant finding was that among Not Hispanic or Latino ethnicity  
42 applications, race of ERAS® applicants had a very weak association with them being called for  
43 interviews or them becoming residents during ERAS® 2019 season as well as during the entire  
44 three-season-period (2018-2020).

45 Conclusion: Race of Not Hispanic or Latino ethnicity ERAS® applicants had a very weak  
46 association with them being called for interviews or them becoming residents at the analyzed  
47 internal medicine residency program.

48

## 49 Introduction

50

51 The benefits<sup>1</sup> of diversity in workforce have been made well-known for quite some time. The  
52 factors<sup>2</sup> interfering with recruitment and retention of diverse workforces are discussed all the  
53 time. The unclear and abstract diversity statements<sup>3</sup> have often raised eyebrows. However, the  
54 discussions have rarely focused on type III errors<sup>4-6</sup> and type IV errors when quantifying  
55 diversity or factors affecting diversity in workforce. Henceforth, solving the wrong problem  
56 accidentally (type III error) or intentionally (type IV error) ignores the critical thought process  
57 needed to identify and define the right problem. For example, while testing the null hypothesis  
58 that diversity is not significantly different among physicians compared to that among general  
59 population, it is important to not only avoid (a) false positives (type I error) and (b) false  
60 negatives (type II error) but also recognize (c) type III errors among true positives of wrong null  
61 hypothesis warranting the correction of null hypothesis as a comparison of diversity among  
62 graduating medical residents with that among graduating medical students and (d) type IV errors  
63 of interpretations about true positives leading to administrative actions over-correcting or even  
64 reversing diversity among matriculating medical residents as compared to that among  
65 matriculating medical students. Therefore, it is our hypothesis (previously published as  
66 perspective paper)<sup>7</sup> that the right problem for graduate medical education (GME) program  
67 directors is whether diversity in their GME programs<sup>8</sup> is as good as diversity in feeder entities to  
68 their GME programs. Generally, the feeder entities to GME residency programs are their  
69 affiliated medical schools. However, the specific feeder entities to GME residency programs are  
70 the unfiltered applicants' pool who apply to these programs through Electronic Residency  
71 Application Service® (ERAS®). More specifically, the truest feeder entities are the filtered

72 ERAS® applicants' pools who meet GME residency programs' specific filtering eligibility  
73 criteria because their applications are the only ones eventually reviewed in detail by GME  
74 residency programs' recruitment teams.

75

76 To test our hypothesis, the current study was designed to analyze associations in diversity among  
77 the GME applicants, the GME interviewees and the GME residents within an internal medicine  
78 residency program where co-author (SK) has been residency program director since 2017.

79

80

81 Methods

82

83 After being approved by institutional review board as non-human participant research, ERAS®  
84 database of co-author (SK) as program director was mined and anonymously tabulated by GME  
85 applicants', interviewees' and residents' age group, gender, ethnicity and race at SK's internal  
86 medicine residency program for ERAS® 2018-2020 seasons and for this anonymized data-  
87 mining, permission from or manuscript review by ERAS® did not seem warranted. These  
88 anonymized tabulations were then analyzed to decipher Cramer's V as association coefficients  
89 ("diversity scores") among applicants, interviewees and residents: age-group-specific "diversity  
90 scores", gender-specific "diversity scores", ethnicity-specific "diversity scores" and race-specific  
91 "diversity scores" for each year as well as for the entire three-season-period (2018-2020).

92 Although a multivariable Cox proportional hazards regression model to calculate hazard ratio  
93 would have been better, we chose Cramer's V for its simplicity and accessibility at Vassar Stats  
94 as up to 5x5 contingency table online calculators with age group/gender/ethnicity/race categories  
95 as independent variables on its rows and applicants/interviewees/residents as "outcomes" on its  
96 columns.

97

98 Statistically, Cramer's V tells the strength of association only when Chi-squared test evaluation  
99 of proportions is significant ( $P < 0.05$ ). Applicably, Cramer's  $V=0$  being no association among the  
100 variables was presumed to mean that age group, gender, ethnicity and race of ERAS® applicants  
101 had never been associated with them being called for interviews or them becoming residents  
102 while Cramer's  $V=1$  being perfect association among the variables was presumed to mean that  
103 age group, gender, ethnicity and race of ERAS® applicants had always been associated with

104 them being called for interviews or them becoming residents. Essentially, we christened the  
105 numerical values (0-1) of only statistically significant Cramer's V as analyzed program's  
106 "diversity scores" with statistically significant Cramer's V=0 meaning diversity perfectly  
107 matched among applicants, interviewees and residents while statistically significant Cramer's  
108 V=1 meaning diversity perfectly denied among applicants, interviewees and residents.

109  
110 During data-mining, it was observed that certain number of applications had missing  
111 documentations about applicants' age, gender, ethnicity and race. Although this might have been  
112 most likely due to applicants choosing ERAS® blinded recruitment processes, we excluded all  
113 these missing documentations during our "diversity scores" analysis as tabulated in Tables 1-  
114 2(a). Additionally, as ERAS® allows free-form unstructured text along with multiple choices  
115 from among too many pre-defined options for ethnicity and race segments, we had to simplify  
116 documented ethnicity and race data before tabulating them as Tables 2 according to pre-defined  
117 options as outlined by United States Census Bureau™.<sup>9</sup> While Not Hispanic or Latino ethnicity  
118 applications had their documented race, Hispanic or Latino ethnicity applications had no clearly  
119 documented race. Therefore, while ethnicity-specific "diversity scores" were possible among two  
120 ethnic groups (Table 2(a)), race-specific "diversity scores" were possible only among Not  
121 Hispanic or Latino ethnicity (Table 2(b)). Moreover, age-group >54 years (Table 1(a)) and  
122 among Not Hispanic or Latino ethnicity, American Indian and Alaskan Native alone race (Table  
123 2(b)) were also excluded due to their almost zero numbers.

124

125 Results

126

127 As shown in Tables 1-2(a), age-group-specific, gender-specific and ethnicity-specific “diversity  
128 scores” were not statistically significant and thus their Cramer’s V (level of association’s  
129 strength) were irrelevant during ERAS® 2018-2020 seasons at the analyzed internal medicine  
130 residency program. Among race-specific “diversity scores” among Not Hispanic or Latino  
131 ethnicity applications as shown in Table 2(b), P-value was significant (<0.01) during ERAS®  
132 2019 season as well as during the entire three-season-period (2018-2020) but Cramer’s V  
133 (“diversity scores”) in both instances were very close to 0 as shown in Table 2(b) (0.07 for 2019  
134 season; 0.04 for 2018-2020 period) indicating that race of ERAS® applicants had a very weak  
135 association with them being called for interviews or them becoming residents during ERAS®  
136 2019 season as well as during the entire three-season-period (2018-2020).

137

138

139 Discussion

140

141 The only key finding of our non-human participant research was that among Not Hispanic or  
142 Latino ethnicity applications, race of ERAS® applicants had a very weak association with them  
143 being called for interviews or them becoming residents. To validate or refute this finding in the  
144 future, we strongly urge ERAS® that instead of allowing applicants to leave age, gender,  
145 ethnicity and race segments empty, it should mandate such documentations by all applicants  
146 wherein they can always choose the pre-defined option “Decline to answer” when documenting  
147 their age, gender, ethnicity and race. Similarly, instead of free-form unstructured text, ERAS®  
148 ethnicity and race segment should have structured text format allowing only one option to choose  
149 from among pre-defined options as outlined by United States Census Bureau™.<sup>9</sup> This will allow  
150 ERAS® ethnicity and race data to be more appropriately followed for tracking and aligning  
151 future “diversity scores” by individual GME programs locally and by ERAS® regionally-  
152 nationally.

153

154 During our current analysis, we could not investigate associations among the filtered ERAS®  
155 applicants’ pools during 2018-2020 seasons who had met the analyzed internal medicine  
156 residency program’s specific filtering eligibility criteria. Therefore, our current analysis had to  
157 use the unfiltered applicants’ pools who had applied to the analyzed internal medicine residency  
158 program during ERAS® 2018-2020 seasons. In the future ERAS® seasons, we are hoping to  
159 utilize our “diversity scores” methodology by prospectively investigating association among all  
160 filtered applications (all reviewed applicants) in terms of age-group, gender, ethnicity and race  
161 with those among all scheduled interviewees. Thereafter, the analyzed internal medicine program

162 can average the lowest rank numbers at which the ranked interviewees had matched as residents  
163 over the prior three years. Assuming that the analyzed internal medicine program matriculates 12  
164 residents annually and the three-year averaged number for the lowest rank at which the ranked  
165 interviewees had matched is 50, the “diversity scores” methodology can be further applied to  
166 investigate association among all scheduled interviewees with the first 50-ranked interviewees  
167 regarding their age-group, gender, ethnicity and race. Essentially, compared to current “diversity  
168 scores” methodology retrospectively investigating association among unfiltered applicants with  
169 scheduled interviewees with matched residents, the future age-group-specific, gender-specific,  
170 ethnicity-specific and race-specific “diversity scores” may attempt to prospectively correct  
171 inadvertent association among age-group, gender, ethnicity and race of filtered-and-reviewed  
172 applicants with scheduled interviewees with first 50-ranked interviewees because these are the  
173 ones which are under primary control of residency programs’ recruitment teams. Alternatively,  
174 diversity can be simply realigned prospectively with our diversity realignment score as explained  
175 in our supplementary excel file of simulated data along with its explanation in our supplementary  
176 video file whose detailed and expanded version can be additionally viewed on YouTube.<sup>10</sup>

177

178 There were few additional limitations to our retrospective analysis. A multivariable Cox  
179 proportional hazards regression model to calculate hazard ratio as applicants evolve to being  
180 interviewees to finally become residents may have been better than Cramer’s V for determining  
181 the effect of each independent variable while controlling the effect of other independent  
182 variables (age-group/gender/ethnicity/race). Data imputation approach may have prevented  
183 exclusions of missing documentations about ethnicity/race. Our single-institution analysis for  
184 brief period of three years needs further validation by multi-institutional prospective studies

185 utilizing the above-mentioned better statistical methods. Once our results have been validated by  
186 multi-institutional prospective studies, our “diversity scores” methodology may turn out to be a  
187 simple method to guide GME program directors during their attempts to ensure age-group-  
188 specific, gender-specific, ethnicity-specific and race-specific diversity among the recruited GME  
189 workforce in the future.  
190

191 Conclusion

192

193 Race of Not Hispanic or Latino ethnicity ERAS® applicants had a very weak association with  
194 them being called for interviews or them becoming residents at the analyzed internal medicine  
195 residency program.

196

197   References

198

199       1. Gomez LE, Bernet P. Diversity improves performance and outcomes. *J Natl Med Assoc.*  
200           2019;111(4):383-392.

201       2. Aysola J, Barg FK, Martinez AB, et al. Perceptions of Factors Associated With Inclusive  
202           Work and Learning Environments in Health Care Organizations: A Qualitative Narrative  
203           Analysis. *JAMA Netw Open.* 2018;1(4):e181003.

204       3. Carnes M, Fine E, Sheridan J. Promises and Pitfalls of Diversity Statements: Proceed  
205           With Caution. *Acad Med.* 2019;94(1):20-24.

206       4. Özdemir V, Springer S. What does "Diversity" Mean for Public Engagement in Science?  
207           A New Metric for Innovation Ecosystem Diversity. *OMICS.* 2018;22(3):184-189.

208       5. Singer DA. Solving the Wrong Resource Assessment Problems Precisely. In: Daya Sagar  
209           B, Cheng Q, Agterberg F, eds. *Handbook of Mathematical Geosciences.* Springer, Cham;  
210           2018.

211       6. Silvers A, Mitroff II. *Dirty Rotten Strategies: How We Trick Ourselves and Others Into*  
212           *Solving the Wrong Problems Precisely.* United States: Stanford Business Books; 2010.

213       7. Gupta D, Kumar S. Diversifying your GME program. *Residency Program Alert*, April 8,  
214           2020. <http://www.hcpro.com/RES-333644-2699/Diversifying-your-GME-program.html>  
215           Accessed April 14, 2020.

216       8. Deville C, Hwang WT, Burgos R, Chapman CH, Both S, Thomas CR Jr. Diversity in  
217           Graduate Medical Education in the United States by Race, Ethnicity, and Sex, 2012.  
218           *JAMA Intern Med.* 2015;175(10):1706-1708.

- 219 9. United States Census Bureau<sup>TM</sup>. Race & Ethnicity.  
220 <https://www.census.gov/topics/population/race/about.html> Accessed August 26, 2020.
- 221 10. YouTube. Kumar-Gupta envisaged diversity realignment score explained.  
222 <https://www.youtube.com/watch?v=nzHh7hl3B38> Accessed November 24, 2020.  
223

224 Table 1 (a): Age-Group-Specific “Diversity Scores” During Recruitment At An Internal Medicine Residency  
 225 Program In 2018-2020

AGE GROUP	<25 years	25-34 years	35-44 years	45-54 years	“Diversity Score” Cramer’s V (P Value)	EXCLUDED	
						>54 years	Missing Data
2018							
Applicants (n=3294)	254 (8%)	2822 (86%)	191 (6%)	18 (1%)	0.03 (0.62)	0 (0%)	9 (0%)
Interviewees (n=107)	10 (9%)	94 (88%)	3 (3%)	0 (0%)		0 (0%)	0 (0%)
Residents (n=13)	2 (15%)	11 (85%)	0 (0%)	0 (0%)		0 (0%)	0 (0%)
2019							
Applicants (n=3820)	300 (8%)	3262 (85%)	233 (6%)	23 (1%)	0.03 (0.23)	0 (0%)	2 (0%)
Interviewees (n=128)	6 (5%)	119 (93%)	3 (2%)	0 (0%)		0 (0%)	0 (0%)
Residents (n=12)	0 (0%)	12 (100%)	0 (0%)	0 (0%)		0 (0%)	0 (0%)
2020							
Applicants (n=3881)	279 (7%)	3266 (84%)	236 (6%)	23 (1%)	0.02 (0.89)	2 (0%)	75 (2%)
Interviewees (n=138)	12 (9%)	117 (85%)	7 (5%)	0 (0%)		0 (0%)	2 (1%)
Residents (n=12)	1 (8%)	11 (92%)	0 (0%)	0 (0%)		0 (0%)	0 (0%)
2018-2020							
Applicants (n=10995)	833 (8%)	9350 (85%)	660 (6%)	64 (1%)	0.02 (0.17)	2 (0%)	86 (1%)
Interviewees (n=373)	28 (8%)	330 (88%)	13 (3%)	0 (0%)		0 (0%)	2 (1%)
Residents (n=37)	3 (8%)	34 (92%)	0 (0%)	0 (0%)		0 (0%)	0 (0%)

226

227

228 Table 1 (b): Gender-Specific “Diversity Scores” During Recruitment At An Internal Medicine Residency  
 229 Program In 2018-2020

GENDER	Male	Female	“Diversity Score” Cramer’s V (P Value)	EXCLUDED
				Missing Data
2018				
Applicants (n=3294)	1961 (60%)	1333 (40%)	0.01 (0.89)	0 (0%)
Interviewees (n=107)	65 (61%)	42 (39%)		0 (0%)
Residents (n=13)	7 (54%)	6 (46%)		0 (0%)
2019				
Applicants (n=3820)	2168 (57%)	1652 (43%)	0.03 (0.29)	0 (0%)
Interviewees (n=128)	64 (50%)	64 (50%)		0 (0%)
Residents (n=12)	6 (50%)	6 (50%)		0 (0%)
2020				
Applicants (n=3881)	2224 (57%)	1656 (43%)	0.01 (0.77)	1 (0%)
Interviewees (n=138)	81 (59%)	57 (41%)		0 (0%)
Residents (n=12)	8 (67%)	4 (33%)		0 (0%)
2018-2020				
Applicants (n=10995)	6353 (58%)	4641 (42%)	0.01 (0.84)	1 (0%)
Interviewees (n=373)	210 (56%)	163 (44%)		0 (0%)
Residents (n=37)	21 (57%)	16 (43%)		0 (0%)

230

231

232 Table 2 (a): Ethnicity-Specific “Diversity Scores” During Recruitment At An Internal Medicine Residency  
 233 Program In 2018-2020

ETHNICITY	Not Hispanic or Latino	Hispanic or Latino	“Diversity Score” Cramer’s V (P Value)	EXCLUDED
				Missing Data
2018				
Applicants (n=3294)	2902 (88%)	125 (4%)	0.01 (0.71)	267 (8%)
Interviewees (n=107)	96 (89%)	4 (4%)		7 (7%)
Residents (n=13)	10 (77%)	1 (8%)		2 (15%)
2019				
Applicants (n=3820)	3352 (87%)	178 (5%)	0.03 (0.21)	290 (8%)
Interviewees (n=128)	114 (89%)	2 (2%)		12 (9%)
Residents (n=12)	9 (75%)	0 (0%)		3 (25%)
2020				
Applicants (n=3881)	3449 (89%)	155 (4%)	0.01 (0.66)	277 (7%)
Interviewees (n=138)	124 (90%)	4 (3%)		10 (7%)
Residents (n=12)	9 (75%)	0 (0%)		3 (25%)
2018-2020				
Applicants (n=10995)	9703 (88%)	458 (4%)	0.01 (0.36)	834 (8%)
Interviewees (n=373)	334 (89%)	10 (3%)		29 (8%)
Residents (n=37)	28 (75%)	1 (3%)		8 (22%)

234

235

236 Table 2 (b): Not-Hispanic-Or-Latino-Race-Specific “Diversity Scores” During Recruitment At An Internal  
 237 Medicine Residency Program In 2018-2020

NOT HISPANIC OR LATINO RACE	White alone	Black or African American alone	Asian alone	Some other race alone	Two or more races	“Diversity Score” Cramer’s V (P Value)	EXCLUDED American Indian and Alaska Native alone
2018							
Applicants (n=2902)	681 (23%)	200 (7%)	1676 (58%)	199 (7%)	146 (5%)	0.04 (0.36)	0 (0%)
Interviewees (n=96)	28 (29%)	3 (3%)	48 (50%)	11 (12%)	6 (6%)		0 (0%)
Residents (n=10)	3 (30%)	0 (0%)	6 (60%)	1 (10%)	0 (0%)		0 (0%)
2019							
Applicants (n=3352)	750 (22%)	233 (7%)	1964 (59%)	231 (7%)	172 (5%)	0.07 (<0.01)	2 (0%)
Interviewees (n=114)	47 (41%)	3 (3%)	48 (42%)	5 (4%)	11 (10%)		0 (0%)
Residents (n=9)	2 (22%)	0 (0%)	5 (56%)	1 (11%)	1 (11%)		0 (0%)
2020							
Applicants (n=3449)	742 (21%)	253 (7%)	2069 (60%)	225 (7%)	159 (5%)	0.03 (0.46)	1 (0%)
Interviewees (n=124)	37 (30%)	10 (8%)	63 (51%)	9 (7%)	5 (4%)		0 (0%)
Residents (n=9)	3 (33%)	0 (0%)	5 (56%)	1 (11%)	0 (0%)		0 (0%)
2018-2020							
Applicants (n=9703)	2173 (22%)	686 (7%)	5709 (59%)	655 (7%)	477 (5%)	0.04 (<0.01)	3 (0%)
Interviewees (n=334)	112 (33%)	16 (5%)	159 (48%)	25 (7%)	22 (7%)		0 (0%)
Residents (n=28)	8 (29%)	0 (0%)	16 (57%)	3 (11%)	1 (3%)		0 (0%)

238 .