Using holistic review to form a diverse interview pool for selection to medical school

Leila E. Harrison, PhD, MA, MEd

Admissions, Recruitment, and Inclusion, Washington State University Elson S. Floyd College of Medicine, Spokane, Washington

ABSTRACT

The holistic review in admissions framework has gained ground in medical schools. Because holistic review is unique at each institution, there is a paucity of evidence about whether it produces a more diverse interview pool than metrics-driven processes. The aim of this quantitative causal-comparative replication study was twofold: (1) to assess whether holistic review produced a more diverse interview group than one based solely on metrics and (2) to assess how the students enrolled through holistic review performed compared to national averages. Participants included 4643 medical school applicants applying for entering years 2011 through 2015. Three interview subgroups included a holistic review group ($n=1505$), an academic group ($n=1505$), and an overlap group ($n=1633$). The sample included 44% women, 11.9% first-generation college students, and 14.9% underrepresented in medicine. Analyses found that in all categories of demographics and experiences, the holistic review group had significantly higher percentages than the academic group. One class performed lower than the national average on both United States Medical Licensing Exam Step 1 and Step 2 Clinical Knowledge; however, the other two classes performed similar to students nationally. This study supports the view that holistic review produces a more diverse interview pool than a metrics pool and is a valuable tool for increasing broad diversity.

KEYWORDS Diversity; holistic review; medical school admissions; medical students

The Association of American Medical Colleges developed the holistic review framework encouraging medical schools to balance the consideration of an applicant's experience, attributes, and metrics to enroll a more diverse student body.¹ Admissions processes that focus heavily on metrics result in less diversity in their classes.² The physician workforce does not reflect the increasingly diverse population of the USA, because physicians tend to come from more affluent and white and Asian racial backgrounds, the latter of which make up more than 75% of the matriculating medical students in the USA.³ Holistic review meets the US Supreme Court's standard of considering how an applicant might contribute to a diverse medical education environment.⁴ Grabowski assessed differences in gender, applicants underrepresented in medicine, first-generation college students, socioeconomic status, age, educational level, hours of employment,

community service hours, and hours of health care exposure between a holistic review interview group and one based on the Medical College Admission Test (MCAT) and undergraduate grade point average (UGPA) and found that the holistic review interview group was significantly more diverse.⁵ Similarly, the current retrospective causal-comparative replication study aimed to assess whether holistic review produced a more diverse interview group than an interview group based on metrics (i.e., MCAT and UGPA). A secondary aim was to assess how the students enrolled through holistic review performed in comparison to national data. Unlike Grabowski's study, which was conducted at a new, privately funded medical school in Michigan that does not distinguish between in-state and out-of-state applicants,⁵ the current study was conducted at a publicly funded medical school in Texas with a state mandate to accept 90% Texas residents.

Corresponding author: Leila E. Harrison, PhD, MA, MEd, Associate Dean for Admissions, Recruitment, and Inclusion, Washington State University Elson S. Floyd College of Medicine, PO Box 1495, Spokane, WA 99210-1495 (e-mail: leila.harrison@wsu.edu)

This research was conducted at Texas A&M University College of Medicine during the author's employment. Findings from this study were presented at an Association of American Medical Colleges Conference in 2018.

Received November 24, 2018; Revised January 12, 2019; Accepted January 17, 2019.

Table 1. Comparison of holistic group ($n = 1505$) and academic group ($n = 1505$) on categorical variables on application

Variable	Holistic group	Academic group	Chi-square	<i>P</i> value	Cramer's V^a
Gender	51.6%	33.2%	113.4	<0.0005	0.156/small
Underrepresented ethnicity	30.4%	5.2%	435.8	<0.0005	0.306/moderately strong
First-generation	17.1%	10.1%	60.9	<0.0005	0.114/small
Socioeconomic status	17.3%	7.3%	104.9	<0.0005	0.150/small
Degree attainment			171.5	<0.0005	0.136/small
Bachelor's	85.1%	94%			
Master's	14.3%	4.8%			
Doctorate	0.6%	1.2%			

^aPer Cohen.⁷

METHODS

Three groups were formed across five cohorts enrolling in 2011 through 2015. The actual holistic review process used by the medical school admissions office included the holistic review interview group for each cohort. To provide a comparison group based on metrics, an artificial academic interview group was formed based on applicants' MCAT and UGPA. This academic group was formed using Grabowski's method of taking the UGPA and multiplying by 10 and then adding the highest total MCAT score to produce an academic score.⁵ Applicants with the highest academic score formed this group and included the same total number as in the holistic review group. Applicants found in both groups were removed to form an overlap group so that applicants were in only one group. Applicants with missing MCAT or UGPA and those selected through special programs were excluded.

Data were collected for each applicant, including UGPA, highest composite MCAT score, gender, ethnicity, first-generation undergraduate student status (neither parent earned a bachelor's degree), socioeconomic status, age, ethnicity, advanced degree, and hours of experience in employment, community service, and health care exposure—all found in the medical school's application portal. Low socioeconomic disadvantaged status was based on the Texas Medical and Dental School Application Service's (TMDSAS) scale of A to D derived from more than 15 socioeconomically related questions within the application, with an A or B rating being the most disadvantaged.⁶ Scores on the United States Medical Licensing Exam (USMLE) Step 1 were collected for the first three cohorts and USMLE Step 2 Clinical Knowledge (CK) scores for the first cohort, mirroring Grabowski's data. These scores as well as graduation year and residency placement information were accessed through Student Affairs records. Each data point was paired from the different sources and the data set was de-identified prior to analysis.

The school's holistic review process includes the review of community service, health care exposure, research, and economic and educational disadvantaged status. Texas utilizes its own application service, the TMDSAS, which provides more information than the American Medical College Application Service. The American Medical College Application Service limits applicants to reporting 15 total experiences, whereas TMDSAS has separate sections for employment, community service, and health care activities with no limitations, allowing for a more robust explanation of how applicants spent their time. Academic information such as highest composite MCAT score, UGPA, science GPA, and graduate GPA, where applicable, were also reviewed. The interview process includes the assessment of attributes like communication, maturity, service to others, and motivation.

RESULTS

There were 4643 applicants in the five cohorts. The holistic review and academic groups had 1505 applicants each, and the overlap group had 1633 applicants. The number of non-Texas residents was the same in the holistic review and academic groups. The total sample included 56% men, 44% women, 11.9% first-generation college students, 14.9% underrepresented in medicine, and 8% with an advanced degree (master's and above). Their mean UGPA was 3.75; mean MCAT, 31.6; and mean age at application, 22.5 years.

A chi-square test of independence was used to determine whether there was an association between interview groups and the categorical variables of gender, ethnicity, first-generation college student, socioeconomic status, and educational attainment. Results revealed a statistically significant association between interview groups and all variables (*Table 1*).

A one-way analysis of variance was conducted to determine whether age, employment hours, community service hours, and health care exposure hours were different for the interview groups. A statistically significant association was

Table 2. Comparison of holistic group ($n = 1505$), academic group ($n = 1505$), and overlap group ($n = 1633$) on age and experience variables on application

Variable	Group (mean $\pm$ SD)			Welch's F	P value
	Holistic	Academic	Overlap		
Age	23.2 $\pm$ 3.3	22.3 $\pm$ 2.7	22.1 $\pm$ 2.6	52.3	<0.0005
Hours employed	5336 $\pm$ 8070	3198 $\pm$ 5549	3314 $\pm$ 6463	40.0	<0.0005
Community service hours	545 $\pm$ 1369	383 $\pm$ 1260	425 $\pm$ 1097	6.1	0.002
Health care exposure hours	1505 $\pm$ 3232	642 $\pm$ 2074	795 $\pm$ 2228	39.0	<0.0005

Table 3. Comparison of enrolled class score and national average on the United States Medical Licensing Exam

Cohort	Step 1			Step 2		
	Enrolled	National	P value	Enrolled	National	P value
2011	225.06	228	0.022	237.14	240	0.013
2012	228.53	230	0.199			
2013	226.59	229	0.064			

found between interview groups and each of those variables (*Table 2*).

For the enrolled classes selected through the holistic review group with USMLE Step 1 and Step 2 CK scores, a one-sample t test analysis was used to assess the means compared with known national means. As shown in *Table 3*, though the 2011 class had significantly lower scores, the difference for the 2012 and 2013 classes were not statistically significant. Four-year graduation rates reported by the Association of American Medical Colleges included 2007 to 2009; assuming that the national graduation rates would be similar (87%), the cohorts enrolling in 2011 and 2012 had higher graduation rates of 95% and 94% matched to residency.

DISCUSSION

This study provides further evidence that using a holistic review process that balances the consideration of experiences and attributes with metrics to meet the school's mission produces a more diverse interview pool from which to choose applicants. Like Grabowski's findings, this study supports the view that holistic review is not just a framework that is important in a theoretical sense; it actually does produce a more diverse interview pool to make important selection decisions from. Selection for interview is a critical point to implement holistic review because this is where the applicant pool is drastically narrowed for most programs. Though more schools are using holistic review at some stages in the admissions process, in one study, 120 medical schools

reported that MCAT and UGPA were the most important in deciding who to interview.⁸ The results of this study show clear evidence that if a process filters an applicant pool primarily or solely based on MCAT and GPA, the richness of broad diversity is compromised compared to when a balanced consideration of metrics, experience, and attributes is used. Having a broader, diverse interview pool allows schools to broaden access to medical education for groups historically underrepresented in medicine.

There were a few differences between this study and that of Grabowski. Statistically significantly more older students were found in the holistic review group in this study, whereas Grabowski found no differences based on age.⁵ Additionally, one class performed lower than the national average in both USMLE Step 1 and Step 2 CK, whereas the other two classes for whom USMLE Step 1 scores were assessed performed similar to the national pool; Grabowski's study found no differences in these outcomes. It is unclear why this one class performed lower than the national average, but it is worth noting that the class went on to graduate above the national average rate and matched to residency at a 94% rate, which was at the national average (93.9%) reported by the National Residency Match Program for that year.⁹

Limitations of the study include not measuring full outcomes (i.e., USMLE Step 2 CK, graduation rates, and matching rates) for each cohort assessed. This was intentional to fully replicate Grabowski's study; however, having the full set of data would be helpful to know whether there are inconsistencies or outliers in class performance that may provide feedback to the faculty. An additional limitation is that the academic group was artificially selected and not an actual selected group, but it could provide insight to those programs whose selection processes heavily weight metrics. Another limitation is that it is unknown whether the enrolled cohorts reflected the holistic review interview group composition and experiences.

Progress is needed in training a more diverse physician workforce to meet the population's needs. Being a competent, compassionate physician is not dependent on metrics and academics alone. Students bring with them many experiences and attributes that enhance the learning for all. There

are barriers to enrolling a diverse medical student body when metrics are the only factors assessed. Therefore, expansion to a process that considers how experiences and attributes are also important in assessing medical school applicants' preparedness for medical school opens the door for enrolling a more diverse student body.

ORCID

Leila E. Harrison <http://orcid.org/0000-0002-6804-075X>

1. Addams AN, Bletzinger RB, Sondheimer HM, White SE, Johnson LM. *Roadmap to Diversity: Integrating Holistic Review Practices Into Medical School Admission Processes*. Washington, DC: Association of American Medical Colleges; 2010.
2. Heller CA, Rua SH, Mazumdar M, Moon JE, Bardes C, Gotto AM. Diversity efforts, admissions, and national rankings: can we align priorities? *Teach Learn Med*. 2014;26:304–311. doi:10.1080/10401334.2014.910465.
3. Association of American Medical Colleges. *Diversity in Medical Education: Facts and Figures 2016*. <http://www.aamcdiversityfactsand-figures2016.org/>. Accessed May 1, 2017.
4. Mahon KE, Henderson MK, Kirch DG. Selecting tomorrow's physicians: the key to the future health care workforce. *Acad Med*. 2013; 88:1806–1811. doi:10.1097/ACM.0000000000000023.
5. Grabowski CJ. Impact of holistic review on student interview pool diversity. *Adv Health Sci Educ*. 2018;23:487–498. <https://doi.org/10.1007/s10459-017-9807-9>.
6. Texas Medical and Dental Schools Application Service. *School Portal User Guide*. <http://www.tmdsas.com/School%20Portal%20Training%20Guide.pdf>. Accessed May 1, 2017.
7. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.
8. Monroe A, Quinn E, Samuelson W, Dunleavy DM, Dowd KW. An overview of the medical school admission process and use of applicant data in decision making: what has changed since the 1980s? *Acad Med*. 2013;88:672–681. doi:10.1097/ACM.0b013e31828bf252.
9. National Residency Matching Program. *Results and Data: 2015 Main Residency Match*. http://www.nrmp.org/wp-content/uploads/2015/05/Main-Match-Results-and-Data-2015_final.pdf. Accessed October 15, 2017.