JRMC Journal of Regional Medical Campuses

Main and Regional Campus Assessments of Applicants to a Rural Physician Leadership Program: A Generalizability Analysis

Terry Stratton, Clarence Kreiter, PhD, Carol Elam, EdD

DOI: https://doi.org/10.24926/jrmc.v2i1.1981
Journal of Regional Medical Campuses, Vol. *2*, Issue 2 (2019)

Published by University of Minnesota Libraries Publishing

LIBRARIES PUBLISHING

Main and Regional Campus Assessments of Applicants to a Rural Physician Leadership Program: A Generalizability Analysis

Terry Stratton, Clarence Kreiter, PhD, Carol Elam, EdD

Introduction

At its most basic level, the selection of qualified candidates remains a defining characteristic of any profession.[1] In medicine, emphasis has shifted in terms of the desired qualities of future physicians[2] and the means to assess those qualities.[3-5] For example, holistic approaches advocate consideration of multiple factors other than academic performance,[6-8] including an individual's "fit" with a medical program's social mission.[9-10] Yet, underlying most selection criteria or methods are individual ratings or judgments of some personal quality, aptitude, or behavior[11] is information limited by the reliability of the ratings and the representativeness of the encounter.[12,13]

For programs with specific foci, such as those accepting applicants directly into rural paths, tracks, or concentrations, the admission process may accommodate supplemental values, interests, or stakeholder perspectives. As such, training targeted to certain practice locales (e.g. rural or underserved),[14-16] medical specialties (e.g. primary care),[15,16] or career interests (e.g. physician scientists)[17] may demand an expanded approach to selecting qualified applicants for a specific programmatic fit.

A changing dimension of the admission process, and the academic qualifications, personal qualities, and/or demographic considerations it entails, is the increasing presence of regional medical campuses (RMCs). RMCs are defined as "campuses of medical schools at which a portion of pre-clinical or clinical education of medical students occurs",[18] and play a significant role in calls for increased enrollment.[19-20] RMCs are classified into 4 models (basic science, clinical, longitudinal, and combined) based on curricular years taught and/or type of training provided[18] and can target a specific mission, demographic, specialty focus, and/or delivery model (e.g. community-based care). While RMCs may provide all aspects of medical training, they are usually considered extensions of the main or "parent" program with selection decisions made by a single, overarching admission committee.

Located in the southeastern United States (US), the Rural Physician Leadership Program (RPLP) was created in 2008 at the University of Kentucky College of Medicine (UKCOM) to attract and train applicants interested in practicing rural medicine, ideally in the state of Kentucky.[21] Located in a city of approximately 320 000, students' pre-clerkship training (years M1-M2 in our program) occurs at the main urban campus, while their clinical instruction and leadership training (years M3-M4) is completed at a smaller, rural (population ~7,500) medical campus about one hour away. Ten students are admitted annually, with preference given to applicants with rural backgrounds, interests, or experiences. Like their main campus counterparts, RPLP students are free to pursue any medical specialty.

RPLP Admissions Process

While myriad factors underlie the selection of applicants to medical school,[22] the challenge for more focused programs like the RPLP is twofold: To gauge preparedness for and fit within the medical profession, and to discern interest in rural medical practice.[23,24] With final admission decisions made by a single committee, regional input into these assessments was deemed essential.

To compliment applicants' written responses to items contained on our secondary form, semi-structured, face-to-face interviews are conducted with applicants meeting certain academic standards. Interviews typically last between 30-45 minutes and are conducted by a wide range of individuals including active and retired faculty, administrative staff, community members, and current medical students and residents. Interviewer assignment is not systematic, though specific individuals may be paired with applicants with similar personal (e.g. geographic area) or professional (e.g. medical specialty or research area) backgrounds. Standardized training includes a review of program missions, the admission process (including instrumentation), and interview protocol. Over 2 consecutive days, RPLP applicants complete interviews at both main and regional medical campuses. At each site, 2 interviewers with access to standardized applicant data (e.g. prior academic performance, standardized test scores, demographic characteristics, residency status, relevant activities/experiences, and letters of evaluation) independently offer subjective, narrative assessments of applicants' backgrounds and qualifications as well as a global (overall) assessment. This assessment is assigned a 1-7 numeric rating ranging from "unacceptable" to "outstanding, clearly superior". Using a scale from 0 ("no chance") to 100 ("absolute certainty"), interviewers at both sites are also

Terry Stratton
Clarence Kreiter, PhD
Carol Elam, EdD

asked to estimate the likelihood that the RPLP applicant will ultimately practice in rural Kentucky.

Several differences between main and regional campus selection procedures are worth noting. First, unlike the review of applications to the regular MD track – which occurs on a "rolling" basis throughout the academic year – all RPLP interviews, as mentioned, are conducted during 3 dedicated, 2-day periods. All decisions to admit or reject interviewed RPLP candidates are made by the main campus admission committee with input from an appointed and voting RPLP faculty member who summarizes the opinions of the regional campus' interview board. Accepted RPLP candidates are granted admission to the regional campus program. RPLP candidates who are either placed on hold for comparison or rejected due to lack of programmatic "fit" may be considered for main campus admission.

Second, the disparate numbers of applicants to the RPLP versus the "regular" MD program, along with the local populations from which to draw interviewers, caused interviewer pools to vary in size. Indeed, across the 9-year study timeframe, the mean number of unique interviewers for RPLP candidates was 24.3 (median = 22.0, SD = 5.2) and 17.8 (median = 18, SD = 3.9) on the main and regional campuses, respectively. Admission interviewers at both campuses consisted of UKCOM faculty and administrators, community members (including practicing and retired physicians), and current medical students. The main campus interviewers also included active and emeritus basic science faculty.

Third, although key metrics and rating scales were identical, variations in the narrative portions of the semi-structured interview forms reflected RPLP interests specific to rural medical practice. For example, whereas main campus interviewers were guided to discuss applicants' general experience/knowledge of the profession and thought into choice of medical schools, RPLP forms prompted interviewers to explore applicants' understanding of rural culture/people and thought into physician/patient relationships.

With raters tasked to assess applicants' qualifications for admission to medical school and fit with the regional campus' rural focus, the guiding research questions were: 1) What is the reliability of ratings issued by regional and main campus admissions interviewers? 2) What contribution is made by regional interviewers to the overall reliability of admission ratings? and 3) What are the optimal numbers of raters at one or both locations?

Methods

The study protocol received institutional review board approval to use preexisting data from 232 RPLP applicants who, from 2009-2017, completed admission interviews. From the 22 re-applicants who interviewed in multiple years, initial rating data were used. Interviews were granted via an internal screening process which included a holistic review of each applicant but tended to emphasize measures of cognitive ability and geographical background of applicants. Of those RPLP applicants interviewed during the 9-year study period, 90 were accepted for admission and subsequently enrolled in the program.

Interviewer ratings of likelihood of rural practice and overall applicant acceptability were examined separately using a generalizability framework. Developed by Cronbach[25] and refined by Brennan[26] and others,[27] this method uses analysis-of-variance techniques to partition variance into multiple sources, or measurement facets. For any given facet, the resulting variance components (VCs) reflect how much of the total score variance can be attributed to that source.

These facets can be crossed or nested, depending on whether or not all conditions of one facet are observed with all conditions of another facet.[28] Based on whether or not the results are intended to generalize beyond the observed conditions, facets are also designated as random or fixed.[29] Fixing a facet will typically increase the estimated reliability since it limits the range over which scores are generalized. For the designated object of measurement, the resulting "universe score" is akin to "true score" variance in classical test theory.

The generalizability framework can be used in retrospective (G study) or prospective (D study) applications. In the latter, hypothetical VCs and statistics can be generated based on systematic manipulations of key measurement facets, much in the way that predicted exam reliability can be examined by altering the number of test items. Whereas G studies document what is observed in practice, D studies suggest what is theoretically possible.

Multivariate G studies allow the calculation of separate (multiple) universe scores for each level of a univariate fixed facet,[29] in this case, main and regional campus ratings. Following the convention of similar studies,[30] Brennan's notation[26] is used to represent these models: A solid (●) or empty (○) circle indicates whether a facet is crossed or nested within the multivariate variable, respectively. The resulting notation of the complete multivariate model for this study of main and regional campus admissions ratings, then, is (r○ : p●).

Descriptive statistics are generated using IBM SPSS Statistics for Windows (Version 24). Univariate and multivariate G and D studies were conducted using *GENOVA* (Version 3.1)[27] and *mGENOVA* (Version 2.1),[31] respectively. Composite (combined) universe scores were based on equal a priori weights for main and regional medical campus ratings.

Results

Descriptive Statistics

Complete sets of ratings (2 x 2 = 4) for overall acceptability and likelihood of rural in-state practice were available for 211 and 174 RPLP applicants, respectively, due primarily to main campus interviewers' failure to consistently rate the latter. Applicants with incomplete rating data did not differ significantly by race, gender, or geographic origin. The

proportion of cases excluded due to missing data varied by year, but followed no discernible pattern. Since these represented key variables of interest, and because sample sizes remained sufficient, cases with missing data were excluded rather than substituted with imputed values. Mean ratings of applicants' overall acceptability were 5.27 (median = 5.50, SD = 0.98) and 5.35 (median = 5.50, SD = 1.20) for main and regional campus interviewers, respectively. For the likelihood of practicing in rural Kentucky, mean ratings were 77.7% (median = 87.5%, SD = 20.4%) and 78.5% (median = 85.0%, SD = 19.8%) for main and regional campus interviewers, respectively. Ratings of applicant acceptability and likelihood of rural in-state practice were not significantly correlated for main (r_s = 0.04, p = .61) or regional (r_s = 0.09, p = .25) campus interviewers.

Univariate G Study Results

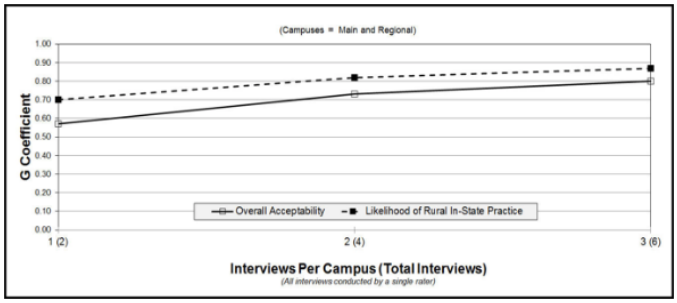
Table 1 displays univariate G study results [r : (p x c)] of overall applicant rating and likelihood of rural in-state practice. As shown, the percentages of "true score" variance associated with the object of measurement (p) were 36% and 51%, respectively. Compared to likelihood of rural practice, the variation in ratings attributable to the person-by-rater (pr) and person-by-campus (pc) interaction was notably larger for overall applicant rating (53% versus 44% and 11% versus 5%, respectively).

Table 1. Univariate Mixed Model G Study Results [r : (p x c)]

Overall Acceptability *(n = 211)*					
Effect	df	Mean Square	Variance Component	Standard Error	% Variance
Persons *(p)*	210	3.59063	0.65440	0.08877	36
Campus ©	1	0.15669	0.00000	0.00000	0
Raters : persons by campus *[r : (p x c, e)]*	422	0.97305	0.97304	0.06683	53
p x c	210	1.36086	0.19391	0.07406	11
Total	843		1.82135		
Likelihood of Rural In-State Practice *(n = 174)*					
Effect	df	Mean Square	Variance Component	Standard Error	% Variance
Persons *(p)*	173	0.13270	0.02727	0.00357	51
Campus ©	1	0.01183	0.00000	0.00000	0
Raters : persons by campus *[r : (p x c, e)]*	348	0.02362	0.02362	0.00179	44
p x c	173	0.02864	0.00251	0.00177	5
Total	695		0.05340		
Campus © is a fixed facet; all others are random.					

Results from a corresponding D study for the [r : (p x c)] mixed model with campus fixed are shown in Figure 1. For overall applicant acceptability, one admissions interview conducted on each of the 2 campuses resulted in an observed reliability of 0.57 which increased to 0.73 when doubled to 2 interviews per campus. Ratings of RPLP applicants' likelihood of practicing medicine in rural Kentucky tended to be more reliable: The reliability of 2 interviews, one on the main campus and one the regional medical campus, was 0.70 and rose to 0.82 when increased to 2 interviews per campus (4 in total). For both measures, the effects on reliability of increasing the number of interviews beyond the present configuration of 2 were modest.

Figure 1. Univariate Mixed Model D Study Results: [r : (p x c)]



Multivariate G Study Results

Table 2 displays the multivariate G study results in a matrix format, with VCs for applicant acceptability ratings and likelihood of rural practice reported by campus on the diagonals of each of the 2 matrices (p, r:p). As shown in the left matrix, the proportion of variance attributable to systematic differences in applicants (p) is considerably greater for regional campus interviewers (53% of the total variance) than their main campus counterparts (30% of the total variance). This implies that the reliability of a single interview is 0.30 for the main campus and 0.53 for the regional campus. The average reliability of one interview across the 2 campuses is 0.42. The relationship between campuses, reflected in the observed covariance (0.56), is reported in the lower left cell of the p matrix. However, more readily interpretable is the universe score correlation (0.83) shown in the upper right cell – which, again, is synonymous with "true score" correlation in classical test theory. This universe score correlation indicates a strong positive relationship between interviewer ratings of overall applicant acceptability at main and regional campuses, suggesting that raters on the 2 campuses were assessing similar but not identical applicant characteristics.

Table 2. Multivariate G Study (r ○: p ●) with 2 Levels (Main and Regional Campus)

Effect	Variance Components Overall Applicant Acceptability *(n = 211)* [%]		Variance Components Likelihood of Rural In-State Practice *(n = 174)* [%]	
	Main	Regional	Main	Regional
p‡	0.41 [30%]	0.83*	273.79 [49%]	0.91*
	0.56†	1.09 [53%]	260.17†	296.69 [61%]
rpΔ	0.97 [70%]		282.64 [51%]	
		0.98 [47%]		189.81 [34%]

‡ Reflects "true score" variance attributable to person (applicant) – the object of measurement.
Δ Reflects error variance from all other sources/facets. Due to the nesting of raters within persons, includes any systematic error due to rater stringency.
** Universe score correlation / † Covariance*

In the right matrix, similar results are presented for RPLP applicants' likelihood of rural in-state practice. Compared to overall applicant acceptability, the proportion of "true score" variance attributable to applicants (p) is somewhat greater and less disparate for both main and regional campus interviewers (49% and 61% of the total variance, respectively). The universe score correlation (0.91), listed in the upper right cell of the top matrix, indicates that raters on the 2 campuses were assessing very similar aspects in arriving at their judgments regarding applicants' likelihood of practicing medicine in rural Kentucky. Since raters were nested within persons, it was not possible to disentangle the specific error attributable to the rater-person interaction and systematic rater stringency. Hence, in this study, the r : p VC reflected the sum of the interaction and systematic VCs.

Multivariate D Study Results

Figures 2 and 3 present multivariate D study results, by campus, for each of the measures collected in the RPLP admissions process: Overall applicant acceptability and likelihood of rural in-state practice, respectively. In addition, composite estimates are presented which combine information on interviewer ratings from both main and regional campuses. Estimated reliability is projected for up to 4 interviews per campus and, correspondingly, 8 total interviews per applicant.

In Figure 2, the G coefficients corresponding to our current protocol of 2 independent raters (one per interview) from each campus are 0.46 (main) and 0.69 (regional) which, respectively, would increase incrementally to 0.56/0.77 and 0.63/0.82 with an additional 2-3 interviews per campus. With main and regional campus ratings weighted equally, the composite reliability averaged across 4 interviews (the current configuration) is 0.73 and would increase to 0.84 if doubled to 8 total interviews (4 per campus).

Figure 2. Multivariate D Study Results: Overall RPLP Applicant Acceptability

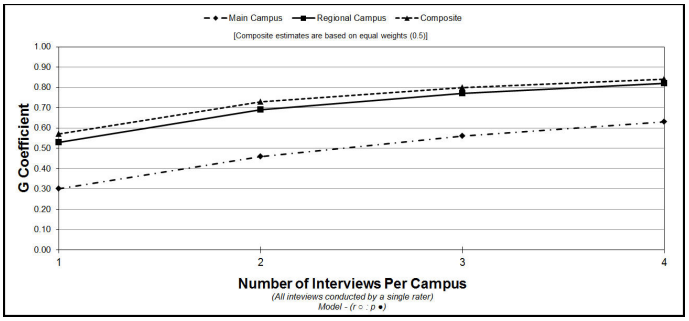
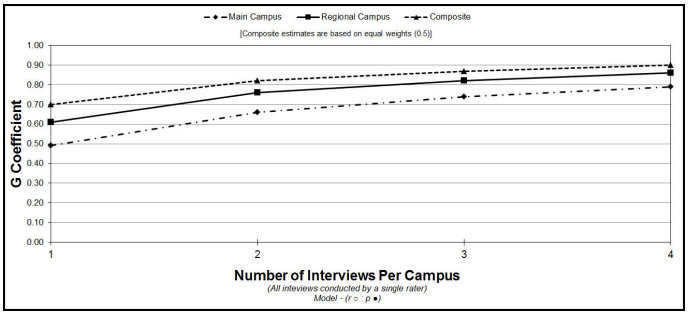


Figure 3 plots comparable generalizability estimates for applicants' likelihood of rural in-state practice. Here, ratings are both more reliable and more comparable, with G coefficients for 2, 3, and 4 interviews being 0.66, 0.74, and 0.79 for main campus interviewers and 0.76, 0.82, and 0.86 for their regional counterparts. Composite reliabilities, again weighted equally across campuses, are 0.82 (4 total interviews), 0.87 (6 total interviews), and 0.90 (8 total interviews) – good to excellent for most purposes.

Figure 3. Multivariate D Study Results: Likelihood of Rural In-State Practice



Discussion

As the growth of regional campuses continues, admission committees may seek to expand or better formalize the roles played by these partner programs. Especially where such programs have uniquely-targeted (e.g. rural) missions, the need may exist to access specific expertise or incorporate local stakeholders in selecting qualified candidates. Coordinating this process in a logistically and psychometrically optimal fashion requires a robust plan for establishing and monitoring this process.

This study examined admission interview ratings from a regional campus which, via a Rural Physician Leadership Program, provides clerkship training to 10 UK COM students interested in rural medical practice, preferably in rural Kentucky.[21] Established in 2008, this "clinical model" RMC[18] is an extension of the "parent" academic medical center located

about one hour away, which retains ultimate authority for all enrollment under a single, centralized admissions committee. However, primary feedback from RPLP stakeholders is routinely provided through a parallel process in which each applicant, over a 2-day period, meets independently with 2 interviewers at each campus.

This structure allowed reliability of interviewer assessments to be examined using a generalizability framework. Weighted equally and averaged across campuses, combined reliabilities of both overall RPLP applicant rating (0.73) and likelihood of rural practice (0.82) were adequate for the purposes at hand. However, main campus assessments tended to contain more error variance for both measures, particularly overall applicant acceptability. While effects of alternative weighting schemes on overall reliability were not formally examined, slight improvements appear possible by assigning a greater contribution to regional campus ratings.

Several possible explanations exist for the observed variation in main and regional campus reliabilities. First, as previously mentioned, the regional campus interviewer pool is smaller and more homogeneous in term of program knowledge and focus; that is, most interviewers are intimately familiar with the RPLP history, curriculum, and objectives. In contrast, main campus interviewers reflect a much broader array of backgrounds, expertise, and interests reflective of an academic medical center. Second, while the association between applicants' overall acceptability and likelihood of rural in-state practice varied somewhat by campus, it tended to be weak: Applicants' academic qualifications were largely unrelated to raters' judgments about their propensity to eventually practice medicine in rural Kentucky.

Due to confounding factors, the data collection design was unable to accommodate occasion and rater as separate random facets. However, the observed results are fairly consistent with other reports on admission interview reliability.[13,32,33] Further, these findings support the utility of having RPLP candidates interview at both the regional campus and main campus. In this study, the observed interview reliability effectively comprised 4 interviews per applicant, the likely reason for the reliable mean interview scores. In addition, results show little benefit to expanding the RPLP admission process beyond 2 interviews per campus.

While the medical school personal interview (MSPI) remains part of the admissions process, it is only one source of data considered in committee assessments of applicant qualifications and professional/program fit.[34] With rare exceptions,[35,36] how this and other information is used in committee deliberations or weighted in decisions to accept, reject, or hold applicants has not been widely examined.[13] Holistically, MSPI ratings are considered part of applicants' overall "dossier", but no algorithm or guidelines standardize their role or degree of influence. As such, it is unknown whether their use consistently constitutes a "high stakes" application, which has obvious implications for the level of rigor required in their collection and measurement. In a holistic review of applicants, a low rating of overall acceptability, based in part on face-to-face interviews, could prove detrimental to admission.

The study objective was not to establish the superiority of a single approach to assessing applicant fit, but rather to empirically explore the apparent tension between the diversity of stakeholder input and the reproducibility of resulting scores. Indeed, recent research suggests a hybrid model containing selected elements from various approaches might be optimal.[33,37] Predicting future events, be it academic performance, specialty choice, or eventual practice locale, remains an inexact science[22] involving both tangible[38] and intangible[39,40] considerations. From prior research on rural medical practice, considerable attention has been paid to applicants' related backgrounds, interests, and experiences.[41-43] Indeed, of the 107 (38.8%) RPLP applicants *not* invited for interviews during the study time frame, most lacked meaningful rural experience and/or sufficient academic performance.[23]

Since the overriding goal of the RPLP is to recruit and train physicians who will practice medicine in rural Kentucky, a shared understanding of program goals is essential. In the case of the RPLP, this was explicit – incorporating the major program outcome (practicing in rural Kentucky) into the actual interview process. The purpose was not to develop a precise measure, but rather to help direct focus on the task at hand. Whether this was effective or caused RPLP interviewers to cognitively approach the process differently is a question for future research. The meager correlation of this measure with overall acceptability suggests interviewers were able to discern between them.

These findings are limited by several factors. First, this study is based on a singular rural track training program at one US institution. As a result, how widely these findings may generalize beyond this context is unknown. Second, although all interviewers follow the same semi-structured format, there is some flexibility in the specific questions that can be asked. Moreover, regional campus interviewers used a slightly different interview narrative form. Lastly, interviewers on both campuses were guided by "rural qualities" gleaned inductively via a nominal group process – not a standardized, demographic definition of "rural". While by design, this more qualitative operationalization was likely implicated in interviewers' assessments of likelihood of rural practice.

Another issue worth mentioning is the potential disconnect between the composite reliability estimates averaged across campus and, in actual practice, the disaggregated use of interviewer ratings by the admissions committee. That is, the informal assessment of agreement made by committee members in their review and comparison of individual interviewer ratings,some of which, not being anonymous, may be afforded more credence than others.

Conclusion

Dedicated rural medical tracks or programs have been shown to be effective strategies in producing primary care physicians

for practice in rural, often underserved areas,[42] especially when provided in settings (like RMCs, for example) that offer meaningful learning experiences outside the larger, urban environment.[44] Key to the success of these efforts is the selection of candidates most qualified to meet programmatic goals. In Kentucky, the RPLP was designed to meet this need by admitting applicants who prefer rural practice and training them in settings with appropriate physician and community role models.

Study results found composite (combined) reliabilities of RPLP applicants' overall acceptability and likelihood of rural in-state practice to be encouraging. On both measures, however, ratings from regional campus interviewers tended to have less error variation than their main campus counterparts. It is possible that better training and calibration, perhaps combining interviewers from both campuses, might narrow the observed differences in reliability. Mean ratings, it should be noted, did not differ between campuses. Various weighting schemes could, in concert with the number of interviews, be more closely examined as a means of maximizing overall reliability.

This study highlights a methodology for developing and monitoring the inclusion of additional stakeholders to the admission process, and may prove useful for programs seeking to strategically tap a wider range of perspectives – especially as they relate to a specific, targeted mission. In the present context, eliciting input from interviewers with complimentary backgrounds resulted in more reliable composite ratings of applicants' acceptability and likelihood of practicing in rural Kentucky. In addition to estimating the reliabilities of these combined scores, the multivariate approach allowed estimation of the relationship between the 2 groups of interviewers.

While broadening the universe of generalization is typically associated with reduced reliability, in this application, the increase in reliability from additional interviews was found to outweigh these effects. Put another way, the addition of regional stakeholders to the pool of potential interviewers may not necessarily result in a less reliable composite measure. As regional campuses proliferate, the use of multivariate generalizability approaches to examine assessments of applicants or students at multiple "fixed" locales may hold promise.

Future studies should include a focus on the validity of interviewer ratings. That is, whether or not assessments validly reflect applicants' academic success or eventual likelihood of practicing in rural Kentucky, for example. Although early data appear encouraging, a continued follow-up of RPLP graduates will help determine the accuracy of these long-term projections made during admission into the profession. On a broader scale, studies may wish to explore underlying differences and similarities in determining the psychometric impact of expanding stakeholder input.

References

1. Freidson E. *Profession of medicine: a study in the sociology of applied knowledge*. Chicago, IL: University of Chicago Press, 1970/1988.
2. 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? *Academic Medicine*. 2013;88(5):672-681. DOI: 10.1097/ACM.0b013e31828bf252.
3. Lemay JF, Lockyer JM, Collin VT, Brownell AK. Assessment of non-cognitive traits through the admissions multiple mini-interview. *Medical Education*. 2007;41(6):573-579. DOI: 10.1111/j.1365-2923.2007.02767.x.
4. Siu E, Reiter HI. Overview: what's worked and what hasn't as a guide toward predictive admissions tool development. *Advances in Health Science Education*. 2009;14(5):759-775. DOI: 10.1007/s10459-009-9160-8.
5. Patterson F, Cousans F, Edwards H, Rosselli A, Nicholson S, Wright B. The predictive validity of a text-based situational judgment test in undergraduate medical and dental school admissions. *Academic Medicine*. 2017;92(9):1250-1253. DOI: 10.1097/ACM.0000000000001630.
6. Kreiter CD. A proposal for evaluating the validity of holistic-based admission processes. *Teaching and Learning in Medicine*. 2013;25(1):103-107. DOI: 10.1080/10401334.2012.741548.
7. Association of American Medical Colleges (AAMC). *Roadmap to diversity: integrating holistic review practices into medical school admission processes*. Washington, DC: AAMC, 2010.
8. Grabowski CJ. Impact of holistic review on student interview pool diversity. *Advances in Health Sciences Education*. 2018;23(3):487-498. DOI: 10.1007/s10459-017-9807-9.
9. Ellaway RH, Mahli R, Bajaj S, Walker I, Myhre D. A critical scoping review of the connections between social mission and medical school admissions: BEME guide no. 47. *Medical Teacher*. 2018;40(3):219-226. DOI: 10.1080/0142159X.2017.1406662.
10. Reiter HI, Eva KW. Reflecting the relative values of community, faculty and students in the admissions tools of medical school. *Teaching and Learning in Medicine.* 2005;17(1):4-8. DOI: 10.1207/s15328015tlm1701_2.
11. Kreiter C, O'Shea M, Bruen C, Murphy P, Pawlikowska T. A meta-analytic perspective on the valid use of subjective human judgment to make medical school admissions decisions. *Medical Education Online (Online)*. 2018;23(1): DOI: 10.1080/10872981.2018.1522225.
12. Pau A, Chen YS, Lee VKM, Sow CF, De Alwis R. What does the multiple mini-interview have to offer over

the panel interview? *Medical Education Online (Online)*. 2016;21:29874. DOI: 10.340/meo.v21.29874.
13. Kreiter CD, Axelson RD. A perspective on medical school admission research and practice over the past 25 years. *Teaching and Learning in Medicine*. 2013;25(sup1):S50-S56. DOI: 10.1080/10401334.2013.842910.
14. Sokal-Gutierrez K, Ivey SL, Garcia RM, Azzam A. Evaluation of the Program in Medical Education for the Urban Underserved (PRIME-US) at the UC Berkeley-UCSF Joint Medical Program (JMP): the first 4 years. *Teaching and Learning in Medicine*. 2015;27(2):189-196. DOI: 10.1080/10401334.2015.1011650.
15. Avery DM JR, Wheat JR, Leeper JD, McKnight JT, Ballard BG, Chen J. Admission factors predicting family medicine choice: a literature review and exploratory study among students in the Rural Medical Scholars Program. *Journal of Rural Health*. 2012;28(2):128-136. DOI: 10.1111/j.1748-0361.2011.00382.x.
16. Cathcart-Rake W, Robinson M, Paolo A. From infancy to adolescence: the Kansas University School of Medicine-Salina: a rural medical campus story. *Academic Medicine*. 2017;92(5):622-627. DOI: 10.1097/ACM.0000000000001455.
17. Harding CV, Akabas MH, Andersen OS. History and outcomes of 50 years of physician-scientist training in medical science training programs. *Academic Medicine*. 2017;92(10):1390-1398. DOI: 10.1097/ACM.0000000000001779.
18. Cheifetz CE, McOwen KS, Gagne P, Wong JL. Regional medical campuses: a new classification system. *Academic Medicine*. 2014;89(8):1140-1143. DOI: 10.1097/ACM.0000000000000295.
19. Bunton SA, Sabalis RF, Sabharwal RK, Candler C, Mallon WT. *Medical expansion: challenges and strategies*. Washington, DC: Association of American Medical Colleges (AAMC), 2008.
20. Association of American Medical Colleges (AAMC). *AAMC statement on the physician workforce*. Washington, DC: AAMC, 2006.
21. Arnett PK, Stratton TD, Weaver AD, Elam CL. University of Kentucky Rural Physician Leadership Program: a programmatic review. *Journal of Regional Medical Campuses (Online)*. 2018;1(3): DOI: 10.24926jrmc.v1i3262.
22. Benbassat J, Baumal R. Uncertainties in the selection of applicants for medical school. *Advances in Health Sciences Education*. 2007;12(4):509-521. DOI: 10.1007/s10459-007-9076-0.
23. Elam CL, Weaver AD, Whittler ET, Stratton TD, Asher LM, Scott KL, et al. Discerning applicants' interests in rural medicine: a textual analysis of admission essays. *Medical Education Online (Online)*. 2015;20:27081. DOI: 10.3402/meo.v20.27081.
24. Mason PB, Cossman JS. Does one school's admission policy help a rural state "grow their own" physicians? *Journal of the Mississippi State Medical Association*. 2012;53(9):284-286,288-292.
25. Cronbach LJ, Glaser GC, Nanda H, Rajaratnam N. *The dependability of behavioral measurements: theory of generalizability of scores and profiles*. New York, NY: John Wiley & Sons, 1972.
26. Brennan RL. *Elements of generalizability theory*. Iowa City, IA: American College Testing Program, 1983.
27. Crick GE, Brennan RL. *GENOVA: a generalized analysis of variance system*. Dorchester, MA: University of Massachusetts at Boston, 1982.
28. Shavelson RJ, Webb NM. *Generalizability theory: a primer*. Newbury Park, CA: SAGE, 1991.
29. Kreiter CD. Generalizability theory. In: SM Downing, R Yudkowsky (Eds); *Assessment in health professions education*. London: Routledge, 2009; 75-92.
30. Kreiter CD, Ferguson KJ. An investigation of the generalizability of medical school grades. *Teaching and Learning in Medicine*. 2016;28(3):279-285. DOI: 10.1080/10401334.2016.1154859.
31. Brennan RL. *Manual for mGENOVA: version 2.1*. Iowa City, IA: Iowa Testing Programs, 2001.
32. Axelson RD, Kreiter CD. Rater and occasion impacts on the reliability of pre-admission assessments. *Medical Education*. 2009;43(12):1198-1202. DOI: 10.1111/j.1365-2923.2009.03537.x.
33. Hanson MD, Kulasegaram KM, Woods NN, Fechtig L, Anderson G. Modified personal interviews: resurrecting reliable personal interviews for admission?*Academic Medicine*. 2012;87(10):1330-1334. DOI: 10.1097/ACM.0b013e318267630f.
34. O'Neill LD, Korsholm L, Wallstedt B, Eika B, Hartvigsen J. Generalizability of a composite student selection program. *Medical Education*. 2009;43(1):58-65. DOI: 10.1111/j.1365-2923.2008.03247.x.
35. Elam CL, Stratton TD, Wilson JF, Scott KL. How admission committees decide: influence of committee members' experience and applicants' academic characteristics. *Academic Medicine*. 2002;77(10 suppl**)**:S26-S28.
36. Elam CL, Johnson MMS. An analysis of admission committee voting patterns. *Academic Medicine*. 1997;72(10 suppl):S72-S75.
37. Bibler Zaidi NL, Santen SA, Purkiss JA, Teener CA, Gay SE. A hybrid interview model for medical school interviews: combining traditional and multisampling formats. *Academic Medicine*. 2016;91(11):1526-1529. DOI: 10.1097/ACM.0000000000001218.
38. Daniels ZM, VanLeit BJ, Skipper BJ, Sanders ML, Rhyne RL. Factors in recruiting and retaining

professionals for rural practice. *Journal of Rural Health*. 2007;23(1):62-71. DOI: 10.1111/j.1748-0361.2006.00069.x.

39. Cutchin MP. Community and self: concepts for rural physician integration and retention. *Social Science and Medicine*. 1997;44(11):1661-1674.
40. Hancock S, Steinbach A, Nesbitt TS, Adler SR, Auerswald CL. Why doctors choose small towns: a developmental model of rural physician recruitment and retention. *Social Science and Medicine*. 2009;69(9):1368-1376. DOI: 10.1016/j.socscimed.2009.08.002.
41. Greer T, Kost A, Evans DV, Norris T, Erickson J, McCarthy J, et al. The WWAMI Targeted Rural Underserved Track (TRUST) Program: an innovative response to rural physician workforce shortages. *Academic Medicine*. 2016;91(1):65-69. DOI: 10.1097/ACM.0000000000000807.
42. Florence JA, Goodrow B, Wachs J, Grover S, Olive KE. Rural health professions education at East Tennessee State University: survey of graduates from the first decade of the community partnership program. *Journal of Rural Health*. 2007;23(1):77-83. DOI: 10.1111/j.1748-0361.2006.00071.x.
43. Roseamelia C, Greenwald JL, Bush T, Pratte M, Wilcox J, Morley CP. A qualitative study of medical students in a rural track: views on eventual rural practice. *Family Medicine*. 2014;46(4):259-266.
44. Brokaw JJ, Mandzuk CA, Wade ME, Deal DW, Johnson MT, White GW, et al. The influence of regional basic science campuses on medical students' choice of specialty and practice location: a historical cohort study. *BMC Medical Education (Online)*. 2009;9:29. DOI: 10.1186/1472-6920-9-29.