Tracking Surgical Education Survey Research Through the APDS Listserv

Tiffany N. Anderson, MD,* Edmund W. Lee, MD,* Sylvia Bereknyei Merrell, DrPH, MS,*,† and James R Korndorffer, MD, FACS, MHPE*

*Department of Surgery, Stanford University School of Medicine, Stanford, California; and †Stanford-Surgery Policy Improvement Research and Education Center (S-SPIRE), Department of Surgery, Stanford University School of Medicine, Stanford, California

OBJECTIVE: Survey-based studies are cornerstones in medical education research. The Association of Program Directors in Surgery (APDS) listserv offers a method to contact program directors (PD) and residents for such research. To facilitate research beneficial to the APDS, improve the quality of survey-based research and minimize survey fatigue, the APDS research committee (ARC) developed a survey review process to grant access to the listserv for research. This study was conducted to determine the impact of the review process on the quality of survey-based research and eventual publication.

DESIGN: This log was systematically reviewed identifying publications resulting from accepted surveys. Publications were categorically analyzed to determine the components of their survey tool methodology, response rate (RR), and medical education research study quality instrument (MERSQI) score.

SETTING: The ARC used a 2-reviewer peer-review process for survey distribution requests. The request was either accepted, rejected, or returned for revision. Accepted surveys were distributed through the listserv with an ARC attestation of approval.

PARTICIPANTS: A log of all survey requests maintained from 2014 to 2017 and subsequent publications.

RESULTS: Thirty-five requests were accepted (40%), 30 were reviewed discovering 10 surveys that led to 12 publications (publication rate of 33%). The average RR was 60% (SD = 29%). Detailed explanations of survey development strategies were reported in 5 (42%), consisting of methods building validity evidence such as expert consensus, modified Delphi method, and pilot group sampling. Half of study participants were PD (50%). MERSQI scores averaged 10 (SD = 1.6).

CONCLUSION: Based on those survey research published to date, the ARC survey peer-review process has enabled most accepted surveys to achieve adequate RR. Although the pool of accepted requests is small, it does highlight areas of improvement. With further refinement of the process, including questioning the survey development methods, the process and listserv can be a powerful tool for further research. (J Surg Ed 76:e41−e48. © 2019 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

ABBREVIATIONS: APDS, Association of Program Directors in Surgery, APD, Associate Program Director, ARC, APDS Research Committee, EP, Educational Policy, IP, Institutional Policy, MERSQI, Medical Education Research Study Quality Instrument, PD, Program Director, PHP, Public Health Policy, RF, Resident or Fellow, RR, Response Rate

KEY WORDS: MERSQI, APDS Listserv, Response rate, Survey fatigue

COMPETENCIES: Medical Knowledge, Professionalism, Systems-Based Practice

INTRODUCTION

Since 1977 the Association of Program Directors in Surgery (APDS) has operated as a forum for the exchange of information, research, and discussion on a wide range of subjects related to postgraduate surgical education. With the advent of electronic methods for dissemination of information, the APDS developed a Listserv to foster group communication between program directors and their associated faculty and residents. This Listserv has been particularly amenable to the conduction of survey-based research.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Correspondence: Inquiries to Tiffany N. Anderson, MD, Department of Surgery, Stanford University School of Medicine, 300 Pasteur Drive, H3552, MC 5655, Stanford, CA 94305; e-mail: tnanders@stanford.edu

© 2019 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.
1931-7204/$30.00
https://doi.org/10.1016/j.jsurg.2019.07.006

Survey-based studies are an integral facet of medical education research.[1,2] Information is gathered by asking a subset of people questions on a specific topic and generalizing the results to a larger population.[2] It is the relative ease of this research methodology that has led it to be perhaps the most misused resource amongst academic researchers.[3] The nature of medical education research is such that samples are often chosen for convenience as opposed to a strictly scientific process of selection. The APDS Listserv, an ideal means through which survey participants might be recruited, became steadily more populated by such requests for research of questionable value. This prompted the APDS Research Committee (ARC) to implement a peer-review process in 2014 to vet survey-based researchers requesting access to the Listserv. It is currently unknown whether survey instruments submitted through the APDS Listserv have a similarly poor response rate or what their subsequent publication outcomes are.[4-6]

The ARC's goal is to facilitate research beneficial to the APDS, improve the quality of survey-based research, and minimize survey fatigue. This systematic review of survey distribution requests was undertaken to examine the degree to which the review process is achieving these goals. We hypothesize that surveys that underwent this evaluation process will have an improved response rate (compared to the typical medical education response rate) and will have improved outcomes in peer-reviewed publication.

MATERIAL AND METHODS

Setting

Survey distribution requests containing the survey instrument, request form, and Institutional Review Board (IRB) approval are submitted to the APDS webmaster and then forwarded to the APDS Research Committee (ARC) Listserv director responsible for the survey review process. Screened requests are given to 2 ARC members for independent blinded peer review to assess the quality of the instrument. These requests are then either accepted, rejected, or returned to the submitter for revisions. Upon acceptance, the survey instrument is distributed through the Listserv with a statement attesting to the instrument's quality ("approved by the APDS research committee and all members of the APDS are encouraged to support this survey") and follows the survey distribution protocol outlined by the researcher. The ARC has maintained a log of all survey distribution requests since May 2014.

Study Design

A systematic review was conducted between December of 2018 and March 2019 of all approved requests made between May 2014 and May 2017. The following digital resources were used for this review: PubMed and Google Scholars. Search terms for each item consisted of the lead author's full name and title of project (if known) or combinations targeting titles and/or author keywords. For those requests that contained insufficient information, attempts were made to directly contact the listed point of contact. The data extraction tools were developed by all authors. Because this article reviewed previously published studies and did not involve human participants, it did not qualify as human subjects research by the Stanford IRB (Fig. 1).

Survey Instrument Data Abstraction Tool

The survey data extraction form for survey instruments characterized instruments by year of request, number of survey items, region of requesting institutions (United States census regions), study participant(s), study subjects, study topic category.

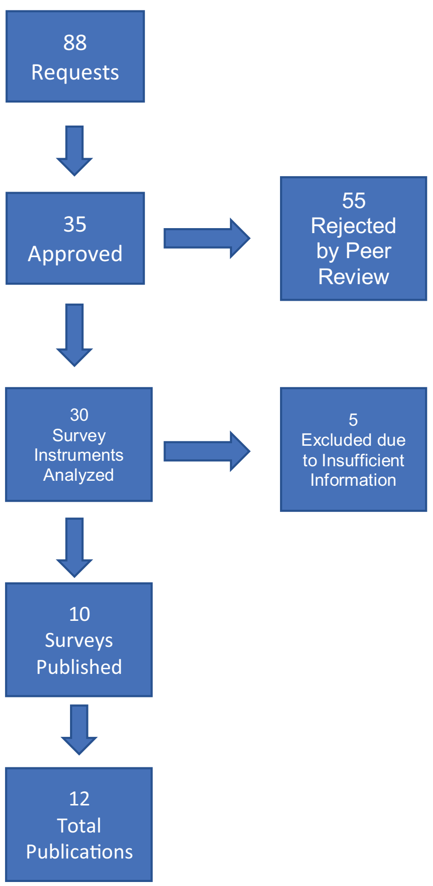

FIGURE 1. Selection strategy for APDS Listserv requests for distribution.

Publication Data Abstraction Tool

We reviewed subsequent publications using the Medical Education Research Study Quality Instrument (MERSQI).[7] The MERSQI contains 6 domains (max for each domain 3, for a max total of 18) of study quality: study design, sampling, data type, validity of assessments, data analysis, and outcomes, and has been found to have a modest overall Cronbach alpha of all items and acceptable inter-rater reliability respondent range.[7] See Appendix for full data abstraction tool.

In addition to the MERSQI, we further characterized publications by year of publication, response rate (as defined as "the percentage of participants who completed the evaluation component of the study),"[7] use of conceptual framework(s), report of development process of the survey instrument, survey duration, and the journal impact factor at the time of publication (InCites Journal Citation Reports by Clarivate).

Data Analysis

Publications and associated survey instruments were reviewed, classified, and scored by 2 reviewers (TNA and JRK) using the data abstraction tools. Only full text articles and complete survey instruments were used in the analysis. Abstracts, oral or poster presentations were not included. To assess the reliability of scoring and classification, all publications were examined by a second blinded reviewer (JRK) using the publication data abstraction tool. The data were examined with descriptive statistic.

Publications and associated survey instruments were reviewed, classified, and scored by 2 reviewers (TNA and JRK) using the data abstraction tools. Only full text articles and complete survey instruments were used in the analysis. Abstracts, oral or poster presentations were not included. To assess the reliability of scoring and classification, all publications were examined by a second blinded reviewer (JRK) using the publication data abstraction tool. The data were examined with descriptive statistics.

RESULTS

Description of the Sample

A total of 88 requests for Listserv approval were made between May 2014 and May 2017; 35 surveys were approved for distribution. Ten of the approved instruments were linked to a possible 14 publications; however, 2 publications were discarded after contacting the authors for further clarification, leading to a project publication rate of 33%. Discrepancies between reviewers

TABLE 1. Characteristics of Survey Requests and Publications

Characteristics	N (%)	Mean (SD)	Range
All requests (n = 30)			
Requests	88		
Accepted	35 (40)		
Evaluated	30		
Published	10		
Publications	12		
Year of request			
2014	3 (10)		
2015	12 (40)		
2016	8 (27)		
2017	7 (23)		
Region of institutions			
Midwest	9 (30)		
Northeast	13 (43)		
South	7 (23)		
West	8 (27)		
Survey participants			
Program directors/ APD	17 (57)		
Faculty	3 (10)		
Fellows/Residents	14 (47)		
Students	1 (3)		
Study subjects			
Program directors/ APD	5 (17)		
Faculty	1 (3)		
Fellows/Residents	16 (53)		
Institution/Program	13 (43)		
Students	1 (3)		
Topic category			
Wellness	6 (20)		
Program evaluation	3 (10)		
Curriculum development	1 (3)		
Public health policy	4 (13)		
Education policy	16 (53)		
Faculty development	1 (3)		
Publications (n = 12)			
Request year			
2014	2 (16)		
2015	5 (42)		
2016	5 (42)		
Publication year			
2015	1 (8)		
2016	2 (17)		
2017	2 (17)		
2018	7 (58)		
Study participant			
Program directors/ APD	6 (50)		
Faculty	1 (8)		
Resident/Fellows	4 (13)		
Students	1 (8)		
Study subject			
Program directors/ APD	3 (25)		
Faculty	1 (8)		
Residents	5 (42)		

(continued)

TABLE 1 (*continued*)

Characteristics	N (%)	Mean (SD)	Range
Institution/Program	4 (33)		
Topic category			
Program directors/ APD	5 (42)		
Students	3 (25)		
Institutional/Program	3 (25)		
Faculty	1 (8)		
Reported response rate	10 (83)	60% (29%)	
Journal impact factor (at time of publication)	12	3.76 (4.49)	
Survey items	11 (92)	40 (27)	13-74
Study duration (in days)	7 (58)	52 (42)	28-133

were discussed until 100% consensus was achieved (Table 1).

Population Characteristics

All of the distributed surveys were part of cross-sectional study designs. Accepted requests for the 3-year period were fairly evenly distributed after May 2015, 30% in 2015, 27% in 2016, and 23% in 2017. Survey requests by region were also relatively equally distributed across the country with 24% from midwest, 35% from the northeast, 19% from the south and 22% from the west. Most survey participants were program directors (PD) or associate program directors (APD) (47%) or residents and fellows (RF) (39%). The study subjects were most commonly RF (44%), then institution or program (IP) (31%), and PD/APD (14%). Study topic categories most commonly pertained to educational policy (EP) (50%) and wellness (19%). The number of survey instrument items averaged 29 (SD = 20, range 9-74) (Table 1).

Publications

Over half of the publications were published in 2018 (58%). The years with the greatest number of survey requests and distributions were 2015 and 2016 (42% each). Half of the participants were PD/APD (50%), followed by RF (34%). Study subjects were most commonly RF (38%), IP (31%), and PD/APD (23%). Study topic categories were Wellness (42%), public health (25%), education policy (25%), and faculty development (8%).

The average MERSQI score was 10 (SD = 1.62, range 8-12). The weakest domain scores were in study design (mean 1, SD = 0), type of data (mean 1.16, SD = .58), validity of evidence of evaluation instrument (mean 1.33, SD = .78), and outcomes (mean 1.04, SD = .14). However, one study, Winkelman et al. (2015), received full marks and therefore succeeded in providing objective data (Table 2 and Table 3).

TABLE 2. Medical Education Research Study Quality Instrument (MERSQI) Scores

MERSQI Item	Mean (SD)
Study design (M1)	1 (0)
Sampling: Institution (M2)	2.08 (.19)
Response rate (M3)	
Type of data (M4)*	1.16 (0.58)
Validity evidence for evaluation instrument scores (M5)	1.33 (.78)
Data analysis: Sophistication (M6)	1.83 (.39)
Data analysis: Appropriateness (m7)	1 (0)
Outcome (M8)	1.04 (.14)
MERSQI score	10 (1.61)

The average reported response rate was 60% (SD = 29%, range 26%-100%). For 3 publications[9-11] the same survey tool was used, and the authors were able to achieve their targeted survey response as calculated by power analysis. Three (25%) publications reported the use of a conceptual framework to guide their project.[20,21,26] The average journal impact factor (at the time of publication) was 3.76 (SD = 4.49, range 2.051-8.498), over 6 unique journals. Detailed explanations of survey development strategies were reported in 6 (50%), consisting of methods building validity evidence such as expert consensus, modified Delphi method, and pilot group sampling. Seven (58%) publications reported survey duration which averaged 52 days (SD = 42, range 28-133). No internal or external funding sources were reported for any of the publications (Table 1).

DISCUSSION

This study demonstrates that the quality of the publications resulting from surveys distributed through the APDS Listserv compares favorably with published quality reviews of other medical education research studies across specialties.[12-16] That being said, these results continue to support the attestation that research in medical education remains poor to marginal at best. The lower threshold for good quality research is defined as a MERSQI of 12.[7] Rangel et al. was able to achieve this with one well-made tool for 3 publications.[9-11]

Although there is no consensus on what percentage constitutes an adequate response rate, Babbie et al. (1990) suggest that a response rate of at least 50% is adequate. Based on the survey research published to date, the APDS[17] research committee peer-review process has enabled most accepted surveys to achieve an adequate response rate, perhaps by limiting survey fatigue and

TABLE 3. Summary of Publications

Source	Response Rate	Survey Participants	Study Subjects	Topic Category	MERSQI Score	Journal Impact Factor
Winkelman et al. (2015)[8]	51.7	Medical students	IP	PH	11.5	3.494
Sandler et al. (2016)[19]	26	PD/APD	PD/APD, IP	Wellness	10	4.307
Langenfeld et al. (2016)[20]	42.5	PD/APD	PD/APD	EP	8	2.163
Damewood et al. (2017)[21]	54.7	PD/APD	IP	EP	10.5	2.3
Yorkgitis et al. (2018a)[22]	44.5	PD/APD	PD/APD	PHP	9	2.051
Yorkgitis et al. (2018b)[23]	44.5	PD/APD	RF	PHP	9	2.3
Kimbrough et al. (2018)[24]	NR	PD/APD	IP	EP	8	2.3
Rangel et al. (2018a)[9]	100	RF	RF	Wellness	12	8.498
Rangel et al. (2018b)[10]	100	RF	RF	Wellness	12	8.498
Rangel et al. (2018c)[11]	100	RF	RF	Wellness	12	2.141
Deal et al. (2018)[25]	35	Faculty	Faculty	Program Evaluation	8	2.3
Lebares et al. (2018)[26]	NR	RF	RF	Wellness	10	4.767

APD, Associate Program Director; EP, Educational policy; IP, Institutional policy or program; PHP, public health policy; NR, Not reported; PD, Program Director; RF, Resident or Fellow.

ensuring that the research question is of importance to the membership.

Although the pool of accepted requests is small, it does highlight an interesting pattern of publication. The publication rate thus far (33%) is consistent with publication of medical education as described across specialties (rates of 22% to 46%) in the past 20 years.[18] Historically, multi-institutional authorship, hypothesis-driven research, and completed studies were found to be predictive of publication.[14,16] Our exploration suggests a similar phenomenon.

The limitations of this study are the small sample size of surveys, which may have limited our ability to identify differences in MERSQI scores between publication and the narrow time period. Also, although we attempted to compensate the limitation of the MERSQI in its ability to assess quality of reporting by incorporating best practice recommendations for survey research reporting, caution should be taken in extrapolating beyond the current sample.

The small sample size may have limited our ability to identify differences in MERSQI scores between publication and the narrow time period. This time period does not take into account those manuscripts that are presently undergoing consideration for publication. Also, though we attempted to compensate for MERSQI's limitations in its ability to assess reporting quality (by incorporating best-practice recommendations for survey research reporting), caution should be taken when extrapolating findings.

Our results highlight opportunities for improving the quality of surgical education research. With further refinement of the approval process—e.g., vetting the survey development methods—the Listserv could prove a powerful tool for further research.

AUTHOR CONTRIBUTIONS

All authors drafted or critically revised the manuscript. Additional authorship contributions include:

- Study concept and design: Anderson TN, Korndorffer JR.
- Acquisition of data: Anderson TN, Korndorffer JR MD.
- Analysis and interpretation: Anderson TN, Korndorffer JR, Bereknyei Merrell B, Lee EW.
- Study supervision: Korndorffer JR.

REFERENCES

1. Artino AR Jr, La Rochelle JS, Dezee KJ, Gehlbach H. Developing questionnaires for educational research: AMEE Guide No. 87. *Med Teach*. 2014;36:463–474.

2. Bennett C, Khangura S, Brehaut JC, et al. Reporting guidelines for survey research: an analysis of published guidance and reporting practices. *PLoS Med*. 2010;8:e1001069.

3. Draugalis JR, Coons SJ, Plaza CM. Best practices for survey research reports: a synopsis for authors and reviewers. *Am J Pharm Educ*. 2008;72:11.

4. Cook JV, Dickinson HO, Eccles MP. Response rates in postal surveys of healthcare professionals between 1996 and 2005: an observational study. *BMC Health Serv Res*. 2009;9:160.

5. Kellerman SE, Herold J. Physician response to surveys: a review of the literature. *Am J Prev Med*. 2001;20:61–67.

6. Phillips AW, Reddy S, Durning SJ. Improving response rates and evaluating nonresponse bias in surveys: AMEE Guide No. 102. *Med Teach*. 2016;38: 217–228.

7. Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. *JAMA*. 2007;298:1002–1009.

8. Winkelman TNA, Lehmann LS, Vidwan NK, et al. Medical Students' Views and Knowledge of the Affordable Care Act: a survey of eight U.S. medical schools. *J Gen Intern Med*. 2015;30:1018–1024.

9. Rangel EL, Lyu H, Haider AH, Castillo-Angeles M, Doherty GM, Smink DS. Factors associated with residency and career dissatisfaction in childbearing surgical residents. *JAMA Surg [Internet]*. 2018. Available from: https://doi.org/10.1001/jamasurg.2018.2571.

10. Rangel EL, Smink DS, Castillo-Angeles M, et al. Pregnancy and motherhood during surgical training. *JAMA Surg*. 2018;153:644–652.

11. Rangel EL, Castillo-Angeles M, Changala M, Haider AH, Doherty GM, Smink DS. Perspectives of pregnancy and motherhood among general surgery residents: a qualitative analysis. *Am J Surg [Internet]*. 2018. Available from: https://doi.org/10.1016/j.amjsurg.2018.07.036.

12. Reed DA, Beckman TJ, Wright SM. An assessment of the methodologic quality of medical education research studies published in The American Journal of Surgery. *Am J Surg*. 2009;198:442–444.

13. Smith RP, Learman LA. A plea for MERSQI: the medical education research study quality instrument. *Obstet Gynecol*. 2017;130:686–690.

14. Walsh CM, Fung M, Ginsburg S. Publication of results of abstracts presented at medical education conferences. *JAMA*. 2013;310:2307–2309.

15. Papp KK, Baker EA, Dyrbye LN, et al. Analysis and publication rates of Clerkship Directors in Internal Medicine (CDIM) annual meeting abstracts 1995-2005. *Teach Learn Med*. 2011;23:342–346.

16. Smith S, Kind T, Beck G, et al. Further dissemination of medical education projects after presentation at a Pediatric National Meeting (1998−2008). *Teach Learn Med*. 2014;26:3–8.

17. Babbie E. Survey Research Methods. 2nd ed Calif: Wadsworth: Belmont; 1990.

18. Sawatsky AP, Beckman TJ, Edakkanambeth Varayil J, Mandrekar JN, Reed DA, Wang AT. Association between study quality and publication rates of medical education abstracts presented at the Society of General Internal Medicine Annual Meeting. *J Gen Intern Med*. 2015;30:1172–1177.

19. Sandler BJ, Tackett JJ, Longo WE, Yoo PS. Pregnancy and Parenthood among Surgery Residents: Results of the First Nationwide Survey of General Surgery Residency Program Directors. *J Am Coll Surg*. 2016;222:1090–1096.

20. Langenfeld SJ, Vargo DJ, Schenarts PJ. Balancing Privacy and Professionalism: A Survey of General Surgery Program Directors on Social Media and Surgical Education. *J Surg Educ*. 2016;73:e28–e32.

21. Damewood RB, Blair PG, Park YS, Lupi LK, Newman RW, Sachdeva AK. "Taking Training to the Next Level": The American College of Surgeons Committee on Residency Training Survey. *J Surg Educ*. 2017;74:e95–105.

22. Yorkgitis BK, Raygor D, Bryant E, Brat G, Smink DS, Crandall M. Surgery program directors' knowledge of opioid prescribing regulations: a survey study. *J Surg Res*. 2018;227:194–197.

23. Yorkgitis BK, Bryant E, Raygor D, Brat G, Smink DS, Crandall M. Opioid Prescribing Education in Surgical Residencies: A Program Director Survey. *J Surg Educ*. 2018;75:552–556.

24. Kimbrough MK, Thrush CR, Smeds MR, Cobos RJ, Harris TJ, Bentley FR. National Landscape of General Surgery Mock Oral Examination Practices: Survey of Residency Program Directors. *J Surg Educ*. 2018. Available from: https://doi.org/10.1016/j.jsurg.2018.07.012.

25. Deal SB, Alseidi AA, Chipman JG, et al. Identifying priorities for faculty development in general surgery using the Delphi consensus method. *J Surg Educ [Internet]*. 2018. Available from: https://doi.org/10.1016/j.jsurg.2018.05.003.

26. Lebares CC, Guvva EV, Ascher NL, O'Sullivan PS, Harris HW, Epel ES. Burnout and stress among US surgery residents: psychological distress and resilience. *J Am Coll Surg*. 2018;226:80–90.

APPENDIX

Full data abstraction tool

Domain: Item	Response options: Scores	Operational definitions
Study design (M1)*	• Single-group cross-sectional or single-group posttest only: 1 • Single-group pretest and posttest: 1.5 • Nonrandomized, 2 group: 2 Randomized controlled trial: 3	• Survey studies are cross-sectional. • Case−control and cohort studies (2 or more defined cohorts) are considered 2-group nonrandomized.
Study design Type	Specify study design	Crosssectional, pre-test/post-test, randomized, etc.
Sampling: Institution (M2)*	• 1 institution: 0.5 • 2 institutions: 1 • 3 or more institutions: 1.5	Number of institutions refers to origin of study participants (not study authors).
Region of institution**	• Midwest: 1 • Northeast: 2 • South: 3 • West: 4	• Midwest: ND, SD, NE, KS, MN, IA, MO, WI, IL, MI, IN, OH • Northeast: ME, VT, NY, MA, PA, CT, NJ, RI, NH • South: FL, GA, LA, TX, MS, AL, OK, AR, TN, KY, WV, VA, NC, SC, MD, DE • West: WA, OR, CA, ID, NV, AZ, MT, WY, CO, NM, AK, HI, UT • Code the region for each participating institution • Code each region only once
Authors' background	List the name of the institutions that the authors are affiliated.	List a given institution only once
Sampling: Response rate (M3)*	• Not applicable • <50% or not reported: 0.5 • 50%-74%: 1 • ≥75%: 1.5	• Response rate is the proportion of those eligible who completed the posttest or survey. For intervention studies, this is the proportion of those enrolled who completed the intervention evaluation. • Use "not applicable" only if a response rate truly does not apply (e.g., data obtained from a medical record or professional organization database). • If power analysis is used, then response rate calculated.
Response rate	Reported value	If not reported: 0
Type of data (M4)*	• Assessment by study participant: 1 • Objective: 3	Observer ratings are considered objective.
Study participant	• Program Directors/Assistant Program Director: 1 • Faculty: 2 • Fellows/Residents: 3 • Students: 4 • Others: 5	• Who completed the survey? • Faculty includes both administrative and non-administrative faculty • Code all groups responsible for completion of the survey • Specify who the other group
Study subject	• Program Directors/Assistant Program Director: 1 • Faculty: 2 • Fellows/Residents: 3 • Students: 4 • Institution/Program: 5 • Others: 6	• Who is the intended subject of the study? • Faculty includes both administrative and non-administrative faculty • Code all intended subjects of the survey • Specify who the other group
Validity evidence for evaluation instrument scores (M5)*	• Not applicable • Content: 1 • Internal structure: 1 • Relationships to other variables: 1	• Relevant content evidence would include using theory, guidelines, experts, and existing instruments to identify or refine the instrument. • Relevant internal structure evidence would include all reliability (internal consistency, interrater, interstation, and test−retest) and factor analysis. • Relevant evidence of relationships to other variables would include expert−novice comparisons and concurrent or predictive correlation with other variables. • Use "not applicable" only if the study does not measure a psychological construct *and* there is no instrument to rate (e.g., gender as the sole outcome); should be used very rarely.
Data analysis: Sophistication (M6)*	• Descriptive analysis only: 1 • Beyond descriptive analysis: 2	• Descriptive analyses include frequency, mean, and median. • Any test of statistical inference is considered "beyond descriptive."

(*continued*)

(*continued*)

Full data abstraction tool		
Domain: Item	**Response options: Scores**	**Operational definitions**
Data analysis: Appropriateness (m7)*	Data analysis appropriate for study design and type of data: 1	• Considered "no" if there is a statistical error or if authors failed to analyze data at all. • If no: 0
Outcome (M8)*	• Satisfaction, attitudes, perceptions, opinions, general facts: 1 • Knowledge, skills: 1.5 • Behaviors: 2 • Patient/health care outcome: 3	• General facts include participant demographics. • Knowledge/skills are in a test setting (paper, computer, simulation, or patients in a nonauthentic setting). • Behaviors are physician actions with real patients in a clinical context, or other activities in a real context. • Patient/health care outcomes are actual effects on real patients, programs, or society.
MERSQI score	The sum of M1-M8	Maximum total 18
Journal impact Factor at time of publication	Impact value	As per the journal citation report for the given year
Conceptual framework of the study	Specify framework if reported	• All named frameworks • If none: 0
Publication year	Year of publication	Not the year the survey was administered
Mode of distribution	• APDS Listserv: 1 • Mail: 2 • Social Media: 3 • Other: 4	Code all manner of distribution.
Funding	• No: 0 • Yes: 1	• External or Internal funding to conduct the study • If not specified: 0
Topic category	• Wellness: 1 • Program Evaluation:2 • Curriculum development:3 • Public health policy: 4 • Education policy: 5 • Faculty development: 6	• Overall theme of the project • Code all that apply

*MERSQI item scores are weighted to reflect features of higher study quality.
**Based off of US Census Regions.