

The Best Graduate Medical Education Articles From 2021—in Our (Humble) Opinions

Gail M. Sullivan, MD, MPH
Deborah Simpson, PhD
Anthony R. Artino Jr, PhD
Lalena M. Yarris, MD, MCR
Nicole M. Deiorio, MD

This past year was another banner year for graduate medical education (GME) publications. With the switch to a virtual format, our annual editors' best medical education papers session at the 2022 Accreditation Council for Graduate Medical Education Annual Educational Conference has shifted from a live to a published format. Here we present the best non-*Journal of Graduate Medical Education* (JGME) 2021 papers relevant to GME (BOXES 1–3). As usual we employ a highly “scientific” approach considering quality and practical relevance, which boils down to: *Do I like this paper?* In this editorial we present arguments for why *our* paper is indisputably the best in 2021. You can join the conversation by presenting your argument for the best paper of 2021 on Twitter and tagging @JournalofGME.

Tony Artino's Pick

The Ethics of Health Professions Education Research: Protecting the Integrity of Science, Research Subjects, and Authorship¹

Have you ever, in the context of your own educational or clinical research, done something that some might consider unethical or which lies in a “gray area”? For example, have you added an author to a paper who did not really deserve authorship—so-called honorary authorship? I have, along with many other health professions education (HPE) scholars.² In this invited commentary, Olle ten Cate discusses the ethics of HPE research based on his own experiences, published guidelines, and discussions with international colleagues.¹ He offers practical advice for mentors and emerging HPE scholars.

ten Cate categorizes research ethics into 3 areas: protecting the integrity of *science*, *research subjects*, and *authorship*. Regarding science, he notes that outright fraud is relatively infrequent in HPE, whereas questionable practices and “sloppy science”

are much more prevalent. He suggests that the risk of sloppy science has increased over the past decade due to factors such as the rapid growth of publications, publication pressures, and predatory journals with low quality standards. In protecting research participants, ten Cate notes that “the avoidance of harm” is a common rule across all scientific fields. He discusses the importance of ethical review committees, especially in research that involves patients or trainees. In the last area, authorship, ten Cate notes that, although authorship practices do not affect the integrity of science or research participants, they do affect “... the recognition and esteem researchers deserve for their creative productivity.” Problematic authorship practices include plagiarism, scooping ideas from other scholars after confidential review, adding authors who did not qualify for authorship, and excluding authors who did qualify for authorship.

Throughout his commentary ten Cate suggests a number of important lessons for mentors and emerging scholars. His lessons include better training with more discussions about how to conduct HPE research with integrity, use of reporting guidelines, prioritizing participant well-being over any individual study goal, scrupulous citation practices, and taking personal responsibility to never accept honorary authorship.

My take-home from this thoughtful commentary: Although ethical rules and regulations can guide research behavior, ethical conduct cannot be strictly enforced. Instead, it is up to individual researchers, editors, peer reviewers, supervisors, funders, and institutions to lead by example and work to create healthy research *cultures* that nurture ethical research *practices*. If we consider the COVID-19 pandemic and negative effects of scientific distrust and disinformation upon health, it becomes clear that we must strive to conduct HPE research with integrity to maintain the credibility of HPE science. Read this article by ten Cate for a clear, concise look at research

DOI: <http://dx.doi.org/10.4300/JGME-D-21-01209.1>

BOX 1 Best Non-JGME Papers of 2021

- Ilgen JS, de Bruin ABH, Teunissen PW, et al. Supported independence: the role of supervision to help trainees manage uncertainty. *Acad Med.* 2021;86(suppl 11):81–86. doi:10.1097/ACM.0000000000004308
- Janssen Stein J, Langerhuizen DWG, Kerkhoffs GMMJ, Ring D. Payments by industry to residency program directors in the United States: a cross-sectional study [published online ahead of print May 18, 2021]. *Acad Med.* 10.1097/ACM.0000000000004166
- Kelleher M, Kinnear B, Sall DR, et al. Warnings in early narrative assessment that might predict performance in residency: signal from an internal medicine residency program. *Perspect Med Educ.* 2021;10(6):334–340. doi:10.1007/s40037-021-00681-w
- Liang KE, Dawson JQ, Stolan MD, et al. A carbon footprint study of the Canadian medical residency interview tour. *Med Teach.* 2021;43(11):1302–1308. doi:10.1080/0142159X.2021.1944612
- ten Cate O. The ethics of health professions education research: protecting the integrity of science, research subjects, and authorship [published online ahead of print September 14, 2021]. *Acad Med.* doi:10.1097/ACM.0000000000004413

ethics in HPE, with practical lessons for mentors and early-career scientists.

Nicole Deiorio's Pick**Payments by Industry to Residency Program Directors in the United States: A Cross-Sectional Study³**

As noted in Tony's paper, temptations that pose threats to ethical conduct abound in academic medicine. In my selected paper, the authors queried 2 open-source databases to link program directors listed in Dximity to the Centers for Medicare & Medicaid Services record of personal payments from industry. A shocking 74% of program directors received personal payments from industry. These may have been for reasons related to research,

BOX 2 Runners-Up: Best Non-JGME Papers of 2021

- Gingerich A, Sebok-Syer SS, Lingard L, Watling CJ. The shift from disbelieving underperformance to recognizing failure: a tipping point model [published online ahead of print October 20, 2021]. *Med Educ.* doi:10.1111/medu.14681
- Hope D, Dewar A, Hay C. Is there a replication crisis in medical education research? *Acad Med.* 2021;96(7):958–963. doi:10.1097/ACM.0000000000004063
- LaDonna KA, Cowley L, Touchie C, LeBlanc VR, Spilg EG. Wrestling with the invincibility myth: exploring physicians' resistance to wellness and resilience-building interventions [published online ahead of print August 10, 2021]. *Acad Med.* doi:10.1097/ACM.0000000000004354

BOX 3 Group Pick: Fun Paper of 2021!

Watling C, Ginsburg S, Lingard L. Don't be reviewer 2! Reflections on writing effective peer review comments. *Perspect Med Educ.* 2021;10(5):299–303. doi:10.1007/s40037-021-00670-z

education, honoraria, entertainment, food, lodging, gifts, or other activities.

While journal editors are not immune from this professional threat,⁴ program directors are in a unique leadership position such that this practice must be examined with a particularly rigorous lens. Program directors are often role models for their residents and responsible for developing their residents' ethical and professional identities. Thus, their practice patterns around payments may have a strong influence on future generations of physicians. Though program directors may believe themselves able to compartmentalize these roles, we know that physicians' behaviors are influenced by gifts and this influence may not be recognized by the physicians themselves.⁵

My take-home from this eye-opening paper: Although not every payment from industry means that a program director's behavior is compromised, a startling number of program directors have potential conflicts of interest that could lead, even subconsciously, to changes in their behaviors. Shining a spotlight on this practice will hopefully advance this discussion in the GME community as it self-regulates around this potential threat.

Deb Simpson's Pick**Warnings in Early Narrative Assessment That Might Predict Performance in Residency: Signal From an Internal Medicine Residency Program⁶**

In 2020 the senior JGME editors, informed by individual interviews and thought leaders, envisioned GME in 2030: Technology and big data were 2 drivers that shaped their perspectives.⁷ In that context, this paper by Kelleher et al using learning analytics results from rating scale data is particularly exciting.⁶ The authors analyzed narrative assessments to identify at-risk residents early in training. While I was hoping the authors would be able to use natural language processing (NLP) for their narrative data sets analysis, their rigorous qualitative analysis moves the needle toward 2030 and affordable NLP.

Have you read narrative assessments comments for first-year residents such as "Resident needs to be more thorough when obtaining a history from the patient" or "Has an inability to recognize when patients aren't

TABLE

Explicit and Implicit Themes From Early Narrative Assessments Associated With Lower Entrusted Internal Medicine Residents⁶

Explicit Themes	Implicit Themes
1. Gaps in residents' attention to details	1. Focus on deficiency rather than growth/opportunity to improve
2. Communication deficiencies with patients	2. Normative comparisons identified residents as behind their peers
3. Difficulty recognizing the "big picture" in a patient's care	3. Warning (via words like mistake, inaccurate, errors, or feelings including worry, concern)

understanding what is being said." Do words such as *worry, errors, or concern* raise alarm? Kelleher et al analyzed quantitative workplace-based entrustment assessment scores for 89 residents over 36 months of residency (3 consecutive cohorts of categorical internal medicine residents) to identify 2 groups: residents who had lower entrustment ratings (N=26) and residents with typical entrustment ratings (N=13 randomly selected from all other residents).⁶ Then they pulled first-year work-based narrative assessments, from July to December, for these 39 residents. The authors performed an inductive thematic analysis of the data using a 2-phase approach cogently described in the paper. While many similarities were identified, the authors focused on identifying explicit and implicit comments (ie, the hidden codes) associated with lower entrusted residents. Six themes were identified, 3 explicit and 3 implicit (TABLE).

My 4 take-homes from this big data study: (1) Narrative assessments reveal real differences that are associated with lower entrusted residents. Faculty advisors, program directors, and Clinical Competency Committees should go beyond the "scores" and carefully review these data sets. Consider asking assessors who have warning-level concerns to reach out to program leaders so data do not get lost; (2) Residency and fellowship programs—even small programs—can collaborate among similar programs to use big data sets for predictive analyses; (3) Rigorous methodology is doable as evidenced by this study getting all "5's" from KeyLIME podcast panelists for both methods and overall value⁸; and (4) 2030 is arriving sooner than we thought!

Runner Up: Wrestling With the Invincibility Myth: Exploring Physicians' Resistance to Wellness and Resilience-Building Interventions⁹

Medical educators and health care systems are struggling with how to address trainee and physician well-being. Most of us have tried numerous interventions at the individual, program, and sponsoring institution levels to address this issue, with mixed

results. Say the word "well-being" and it may be received with eye rolls and sighs. Enter LaDonna et al who conducted semi-structured interviews with 22 Canadian faculty physicians regarding their well-being and burnout experiences, their perceptions of wellness initiatives, and how these influence their uptake of strategies to address well-being.⁹

Their findings: faculty participants' emotions on this topic ranged from frustrated and irritated, to angry. Faculty felt that medicine as a profession was moving away from "the values of caring and compassion that initially inspired them to become physicians"⁹ to valuing those who continuously and efficiently strive to meet performance targets—a "faster and faster treadmill" as one of my physician colleagues noted after reading the article. Combining this production mentality with the medicine invincibility myth produces faculty participants' perceptions that medical profession and health care systems approaches to well-being are "lip service" and "band-aid" solutions. The authors conclude that, "rather than extinguishing burnout...even if well intentioned, resilience rhetoric and wellness interventions may inadvertently accelerate it..."⁹ Their solution? Start with a long-overdue needs assessment. Read this excellent article and share with all who plan wellness initiatives.

Gail Sullivan's Pick

A Carbon Footprint Study of the Canadian Medical Residency Interview Tour¹⁰

As I write, the countries of the world are meeting in Glasgow to find ways to avoid a planet meltdown and the end of life as we humans know it. Health professionals rarely consider the environmental impact of their interventions, whether from procedures, medications, tests, or travel. The US health sector is responsible for 8.5% of national greenhouse gas (GHG) emissions and 25% of the global health sector GHG emissions.¹¹ Preservation of human life, not flora, fauna, or the planet, has been our focus. During the COVID-19 pandemic era, researchers have published estimates of reduced travel costs for residency

applicants and reduced lost work time for faculty and residents, but not potential environmental effects.

In this paper Liang et al measure the typical carbon footprint of Canadian medical student residency interview travel, known as the “CaRMS tour,” and provide estimates for other interview approaches.¹⁰ The authors surveyed all Canadian medical students in the pre-pandemic 2020 CaRMS tour for interview schedules and flight itineraries. The survey response rate was 33% (960 of 2943) of all applicants, with good representation in terms of schools, regions, languages, and specialty choices (the Quebec and Ontario regions were underrepresented in the respondent sample). In addition to these demographic factors, the survey asked for interview itineraries: cities in order, travel methods, and layovers. From this data, GHG emissions were calculated and compared to the average yearly GHG emissions per Canadian household and average GHG emission reductions associated with sustainable changes for the planet.

The authors carefully calculated GHG emissions data using best practices. For example, takeoff and landing contribute the highest emissions as well as airport taxiing and waiting in line at airports; thus, short-haul flights produce the highest emissions per kilometer traveled. Interestingly, all urology interviews in Canada are conducted at just one site, so this data was separately analyzed.

For alternative approaches, the authors created 4 models with the same applicant data: (1) 3 centralized interview sites (Calgary, Toronto, Halifax); (2) one central interview site (Toronto); (3) low-technology virtual interview (interviewers and applicants use laptops); and (4) high-tech virtual interview (applicants use a laptop; interviewers use television screens, high-definition cameras, microphones, and a sound system). They calculated GHG emissions for each scenario.

For the 2020 CaRMS tour, Quebec students and urology applicants had the lowest GHG emissions. For urology, this reflects use of a single interview site. For Quebec, with many French-speaking medical students, a high proportion of applicants did not travel for interviews. The overall flight emissions of applicants represented 35% of the annual Canadian household per capita emissions. Twenty-eight percent of applicants exceeded their *total annual individual carbon budget* required to limit global temperature rise to 2°C (ie, the projected 2050 carbon footprint for each human).

In the 4 interview models, the emission reductions, as compared to the actual 2020 interview tour, were: 14% less for the 3-site interview model; 75% less for the one-site interview model; 98% less for the high-tech virtual interview model; and greater than 99.9% less for the low-tech virtual interview model. The authors

concluded that the current residency application model contributed considerable personal GHG emissions.

My take-home from this well-crafted study: Studies of interventions, whether in medical education or clinical realms, must include environmental effects, not just board passage rates, wellness measures, or improved patient outcomes. Greater attention to our education carbon footprints, which are wasteful in many ways, is appropriate, essential, and urgent.

Runner Up: Is There a Replication Crisis in Medical Education Research?¹²

This paper delineates sources for research integrity problems, the “sloppy science” alluded to in Tony’s pick for best article.¹ Hope et al review the conundrum that many research results, including in medical education, cannot be replicated by others.¹² The authors explore risk factors that may lead to this problem, which may not only reduce research credibility but also risk implementation of unhelpful or even harmful interventions. Key factors endangering replication include small sample sizes; small effect sizes; exploratory research designs; flexibility and variability in design, analysis, and outcome measures; conflicts of interest; and competing research teams, particularly for “hot” areas of research. The authors suggest potential solutions: 2 key ones are more collaboration among medical educators and more careful scrutiny of work by editors.

Medical educators are not unique in being attracted to the next “shiny new thing”—or intervention. I recommend this paper for HPE educators and researchers, to remind us to consider these factors when examining that new thing.

Lainie Yarris’s Pick

Supported Independence: The Role of Supervision to Help Trainees Manage Uncertainty¹³

As GME continues the transition to a competency-based model, there is a consistent emphasis on entrustment, or how clinician educators make real-time decisions regarding what they trust learners to do, and with what degree of supervision. The desired outcome is independence. Learners who are entrusted to perform activities without supervision gain autonomy but lose the direct observation that is necessary for informed feedback, deliberate practice, and corrective action. This tension—between entrustment and supervision—is a hot topic in GME research.

In this thoughtful paper Ilgen et al explore the role of supervision from the learner perspective.¹³ Rather than focusing on how educators make entrustment decisions, the authors consider how supervision

impacts trainees' experiences with uncertainty. This work explores trainee perceptions of the competing tensions of clinical independence, supervisory support, and patient safety.

The authors used a constructivist ground theory approach and applied a conceptual framework that drew upon Vygotsky's zone of proximal development, and the idea that learners may be able to push their limits with the guidance of a trusted supervisor. The authors applied a critical incident technique to elicit instances of uncertainty as a springboard for trainees to discuss perceptions of their supervisor's role during moments of challenge and discomfort. Individual interviews were conducted with 13 emergency medicine residents at 2 training programs (United States and Canada) over a 3-month period, and then analyzed iteratively to determine codes, categories, themes, and relationships among themes. Emergency medicine was selected, as learners encounter uncertainty frequently.

This study found that learners desire "supported independence" that facilitates their ability to work semi-autonomously through challenging situations. Residents described experiencing supported independence as "borrowing comfort" from attendings. Residents specifically invoked this supportive independence and faced challenges when the right level of supervision was not available to them. The authors found the concept of trust was important, particularly in the residents' need to have trust in their supervisors' ability to assist them when needed. Further, appropriate supervision empowered residents with the confidence to push their comfort zones in future encounters.

My take-home from this rigorous qualitative study:

This paper is an excellent example of applying a well-designed and well-executed qualitative methodology to a phenomenon that is important to GME yet little understood. Ilgen and colleagues' findings are thought-provoking and raise questions that are crucial to consider as we move forward in the competency-based medical education era. *How do we balance autonomy and supervision? How do supervision, trust, and confidence affect learner discernment and performance? How can educators facilitate learners' abilities to expand their comfort zones and take risks that are necessary for learning but still safe for patients?* This paper is a must-read for GME educators and leaders.

Runner Up: The Shift From Disbelieving Underperformance to Recognizing Failure: A Tipping Point Model¹⁴

This paper by Gingerich et al, also a constructivist grounded theory study, explores how supervisors

recognize when trainees are underperforming.¹⁴ Supervisors may be reluctant to report failing performances, and the stakes of reporting are high—both for trainees and program directors. But the stakes of *not* doing so may be higher. The authors had previously found that, when supervisors encounter unexpected underperformance, a period of disbelief precedes acknowledgement of the need for remediation or dismissal. This new study explores the experiences of Canadian supervisors with underperforming trainees to shed light on how the shift from disbelieving underperformance to recognizing behavior occurs. They found 3 themes contributing to the shift: *accumulation of significant incidents*, *discovery of an egregious error*, and *having a deficit be pointed out by someone else*. Recognizing failure incited emotions in the supervisor, including anger, certainty, and a desire to prevent harm.

I admired the authors' willingness to explore a difficult question in this provocative study. Although failure is rare, I suspect that serious concern about trainee performance is not. Program directors who are considering trainee remediation or dismissal may experience a sense of isolation and hesitation and may find little guidance in the literature regarding how to navigate the process. Illuminating the phenomenon of supervisor disbelief of underperformance is necessary to recognize and intervene upon failure—which ultimately benefits patients, trainees, and program directors alike.

Editors' Group Pick

Don't Be Reviewer 2! Reflections on Writing Effective Peer Review Comments¹⁵

We greatly enjoyed reading this paper by Watling et al.¹⁵ The authors focus on how to avoid belittling or condescending reviews in order to be perceived as—and actually be—colleagues providing helpful advice to peers. This paper is part of the ongoing *Perspectives on Medical Education* Writer's Craft series, which we highly recommend as well. In this paper, the authors remind us of red flag words, reported in an earlier linguistic analysis of harsh reviews,¹⁶ such as "absurdly," "illogical," and "ridiculous." In an engaging writing style the authors promote several strategies: be mindful of the challenge of unpaid, often unrecognized peer review; have a conversation with the author; redirect negative feedback toward the paper, rather than to the authors; provide editors with clear guidance on your overall take on the paper; couch your assessments as your opinions, not facts; avoid nitpicking and long-winded reflections about the paper (vs specific requests); provide your name as reviewer; and keep your own ego out of the review.

By far the majority of the JGME reviews we read are collegial and nurturing, including when the reviewer recommends rejection. Such reviews ensure that the authors learn something from the process. When writing, imagine that the review is going to a good friend and adjust your language accordingly. This article is a great read for those who review for any type of journal, not just medical education.

We hope you enjoyed this year's "best paper" selections and that you offer some of your own. We look forward to seeing you, in person, in the future. Happy reading!

References

- ten Cate O. The ethics of health professions education research: protecting the integrity of science, research subjects, and authorship [published online ahead of print September 14, 2021]. *Acad Med*. doi:10.1097/ACM.0000000000004413
- Artino AR Jr, Driessen EW, Maggio LA. Ethical shades of gray: international frequency of scientific misconduct and questionable research practices in health professions education. *Acad Med*. 2019;94(1):76–84. doi:10.1097/ACM.0000000000002412
- Janssen Stein J, Langerhuizen DWG, Kerkhoffs GMMJ, Ring D. Payments by industry to residency program directors in the United States: a cross-sectional study [published online ahead of print May 18, 2021]. *Acad Med*. 10.1097/ACM.0000000000004166
- Liu JJ, Bell CM, Marelski JJ, Detsky AS, Cram P. Payments by U.S. pharmaceutical and medical device manufacturers to U.S. medical journal editors: retrospective observational study. *BMJ*. 2017;359:j4619. doi:10.1136/bmj.j4619
- Wazana A. Physicians and the pharmaceutical industry: is a gift ever just a gift? *JAMA*. 2000;283(3):373–380. doi:10.1001/jama.283.3.373.
- Kelleher M, Kinnear B, Sall DR, et al. Warnings in early narrative assessment that might predict performance in residency: signal from an internal medicine residency program. *Perspect Med Educ*. 2021;10(6):334–340. doi:10.1007/s40037-021-00681-w
- Simpson D, Sullivan GM, Artino AR, Deiorio NM, Yarris LM. Envisioning graduate medical education in 2030. *J Grad Med Educ*. 2020;12(3):235–240. doi:10.4300/JGME-D-20-00292.1
- Sherbino J. Can narrative comments provide educational GPS? Episode 335. Royal College of Physicians and Surgeons of Canada. <https://keylimepodcast.libsyn.com/335-can-narrative-comments-provide-educational-gps>. Accessed December 10, 2021.
- LaDonna KA, Cowley L, Touchie C, LeBlanc VR, Spilg EG. Wrestling with the invincibility myth: exploring physicians' resistance to wellness and resilience-building interventions [published online ahead of print August 10, 2021]. *Acad Med*. doi:10.1097/ACM.0000000000004354
- Liang KE, Dawson JQ, Stolan MD, et al. A carbon footprint study of the Canadian medical residency interview tour. *Med Teach*. 2021;43(11):1302–1308. doi:10.1080/0142159X.2021.1944612
- Dzau VJ, Levine R, Barrett, G, Witty A. Decarbonizing the U.S. health sector—a call to action. *N Engl J Med*. 2021;385(23):2117–2119. doi:10.1056/NEJMp2115675
- Hope D, Dewar A, Hay C. Is there a replication crisis in medical education research? *Acad Med*. 2021;96(7):958–963. doi:10.1097/ACM.0000000000004063
- Ilgen JS, de Bruin ABH, Teunissen PW, et al. Supported independence: the role of supervision to help trainees manage uncertainty. *Acad Med*. 2021;86(suppl 11):81–86. doi:10.1097/ACM.0000000000004308
- Gingerich A, Sebok-Syer SS, Lingard L, Watling CJ. The shift from disbelieving underperformance to recognizing failure: a tipping point model [published online ahead of print October 20, 2021]. *Med Educ*. doi:10.1111/medu.14681
- Watling C, Ginsburg S, Lingard L. Don't be reviewer 2! Reflections on writing effective peer review comments. *Perspect Med Educ*. 2021;10(5):299–303. doi:10.1007/s40037-021-00670-z
- Hyland K, Jiang FK. "This work is antithetical to the spirit of research": an anatomy of harsh peer reviews. *J Engl Acad Purp*. 2020;46:100867. doi:10.1016/j.jeap.2020.100867



Gail M. Sullivan, MD, MPH, is Editor-in-Chief, *Journal of Graduate Medical Education* (JGME), and Associate Director for Education, Center on Aging, and Professor of Medicine, University of Connecticut Health Center; **Deborah Simpson, PhD**, is Director of Education, Academic Affairs at Advocate Aurora Health, and Clinical Adjunct Professor of Family & Community Medicine, Medical College of Wisconsin, University of Wisconsin School of Medicine and Public Health, and Deputy Editor, JGME; **Anthony R. Artino Jr, PhD**, is Professor and Associate Dean for Evaluation and Educational Research, The George Washington University School of Medicine and Health Sciences, and Deputy Editor, JGME; **Lalena M. Yarris, MD, MCR**, is Professor of Emergency Medicine, Oregon Health & Science University, and Deputy Editor, JGME; and **Nicole M. Deiorio, MD**, is Professor, Department of Emergency Medicine, and Associate Dean, Student Affairs, Virginia Commonwealth University School of Medicine, and Executive Editor, JGME.

Corresponding author: Gail M. Sullivan, MD, MPH, University of Connecticut Health Center, gsullivan@uchc.edu, Twitter @DrMedEd_itor