

Impact of medical education on patient safety: finding the signal through the noise

Jasmine Hwang, Rachel Kelz 

Department of Surgery,
University of Pennsylvania,
Philadelphia, Pennsylvania, USA

Correspondence to

Dr Rachel Kelz, Surgery,
University of Pennsylvania,
Philadelphia, PA 19104, USA;
rachel.kelz@uphs.upenn.edu

Accepted 5 September 2022

Medical education and patient care are inextricably linked. At this time, with the limitations of simulation training and our continued dependence on educated physicians for clinical decision-making, one domain cannot exist without the other. The involvement of medical trainees in patient care means it is vital that the impact of changes to medical training programmes on patient outcomes are assessed with well-designed studies. The study, “National improvements in resident physician-reported patient safety after limiting first-year resident physicians’ extended duration work shifts,”¹ by Weaver and colleagues published in this month’s issue of *BMJ Quality and Safety* signals the need for a robust discussion of education policy research within the field of medicine.

Before specifically addressing the approach used by Weaver *et al*,¹ a review of the complexities involved in studying the impact of medical education policy is warranted.² The field of ‘education outcomes research’ examines the impact to patients of changes to the educational process and clinical learning environment. As with all robust research, studies must be designed to answer a specific question. Typically, this necessitates an explicit definition of the exposure and outcomes to be examined. The exposures under investigation such as the change in education policy evaluated by Weaver *et al* are often straightforward, but the outcomes are frequently complex. For example, changes to education policy may influence the care delivered by trainees immediately or years later depending on the nature of the change. A further challenge to a well-conceived education outcomes study is the selection of the outcome measure to be evaluated. Ideally, studies on the impact of patient safety should

focus on an objective outcome measure rather than a proxy. However, this may not always be feasible. Finally, changes to education policy or practice are not made in isolation. There are usually multiple other influences occurring at any one time that can impact any level of training (figure 1). As such, each study must carefully consider how other simultaneous influences such as new technology (eg, robotic surgery) and alterations in practice patterns (eg, shift to outpatient procedures) might impact the study findings. Each of the aforementioned challenges must be addressed in order to permit the conduct of robust, reliable and reproducible works.

To provide context for international readers, in the USA, first-year residents are typically in the first year of post-graduate training after completion of an undergraduate degree and medical school. These physicians-in-training are often referred to as interns or first-year residents. They serve as frontline physicians who place orders for patients and are often the first to receive calls from the nursing staff in the inpatient setting. Long work hours have been associated with resident impairment and medical errors.^{3 4} In 2011, the Accreditation Council for Graduate Medical Education (ACGME) implemented restrictions to resident duty hours which included an 80-hour limit to the work week, 16-hour limit of continuous in-hospital duty and at least 1 day off every week on average over 4 weeks. A resident’s shift is typically anywhere from 12 to 30 hours. The 16-hour work limit reduced the consecutive number of hours that first-year residents could work in a single shift from 30 to 16. The duty hour reduction was revoked in 2017 because several large well-designed studies demonstrated



► <http://dx.doi.org/10.1136/bmjqs-2021-014375>



© Author(s) (or their employer(s)) 2022. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Hwang J, Kelz R. *BMJ Qual Saf* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjqs-2022-015054

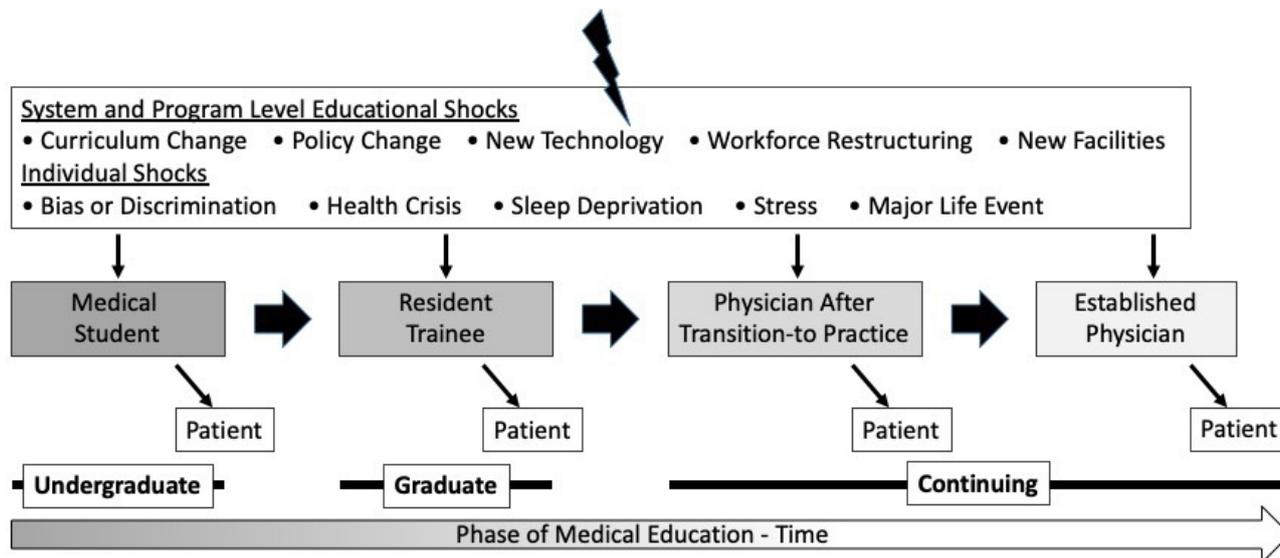


Figure 1 Effects of educational shocks on learners and patients across the educational continuum.

that flexible work hours were non-inferior to those imposed by the 16-hour limit.⁵⁻⁷

With that background, we can now explicitly examine the work by Weaver *et al* which aimed to examine the effects of the ACGME 16-hour work limit for first-year residents on patient safety. The authors found that the 16-hour work rule was associated with a 32% reduced risk (RR) of resident physician reported significant medical errors (RR 0.68; 0.64–0.72), a 34% reduced risk of reported preventable adverse events (RR 0.66; 0.59–0.74) and a 63% reduced risk of reported medical errors resulting in patient death (RR 0.37; 0.28–0.49). The prospective nature of data collection and use of the same survey instrument for two time periods (2002–2007; 2014–2017) are strengths of the study design. Furthermore, the study subjects represent the full complement of ACGME accredited specialty programmes which would normally improve the generalisability of the study findings. However, as only 9% of residents consented to participation, the generalisability may be limited.

There are several design elements of the study that limit complete confidence in the reliability of the findings. First, as acknowledged by the authors, the use of resident perceptions of errors may be subject to recall bias and introduce self-reporting bias. As a proxy, it is also one step removed from the most important outcome—objectively measured medical error. This study relies on participants to self-report both the true exposure, duty hours and the outcome of medical errors. It has been well established that residents may falsely report duty hours,⁸ especially early on in training,⁹ and that residents tend to selectively report errors depending on the outcome of the error.¹⁰ Furthermore, the study participants were asked to report ‘significant’ medical errors, which leaves it open to the resident to determine what qualifies as

significant. In addition, residents were asked to self-report whether ‘they’ had personally made significant medical errors. This can be difficult to discern as rarely is one individual to blame for adverse events: they are multistep and complex processes. Adverse events are difficult to capture and physicians in particular are known for self-reporting not only fewer but also a narrower spectrum of adverse events when compared with nurses or patients.^{11 12} Regarding the mode of survey distribution, residents are less likely to report adverse events in email queries, as used in this study, compared with face-to-face interviews, which may have resulted in under-reporting.¹³

An additional measurement concern exists because Weaver *et al* did not consider the expansion of education on quality and safety that occurred in parallel with the implementation of the 16-hour rule. Because discrete educational sessions on safety event reporting result in increased reporting for several weeks,¹⁴ it is possible that the study participants enrolled after the 16-hour rule implementation reported a greater proportion of these events because they had recently been prompted to keep errors on the top of their minds. Moreover, a coincident focus on safety culture, attributed to the implementation of the clinical learning environment initiative in 2012, may have increased the tendency to report adverse events^{10 15} in the post-period as well. If this was the case, it is possible that the protective effect was even greater than estimated by this study. Alternatively, there has been an increased presence of advanced practice providers on inpatient teams and units in the USA who may provide a safeguard for patients and first-year residents against medical errors. This safety feature did not exist prior to duty hour reform and likely expanded alongside the 16-hour work restrictions. These confounders would bias study results in both directions, either overestimating or

underestimating adverse events. It is difficult to evaluate the magnitude of the effect these simultaneous influences had on learners and patients, or the results of this study.

The findings of Weaver and colleagues are at odds with several of the other high-quality, published studies that found little to no association between the introduction of the 16-hour work rule and patient outcomes or resident wellness in surgery and internal medicine.^{16–18} Weaver *et al* discuss this discrepancy by asserting that the outcome measures used in other studies, specifically mortality and morbidity, are not directly related to the work of first-year resident-physicians and, therefore, not sensitive enough to detect the adverse impact of the extended work hours on patient care. We argue against this assumption because first-year residents are often first responders and can provide early interventions that save lives. Similarly, the most common non-technical adverse event in surgery is medication related.¹⁹ As the primary caretaker of their patients responsible for medication reconciliation and inpatient orders, first-year residents are inherently connected with medication-related adverse events.²⁰ Finally, we agree that residents at different levels in their training may impact various safety measures in different areas and to different extents.

Questions related to best practices in medical education are important to our learners, our patients and the public at-large. While there are no magic metrics, education outcomes research should move towards designs that can handle the complexity of the study questions and assessment metrics that objectively quantify the outcomes of interest. Multiple approaches to study the impact of the 16-hour work-hour limit and other changes in medical education are needed as they can serve to complement one another. Pragmatic trials and observational studies can be used to perform many of the patient-centred assessments. Because of concerns regarding the retrospective analysis in this study, a prospectively designed pragmatic trial, such as that employed by the FIRST trial where duty hours would be assigned at the programme level, might have been useful.²¹ Furthermore, the limited response rate could have been addressed by the use of data from mandatory surveys such as the ACGME resident survey. To address issues of self-reporting, objective data on adverse outcomes reported to the state or federal agencies like the Centers for Medicare & Medicaid Services could be used. A difference-in-difference approach would have been helpful to control for temporal factors other than the selected exposure.^{22 23} New educational approaches, materials and rules should be evaluated just as any new treatment that may impact patient care would be tested before receiving approval. Learner-centred assessments require thought and consideration to make sure they are free from bias and discrimination particularly

as there are many educational shocks throughout the educational continuum.

Ultimately, this study provides a new perspective on the effects of the 16-hour consecutive work-hour limit policy on patient care. The authors bring up valid concerns about patient safety with the reversal of this policy in 2017. However, given the inherent design limitations, we suggest caution in the re-implementation of the 16-hour work-hour limit based solely on the findings of this study. Instead, medical errors and workforce safety should be closely monitored to make sure that the clinical environment remains safe following the rollback of the 16-hour work rule. Also, if the authors can collect a more robust sample of responses, a difference-in-difference study design might offer a better understanding of the 16-hour rule implementation on medical error.

Twitter Rachel Kelz @surgeriespice

Contributors Each author contributed to the concept, drafting and critical revision of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Commissioned; internally peer reviewed.

ORCID iD

Rachel Kelz <http://orcid.org/0000-0001-6635-5961>

REFERENCES

- Weaver MD, Landrigan CP, Sullivan JB, *et al*. National improvements in resident physician-reported patient safety after limiting first-year resident physicians' extended duration work shifts: a pooled analysis of prospective cohort studies. *BMJ Qual Saf* 2022. doi:10.1136/bmjqs-2021-014375. [Epub ahead of print: 10 May 2022].
- Sellers MM, Niknam BA, Kelz RR. A framework for the assessment of graduate medical education. *JAMA Surg* 2017;152:715–6.
- Philibert I. Sleep loss and performance in residents and nonphysicians: a meta-analytic examination. *Sleep* 2005;28:1392–402.
- Landrigan CP, Rothschild JM, Cronin JW, *et al*. Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med* 2004;351:1838–48.
- Choma NN, Vasilevskis EE, Sponsler KC, *et al*. Effect of the ACGME 16-hour rule on efficiency and quality of care: duty hours 2.0. *JAMA Intern Med* 2013;173:819–21.
- Bilimoria KY, Chung JW, Hedges LV, *et al*. National cluster-randomized trial of duty-hour flexibility in surgical training. *N Engl J Med* 2016;374:713–27.
- Silber JH, Bellini LM, Shea JA, *et al*. Patient safety outcomes under flexible and standard resident duty-hour rules. *N Engl J Med* 2019;380:905–14.
- Drolet BC, Schwede M, Bishop KD, *et al*. Compliance and falsification of duty hours: reports from residents and program directors. *J Grad Med Educ* 2013;5:368–73.

- 9 Zamani N, Fallon SC, Bremer A, *et al.* Using an anonymous, resident-run reporting mechanism to track self-reported duty hours. *Am J Surg* 2019;218, :225–9.
- 10 Kaldjian LC, Jones EW, Wu BJ, *et al.* Reporting medical errors to improve patient safety: a survey of physicians in teaching hospitals. *Arch Intern Med* 2008;168:40–6.
- 11 Rowin EJ, Lucier D, Pauker SG, *et al.* Does error and adverse event reporting by physicians and nurses differ? *Jt Comm J Qual Patient Saf* 2008;34:537–45.
- 12 Pham JC, Ashton MJ, Kimata C, *et al.* Surgical site infection: comparing surgeon versus patient self-report. *J Surg Res* 2016;202:95–102.
- 13 Weingart SN, Ship AN, Aronson MD. Confidential clinician-reported surveillance of adverse events among medical inpatients. *J Gen Intern Med* 2000;15:470–7.
- 14 Hoffman RL, Morris JB, Kelz RR. The "Quality minute"—a new, brief, and structured technique for quality improvement education during the Morbidity and Mortality Conference. *JAMA Surg* 2017;152:883–4.
- 15 Appelbaum NP, Dow A, Mazmanian PE, *et al.* The effects of power, leadership and psychological safety on resident event reporting. *Med Educ* 2016;50:343–50.
- 16 Rajaram R, Chung JW, Jones AT, *et al.* Association of the 2011 ACGME resident duty hour reform with general surgery patient outcomes and with resident examination performance. *JAMA* 2014;312:2374–84.
- 17 Patel MS, Volpp KG, Small DS, *et al.* Association of the 2011 ACGME resident duty hour reforms with mortality and readmissions among hospitalized Medicare patients. *JAMA* 2014;312:2364–73.
- 18 Desai SV, Asch DA, Bellini LM, *et al.* Education outcomes in a duty-hour flexibility trial in internal medicine. *N Engl J Med* 2018;378:1494–508.
- 19 Schwendimann R, Blatter C, Dhaini S, *et al.* The occurrence, types, consequences and preventability of in-hospital adverse events – a scoping review. *BMC Health Serv Res* 2018;18:1–3.
- 20 Vira T, Colquhoun M, Etchells E. Reconcilable differences: correcting medication errors at hospital admission and discharge. *Qual Saf Health Care* 2006;15:122–6.
- 21 Sox HC, Lewis RJ. Pragmatic trials: practical answers to "real world" questions. *JAMA* 2016;316:1205–6.
- 22 Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. *JAMA* 2014;312:2401–2.
- 23 Kelz RR, Niknam BA, Sellers MM, *et al.* Duty hour reform and the outcomes of patients treated by new surgeons. *Ann Surg* 2020;271:599–605.