



Perspective

Leaping without Looking — Duty Hours, Autonomy, and the Risks of Research and Practice

Lisa Rosenbaum, M.D.

In 2014, Facebook users were furious to discover that they'd unwittingly been experimented on.¹ Researchers had randomly assigned users to news feeds with reduced “positive” content or reduced

“negative” content and found that happy posts beget happy posts and that grim ones beget grim ones.² Although that may now seem obvious, previous evidence had suggested that because we tend to compare ourselves to others, exposure to positive content compromises users' well-being. There was thus no reason to believe that the status quo — news feeds curated by an algorithm tailored to users' viewing habits — was any “safer” than the experimental interventions. And given Facebook's reach, there were compelling reasons to find out. Nevertheless, the results triggered outrage that 700,000 users had been exposed to potential emotional damage without their consent.

Similar accusations have been

leveled at investigators who are comparing the 2011 duty-hour restrictions imposed by the American Council of Graduate Medical Education (ACGME) with more flexible shift lengths for residents. The Flexibility in Duty Hour Requirements for Surgical Trainees (FIRST) trial, whose results are now reported by Bilimoria et al. in the *Journal*, compared 59 surgical training programs randomly assigned to an ACGME-compliant schedule with 58 granted flexibility in designing shift lengths (still within an 80-hour workweek). The ongoing Individualized Comparative Effectiveness of Models Optimizing Safety and Resident Education (iCOMPARE) trial involves internal medicine programs. Both used cluster randomization

at the residency-program level, and neither required consent of residents or patients. That consent waiver has drawn criticism from Public Citizen and the American Medical Student Association, which in open letters to the Office for Human Research Protections (OHRP) accuse the investigators of “egregious ethical and regulatory violations.”^{3,4}

The allegations, focused primarily on “serious health risks” to residents from long shifts, are dizzily tautological. The critics claim it's unethical not to obtain residents' consent; but because pressure on residents to conform makes seeking their consent akin to coercion, that's unethical too. Thus, there's no ethical way to study the duty-hour rules in a randomized fashion. But that's fine, because we already know they're beneficial; we know that because the ACGME made the rules in the first place. And if the trials found otherwise,

their results challenging the status quo would be suspect because the investigators, who have publicly acknowledged the need for data to inform policy, are consequently too biased to generate those data.

To unpack these allegations, it's important to understand that even if the trials are considered human-subjects research, there are circumstances under which federal rules deem it ethical to waive consent. The key one here is that the incremental risk posed by the research should be, at most, minimal. For trials like these that evaluate a standard practice, the question becomes: Is there equipoise between the status quo and investigational groups in terms of possible risks? Though the letters to OHRP claim otherwise, the answer is unequivocally yes. The complaints ignore a considerable body of research suggesting, as Bilimoria et al. point out, that duty-hour reforms have not improved patient safety; some trials have even raised concerns that they've actually worsened quality of care and patient outcomes.

As for risks to residents, the letters cite data suggesting that fatigue causes harms such as increased motor vehicle accidents, needlesticks, and burnout. Yet there's little evidence to suggest that shorter hours have reduced occupational hazards or burnout rates. Though I suspect that these findings partly reflect the emotional toll of "work compression" and the reality that many trainees don't actually sleep more, they also speak to a fundamental challenge in improving care: the factors affecting physicians' performance are so numerous and interdependent that no single var-

iable, such as sleep, can be understood or targeted in isolation. Because of the unknown real-life consequences of such myriad interactions, no drug would be approved solely on the basis of laboratory evidence. Yet we require neither consideration of complexity nor rigorous studies before implementing policies with similarly broad implications. Why?

Bioethicist and legal scholar Michelle Meyer has described our "tendency to view a field experiment designed to study the effects of an existing or proposed practice as more morally suspicious than an immediate, universal implementation of an untested practice." She argues that people in power often rely on intuition in creating and implementing wide-reaching policies. Indeed, neither residents nor patients consented to the ACGME rules, yet no one finds this omission ethically suspect. Moreover, intuition seems particularly salient to debates over duty hours, since everyone knows how it feels to be tired. Unfortunately, few people know how it feels to see a patient through illness, spend a fifth of your time engaged in hand-offs, leave halfway through an operation because your shift's up, or perceive resentment in your supervisors who think you have it easier than they did. Given such trade-offs and uncertainties, it's not just ethical but laudable to comparatively evaluate duty-hours policies. The question then becomes: Can the research be accomplished if consent is required?

The Facebook experiment's results would have been invalid had consent been sought, since we couldn't determine how much users adjusted their emotional

content because they knew it was being monitored. Similarly, requiring residents' consent in duty-hour trials would render the results uninterpretable, given the selection bias that would be introduced if those preferring longer hours were more likely to participate.

The challenges with regard to patients are more pragmatic. Consider, for instance, caring for a man with a myocardial infarction. After obtaining his consent for percutaneous coronary intervention, you'd have to add, "I also need your consent to be cared for by residents who are working longer hours." If he said no, would you have to transfer him, as heart muscle continued to die, to a nonteaching hospital? Surely here the risk posed by seeking consent is greater than that from the research itself.

Moreover, as we examine the implications for efforts to develop "learning health systems," a corollary of this hypothetical situation is worth considering. Imagine telling a patient, "I need your permission to care for you at a hospital where we're using a new electronic health record, are basing your doctor's reimbursement on whether you stay healthy, and are under pressure to discharge you quickly and make sure you don't come back. We don't really know how all this will affect your health, but we believe it's for the better. Can you sign here?"

The point is that our approach to human-subjects research perpetuates a misleading distinction between risks posed by research and those posed by practice, demanding greater scrutiny for investigative efforts while assuming that untested practice is safe. In describing this phenomenon, Meyer cites the moratorium that

the OHRP imposed on a study assessing a checklist designed to reduce catheter-related bloodstream infections because researchers hadn't obtained physicians' or patients' consent. The OHRP explained that its regulations don't apply when institutions are merely "implementing" practices aiming to improve care, but if they're "planning research activities examining the effectiveness of interventions to improve the quality of care, then the regulatory protections are important to protect the rights and welfare of human research subjects." This double standard leaves us, paradoxically, with unregulated practices that may be ineffective and unsafe because we can't surmount the regulatory hurdles to conducting research to improve them.

To address this problem, we must understand the values of the people we're professing to protect. In one relevant study, Halpern and colleagues asked patients undergoing dialysis to imagine two hypothetical scenarios.⁵ In the "research scenario," patients in a trial are randomly assigned to a prespecified dialysis duration of 4.5 hours or a duration at the physician's discretion (both approaches are within the stan-

dard of care). In the "clinical care scenario," patients receive dialysis for a duration determined by a protocol (also common practice). Participants were more willing in the research than the practice setting to give up their own decision-making autonomy, including written informed consent. They recognized the value of research and didn't perceive the hypothetical study as posing higher risk than ordinary care. But they expressed deep reservations about compromising physicians' autonomy to individualize treatment absent compelling reasons for doing so.

This last finding highlights the ultimate irony of both duty-hour restrictions and objections to studying them: we've created an educational system that compromises trainees' freedom to judge for themselves when their patients need them. The value that physicians and patients place on such autonomy is not measurable in mortality rates or hours slept but should remain foremost in our discussions. An essential contribution of the duty-hour trials is that, in assessing flexibility itself, they remind us that autonomy is an ethical concept that matters to both doctors and

patients — in research and in practice.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.

Dr. Rosenbaum is a national correspondent for the *Journal*.

This article was published on February 2, 2016, at NEJM.org.

1. Meyer MN. Two cheers for corporate experimentation: the A/B illusion and the virtues of data-driven innovation. *Colo Tech LJ* 2015;13(2):273.
2. Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proc Natl Acad Sci U S A* 2014;111:8788-90.
3. Carome MA, Wolfe SM, Almashat S, Hall DV. Letter to Jerry Menikoff, director, and Kristina Borrer, director, Division of Compliance Oversight, Office for Human Research Protections, Department of Health and Human Services, regarding iCOMPARE trial. November 19, 2015 (<http://www.citizen.org/documents/2283.pdf>).
4. Carome MA, Wolfe SM, Almashat S, Hall DV. Letter to Jerry Menikoff, director, and Kristina Borrer, director, Division of Compliance Oversight, Office for Human Research Protections, Department of Health and Human Services, regarding FIRST trial. November 19, 2015 (<http://www.citizen.org/documents/2284.pdf>).
5. Kraybill A, Dember LM, Joffe S, et al. Patient and physician views about protocolized dialysis treatment in randomized trials and clinical care. *AJOB Empirical Bioethics* 2015 October 23 (Epub ahead of print).

DOI: 10.1056/NEJMp1600233

Copyright © 2016 Massachusetts Medical Society.