In 2014, Facebook users were furious to discover that they’d unwittingly been experimented on. 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.

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.”

The allegations, focused primarily on “serious health risks” to residents from long shifts, are dizzyingly 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 sta-
tus quo would be suspect be-
cause the investigators, who have
publicly acknowledged the need
for data to inform policy, are
consequently too biased to gen-
erate 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 an-
swer is unequivocally yes. The
complaints ignore a considerable
body of research suggesting, as
Bilimoria et al. point out, that
duty-hour reforms have not im-
proved 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 in-
creased 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 compres-
sion” 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’ per-
formance are so numerous and
interdependent that no single var-
able, such as sleep, can be under-
stood or targeted in isolation. Be-
cause of the unknown real-life
consequences of such myriad in-
teractions, no drug would be ap-
proved solely on the basis of lab-
oratory evidence. Yet we require
neither consideration of complex-
ity nor rigorous studies before
implementing policies with simi-
larly broad implications. Why?

Bioethicist and legal scholar
Michelle Meyer has described our
“tendency to view a field experi-
ment designed to study the ef-
effects of an existing or proposed
practice as more morally suspi-
cious than an immediate, uni-
versal implementation of an untest-
ed practice.” She argues that
people in power often rely on in-
tuition in creating and imple-
menting wide-reaching policies.
Indeed, neither residents nor pa-
tients consented to the ACGME
rules, yet no one finds this omis-
sion ethically suspect. Moreover,
intuition seems particularly sa-
lent 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 en-
gaged in hand-offs, leave half-
way through an operation be-
cause 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 compara-
tively evaluate duty-hours poli-
cies. The question then becomes:
Can the research be accom-
plished if consent is required?

The Facebook experiment’s re-
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, requi-
ring residents’ consent in duty-hour
trials would render the results un-
interpretable, 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 infarc-
tion. After obtaining his consent
for percutaneous coronary inter-
vention, you’d have to add, “I also
need your consent to be cared for
by residents who are working lon-
er 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 cor-
ollary of this hypothetical situa-
tion is worth considering. Imag-
ine telling a patient, “I need your
permission to care for you at a
hospital where we’re using a new
electronic health record, are bas-
ing 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 per-
petuates a misleading distinction
between risks posed by research
and those posed by practice, de-
manding greater scrutiny for in-
vestigative efforts while assum-
ing 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. In the “research 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.


DOI: 10.1056/NEJMp1600233
Copyright © 2016 Massachusetts Medical Society.