Clin Orthop Relat Res (2026) 00:1-4
DOI 10.1097/CORR.0000000000003887

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

Not the Last Word

Published online: 26 February 2026

Not the Last Word: The Best Medical Students Want Residency Programs That Mentor, Not Monitor; Teach, Not Track

Joseph Bernstein MD[1]

The orthopaedic residency match is famously competitive. In 2023, for example, 32% of applicants failed to get a position [3]. Less well known but equally true is that programs compete fiercely to attract top candidates. In other words, the competition runs both ways.

Just as applicants in this scramble try to outdo one another—11 research papers become 12, four away rotations become five [2]—programs engage in a similar arms race. Constrained by fixed salaries, standard curricula, and set locations, programs compete by piling on perks, paying for surgical loupes, textbooks, and conference travel. In other words, the escalation runs both ways.

I propose a novel approach: Programs that seek to differentiate themselves should do so by actually being different.

Specifically, programs should reject the current big-data model of resident evaluation that dominates the landscape today. Right now, residency programs vacuum up vast amounts of data about their trainees, relying heavily on easily accessible but not necessarily meaningful metrics. A program that hopes to appeal to the best students should stop doing that.

To start, many residency programs place an outsized focus on standardized evaluation forms. I am not sure that's really helpful. I know that when I am asked to tick a categorical response to a long list of questions, by the 17th item, I am often focusing more on aligning the cursor with the box than on the response itself.

There is also excessive emphasis on Orthopaedic In-Training Examination (OITE) scores. The OITE is pitched as a so-called formative evaluation—namely, feedback "for the purpose of making improvements." However, in practice, the OITE is used more like a so-called summative evaluation: a higher stakes assessment administered after instruction as an "assessment or judgment of a learner" [5]. At some programs, in fact, certain privileges—such as the right to attend conferences or work moonlighting shifts—are granted only to residents who achieve a sufficiently high OITE score.

Then, there is the menace of digital surveillance. Every keystroke in the electronic medical record is time-stamped and stored. Smartphones emit constant streams of data and metadata. Chip-enhanced ID badges serve as backup bear trackers, even if the phone and its geolocation pinger are shut off. As Foucault noted correctly [4] (a phrase I never thought I'd type, but he's right), the mere possibility of being watched prompts people to internalize the observer's gaze, self-regulate, and live with a constant undercurrent of anxiety.

Worse than the system's ability to collect information is its inability to forget. We frail humans are not always big enough to forgive, but we are mercifully prone to forget. The steel-trap memory of a computer denies us that backdoor path to grace. Without the possibility of imperfect recall, every decision must be made with the knowledge that it could be revisited in some future forum, where the facts remain but the context is gone. This drives residents toward the safest, least educational choices. Program directors

A note from the Editor-in-Chief: We are pleased to present to readers of Clinical Orthopaedics and Related Research® *the next "*Not the Last Word*." The goal of this section is to explore timely and controversial issues that affect how orthopaedic surgery is taught, learned, and practiced. We welcome reader feedback on all of our columns and articles; please send your comments to eic@clinorthop.org.*

The author certifies that there are no funding or commercial associations (consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted article related to the author or any immediate family members.
All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.
The opinions expressed are those of the writers, and do not reflect the opinion or policy of *CORR*® or The Association of Bone and Joint Surgeons®.

J. Bernstein ✉, University of Pennsylvania, 424 Stemmler Hall, Philadelphia, PA 19104, USA, Email: orthodoc@uphs.upenn.edu

[1]Department of Orthopaedic Surgery, University of Pennsylvania, Philadelphia, PA, USA

might like to think that independent thought can be decoupled from independent action, but timid actors tend to become timid thinkers.

In practice, switching to a small-data approach requires a few concrete commitments from residency programs. First, a promise that no evaluation that counts toward a resident's record will rely on simplistic box-checking—if it isn't prose, it doesn't count. Second, program directors should decline to see individual residents' OITE scores; if they want to use OITE scores to measure their teaching, aggregated results will suffice. Third, the program should pledge not to use surveillance data from electronic records or ID badges to judge individual residents. They can't stop the data from being collected, but programs can deem it "privileged" and promise not to use any of it for the purpose of evaluation.

More broadly, a small-data approach requires clearly separating formative from summative evaluations. Day-to-day commentaries (e.g., corrective criticism on rounds, peer-to-peer feedback, review of surgical planning and postoperative debriefing) should be strictly formative and never used in official assessments. On the other hand, summative evaluations should be highly intentional and well thought out. For example, a program might want to incorporate Oxford/Cambridge-style oral exams at the end of each academic year to assess real understanding. Such an examination is good preparation for the American Board of Orthopaedic Surgery (ABOS) Part II test and allows for nuanced evaluation in an era that allows all of us to carry a complete library, AI-enhanced consultants, and other mundane ways to access information—like my go-to method of texting a colleague—in our pockets.

To counter the effects of indelible record keeping, I further propose what might be called a "no-data rotation"—a period in the program when few, if any, data are collected for summative evaluation. In my residency program, we spent 9 months away from the Mother Ship at community hospitals, free from constant assessment. This created a safe space and allowed for a "reputational growth spurt"—the chance to return to the home program with sharper skills, noticed in the same way high school classmates seem taller after summer break.

A small-data approach is not without cost. Generating and reviewing evaluations with intention demands effort, attention, and time. It requires genuine investment by the program. Forgoing big data also risks losing valid warning signals. Patterns such as declining case volume, documentation failures, or poor attendance can reveal burnout and other problems early enough to address them. Selective, human-mediated observation means some issues might go unnoticed until they surface dramatically and irreparably. Companies like Google and Facebook collect terabytes of user data not to play secret policeman, but because in vast oceans of data, valuable signals can be found. Last, critics might argue that reducing audits, surveillance, and metrics would leave residents unprepared for the audit culture of modern medicine.

However, none of these costs is prohibitive. First, there will be offsetting savings, since programs need not spend endlessly on perks if they invest in a better training environment. The loss of signal about resident behavior is likewise offset by the loss of noise. (Companies like Google and Facebook, after all, do not act on every tidbit in their ocean of information, but rather only those data points that have been shown to matter.) Last, it's silly to assume that training has to mimic every hardship of the real world. If those of us who had Medicare audits, EMR dashboards, and productivity metrics foisted on us in middle age could adapt (and we did), certainly a younger cohort of digital natives can manage as well.

In the end, by focusing on meaningful growth and learning rather than surviving the gantlet of big-data collection, programs will train better, more genuinely prepared doctors. That applicants will find this appealing is only a bonus. The benefits will run both ways.

Lori Jia MD

Second-year Resident, Cleveland Clinic Orthopaedic Residency Program

Two years ago, when I was applying to residency, I remember curious medical students everywhere asking about a program's "operative volume" and "resident autonomy." They asked questions like, "How often are residents double-scrubbed?" and "How many cases are residents doing skin-to-skin?"

These questions were usually answered in vague, nebulous terms ("don't worry, they're great!") because the nuances of orthopaedic surgery education simply cannot be captured by the percentage of cases done "skin-to-skin," nor can they be fully captured by OITE scores, vapid checkbox evaluations, or constant digital surveillance, as Dr. Bernstein describes. After all, in the same way that a high Step 2 score does not necessarily make for a great resident, a high OITE score does not necessarily make for a great surgeon. These scores are important, but not sufficient in isolation.

Instead, I think one of the best ways to both teach and assess learning is what

Not the Last Word

one hand attending at my program likes to call the "see one, teach one, do one" method: The first time she operates with a resident, she'll discuss what she's doing while the resident observes. The second time, she'll listen as the resident narrates what she's doing (and what she's about to do next). The third time, if her assessment of the resident's knowledge and understanding has been satisfactory, she'll hand the resident the knife.

This method of teaching and assessment is perhaps a small-scale version of the oral exam that Dr. Bernstein proposes—an adaptive, Socratic-style, teach-back exam performed in real time, intraoperatively. It allows the resident to achieve a higher level of knowledge and understanding of the finer details of surgery, far beyond what rote memorization of Orthobullets could accomplish. It increases the efficiency and effectiveness of learning because it is done in real time with visual and auditory cues using a teach-back format. Progress is individualized and driven by the resident's pace. And, above all, this intraoperative lesson does not require additional time outside of the operating room to complete.

Although universal implementation of these intraoperative assessments may be aspirational, the trend toward real-time, nuanced evaluation and feedback aligns well with the mission of the ABOS Knowledge, Skills, and Behavior (KSB) program [1]. In this program, residents are required to log cases and request evaluations on at least 80 cases per year, in which they are rated between 1 ("Attending provides maximal assistance") and 5 ("Attending did not need to be there") across 10 skills categories. Evaluations must be requested within 48 hours of the case and must be completed by the attending within 48 hours of the request.

While I think KSB evaluations could be simplified (perhaps just "case preparation," "knowledge of procedural steps," and "technical performance" instead of the current 10 skills categories), I think overall these evaluations strike an appropriate balance between the small-data and big-data approaches that Dr. Bernstein describes. KSB evaluations are detailed enough to provide nuance while still being suitable for analysis in the aggregate. I can envision programs generating reports for individual residents to identify relative strengths and weaknesses across the evaluated skills. I can also envision programs generating reports for the residency program, stratified by variables such as PGY level, clinical rotation, and specific skills, so they can address gaps as needed. These results may even provide concrete insights that programs can share with prospective applicants. For example, instead of making abstract claims like "we provide great operative volume and resident autonomy," a program could say, "The average PGY-2 at our program performs X cases across the academic year, with average scores on the shoulder service being a 2 and average scores on the arthroplasty service being a 3…"

As orthopaedic surgery educational tools continue to evolve, I anticipate Dr. Bernstein's prediction will be right. Programs will look toward developing new and better ways of teaching and assessing learning, which will, in turn, attract the most capable candidates.

Scott E. Porter MD, MBA

Department of Orthopaedic Surgery, Prisma Health

I generally enjoy these columns as a refreshing refrain from heavily scientific endeavors. This one, however, has a title and a body that are disjointed. In my opinion, very little of what was written has anything to do with what the best medical students want. Moreover, there is nothing to suggest that any of the remedies presented would be desired by a single resident, let alone the best ones.

In my experience, medical students have no idea what they want beyond the oft-expressed desires of "early operative exposure and autonomy," and "a well-rounded curriculum with all of the subspecialties represented." I would furthermore submit that we, as an academic orthopaedic community, have no idea how to objectively define what is "best." Previously, the USMLE Step I score served as a ubiquitous (albeit terrible) marker of merit that we tried to use as an identifier for medical students that would make good residents and pass their ABOS written examination. Dean's letters, GPA, class ranking, letters of recommendation, and the like are all metrics that have some objective comparative value, but ultimately fail miserably at defining the very nebulous term "best."

As a result, I would like to propose that we as an orthopaedic community stop pursuing "best," as it does not exist in an absolute form. Beautifully, each of us naturally has a different system of evaluation and a different scale for any given metric; with so much variability, an absolute can't exist.

In place of "the best," I propose that we pursue "the best match." This should not be confused with the National Resident Matching Program's Match; rather, this term is meant to capture the process of identifying the greatest compatibility or suitability between two items, which, in this case, is each individual applicant and each orthopaedic residency program. There are 213 orthopaedic training programs in the United States, and while all fight within an identical pool of residents, the fight should not be over the same individuals.

Not the Last Word

Rather, all efforts should go toward identifying the greatest compatibility and suitability between applicants and training programs, or stated another way, the best match. I firmly believe that residencies should do everything possible to promote their brands. Programs must start by identifying traits or resources that enjoy variability as opposed to those that are immutable and then promote the idea that only those who find value in a program's strengths, acceptance of that program's weaknesses, and a desire to join its environment should apply. We should do our very best to encourage residents to move away from the mentality of, "I'll go anywhere, I just want to get into orthopedic surgery," and toward one that critically evaluates prospective programs with an eye on compatibility.

Dr. Bernstein believes that focusing on meaningful growth and learning rather than big-data collection will result in better-trained, genuinely prepared doctors. Instead, I would refocus us on his premise of discussing the best medical student and implore our community to encourage the mutual selection by students and programs of the most compatible and suitable environment in a process that would then, truly, result in "the best match."

References

1. American Board of Orthopaedic Surgery. ABOS knowledge, skills, and behavior program. Available at: https://www.abos.org/abosksb/. Accessed February 3, 2026.
2. Bernstein J. Not the last word: the academic inflation reduction act. *Clin Orthop Relat Res*. 2024;482:1296-1301.
3. Katsma MS, Murray JB, Balazs GC, Speicher MR, Olsen AA. The real match rate? US DO students match into orthopaedic surgery at significantly lower rates than US MD students despite single accreditation. *JB JS Open Access*. 2025;10:e24.00002.
4. Logan S. The needle and the damage done: of haystacks and anxious panopticons. *Big Data & Society*. 2017;4.
5. Pinney SJ, Mehta S, Pratt DD, et al. Orthopaedic surgeons as educators: applying the principles of adult education to teaching orthopaedic residents. *J Bone Joint Surg Am*. 2007;89:1385-1392.