How Data Analytics Can Be Leveraged to Enhance Graduate Clinical Skills Education

Brian T. Garibaldi, MD, MEHP
McKenzie Hollon, MD, FASE
Michelle I. Knopp, MD
Abigail Ford Winkel, MD, MHPE
Jesse Burk-Rafel, MD, MRes
Holly A. Caretta-Weyer, MD, MHPE

Precision medical education (PME) leverages data about individual learners to guide personalized education.[1-3] A PME cycle starts with data inputs that are analyzed to generate insights about a learner. Insights then inform personalized interventions that lead to measurable outcomes. Outcomes inform iterative adjustments to the cycle.[3] PME cycle inputs in graduate medical education (GME) have expanded from traditional sources (eg, knowledge-based examinations, direct observation) to include data from the electronic health record (EHR), real-time location systems, and other technologies. Multimodal data capture may allow for a more complete picture of learners, their behaviors, and their environment to improve educational and patient outcomes. Methods for analyzing these data to generate insights are nascent.[1] This perspective discusses current challenges and opportunities in using data analytics to optimize clinical skills training (TABLE).

US residency programs rely on Clinical Competency Committees (CCCs) to make determinations regarding trainee advancement.[4,5] The cognitive load of the CCC can lead to decision fatigue, groupthink, and bias.[6] CCCs may have a narrow view into a trainee's performance, with at best a limited data set and at worst one that is misrepresentative. For example, internal medicine residents may spend little time in direct contact with patients,[7-10] which will lead to few opportunities for direct observation. Simulation is often used to address this issue,[11,12] but there is renewed interest in observations of real patient interactions.[13-17] One method is to capture frequent assessments of entrustable professional activities (EPAs) during patient care.[18-23] However, this data can be difficult to synthesize in a way that leads to actionable insights, and may be costly due to the need for faculty participation in data collection.[24] There is also no clear consensus on the common elements of clinical competence within and across specialties.

Passively acquired ("no-touch") data can increase the quantity of information about trainees.[13] Real-time tracking systems provide information about movement through the clinical environment, including data about inter-resident and service-based differences in clinical activities.[1,7,10] These data could help suggest improvements in trainee workflow or modifications to specific rotations to generate desired clinical experiences. EHR data includes how users interact with the EHR (metadata), clinical reasoning in notes, clinical exposure, prescribing habits, and patient outcomes.[25-27] Given that the attributions of tasks or clinical outcomes to an individual trainee can be challenging,[28] EHR-derived measures like Trainee Attributable & Automatable Care Evaluations in Real-time (TRACERs) might help more directly link outcomes to individual behaviors and provide feedback to improve performance. For example, TRACERs can give feedback on whether a trainee is ordering anti-hyperglycemic therapies based on current guidelines.[29] The promise of such approaches cannot be fully realized until medical informaticists find solutions to the lack of standardized data collection across different EHRs.[30-32] Common data languages standardize clinical data for research[33,34]; similar data dictionaries could help address this issue in medical education. A downside to grounding assessment in EHR metrics is that it might provide additional incentive for learners to focus on the EHR at the expense of time with patients.

Data analytics, including the use of artificial intelligence (AI), may help to generate insights from vast and disparate data sources. AI is starting to be used in assessment, clinical reasoning, teaching, and other activities.[35] AI could potentially help with data collection, personalized analytics, participatory interventions, and prediction of outcomes.[36] Natural language processing (NLP) can analyze narrative feedback about trainees to potentially improve evaluation processes while reducing administrative burden.[37] NLP can also evaluate trainees' notes for clinical reasoning.[38] Generative AI can incorporate multimodality inputs that may be too difficult for human evaluators to synthesize. This synthesis could guide educational interventions, such as suggesting learning content to review or assigning specific clinical encounters to address a gap in prior experience.[3] Additionally, the ability of

DOI: http://dx.doi.org/10.4300/JGME-D-24-00776.1

TABLE
Improving Data Analytics in Precision Education in Support of Clinical Skills Development

Activity	Description
Develop shared mental models and definitions of clinical competence	Agree on the common elements of clinical competence within and across specialties.
Leverage technology to aid in the collection and validation of assessment data	Utilize data from the EHR, RTLS, and other sources to broaden the quality and quantity of assessment data available about trainees. Study validity evidence for "no" and "low-touch" assessments of clinical skills including novel assessments and technology-based metrics (eg, TRACERs).
Incorporate more direct observation of trainees in the workplace	Incorporate more direct observation of learners in the workplace as part of a renewed emphasis on the importance of physicians spending time in direct contact with patients. Correlating direct observations with novel metrics might help to decrease the cost and burden of in-person observations.
Use data analytics to provide insights	Use AI to analyze large volumes of data to create meaningful insights about trainees. Issues related to data privacy, data access, cost of maintenance, and operation need to be fully explored.
Build trust in assessments	Obtain validity evidence for assessments for different settings and subjects. Ensure assessments are fit for purpose (ie, they capture what they intend to measure). Appropriately manage learner attribution and contribution. Maintain appropriate data privacy.
Develop interventions	Use insights to affect trainee learning and clinical outcomes. Measurements of interventions are critical to help drive this area forward and optimize both learning and patient care.
Safety in assessment	GME programs could use insights to develop individualized learning plans for trainees but must consider potential consequences of data access by fellowship programs, prospective employers, and patients.

Abbreviations: EHR, electronic health record; RTLS, real-time tracking systems; TRACERs, Trainee Attributable & Automatable Care Evaluations in Real-time; AI, artificial intelligence; GME, graduate medical education.

generative AI to develop learning cases at scale could revolutionize simulation and create new avenues for trainee assessment.[39]

Appropriate use of AI in medical education requires ethical frameworks, interdisciplinary collaborations, investment in education, promotion of transparency and accountability, and monitoring to evaluate impact. Key ethical principles include privacy, security, transparency, accountability, and fairness.[40] Data about trainees must be managed to protect learners as well as patient privacy. Challenges include determining who will have access to the insights generated from learner data and for what purpose. For example, should insights be shared with prospective employers or patients? The potential consequences of using GME data to inform subsequent training and practice need to be explored. If learners and assessors suspected that data would be shared outside of the educational environment, this could undermine the entire assessment system. The use of AI must be disclosed to trainees, and appropriate monitoring methods must be instituted. Additionally, AI-based analytics should complement bedside assessments, not replace them. The cost of data collection, storage, and third-party systems must be considered. These costs may be offset by improved clinical performance with reduced need for remediation, or by freeing up faculty time for other tasks.

Mistrust in learner assessments may be another barrier to PME implementation.[41] Interpretation of assessments is hampered by variation among programs in assessment methods used, reduced in-person feedback, and concerns about validity evidence.[42] Many clinicians are unaware of the use of EHR metadata to generate insights about their clinical performance. Some mistrust the data or fear its misuse by employers.[43] This highlights the need to accrue evidence of validity for new analytic techniques, including with direct observation of clinical behavior. The issue of mistrust also emphasizes the importance of engaging stakeholders in the analytic process. With the potential effects of PME on future training and employment, trainees should have a seat at the table when deciding on how data insights are used. One could imagine a hybrid system where GME program directors agree on a minimum specialty-specific standard for clinical competency and then, using personalized data, partner with trainees to identify individual professional goals.

Data analytics are only valuable if the generated insights provide actionable guidance to trainees. Currently that guidance is fragmented, in part reflecting the disorganized manner of data collection and synthesis.

Even high-quality, well-organized data can overwhelm trainees' efforts to make sense of feedback. One hopes that formal coaching programs that translate insights into developmental plans in partnership with trainees will result in a more satisfying professional experience and improvements in clinical skills.[44-47]

Similar to how the mapping of the human genome enabled the launch of the National Institutes of Health Precision Medicine Initiative a decade ago, we must articulate PME's key building blocks to realize its potential. Discussion and clarification about data analysis in the educational setting may enable innovators to define these building blocks. Development of shared data definitions, learner models, educational outcomes of interest, and guidelines for privacy and security are some of the first areas educators should tackle. EHR data and AI will likely play key roles in PME, but we must also prioritize obtaining validity evidence for assessments that incorporate direct observations of clinical skills. These efforts will be enhanced by engaging learners in this process; as stakeholders they can contribute unique perspectives on professional growth and effects on future career opportunities. Finally, as clinical competence is specialty-specific, development of PME tools will require collaboration across multiple specialties to ensure that analytic processes work across the continuum of medical education.

References

1. Garibaldi BT, Hollon MM, Woodworth GE, Winkel AF, Desai SV. Navigating the landscape of precision education: insights from on-the-ground initiatives. *Acad Med*. 2024;99(4 suppl 1):71-76. doi:10.1097/acm.0000000000005606
2. Triola MM, Burk-Rafel J. Precision medical education. *Acad Med*. 2023;98(7):775-781. doi:10.1097/acm.0000000000005227
3. Desai S. Precision medical education: the future of lifelong learning in medicine. *Acad Med*. 2024;99(4 suppl 1):14-20. doi:10.1097/ACM.0000000000005601
4. Green ML, Aagaard EM, Caverzagie KJ, et al. Charting the road to competence: developmental milestones for internal medicine residency training. *J Grad Med Educ*. 2009;1(1):5-20. doi:10.4300/01.01.0003
5. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system–rationale and benefits. *N Engl J Med*. 2012;366(11):1051-1056. doi:10.1056/NEJMsr1200117
6. Dickey CC, Thomas C, Feroze U, Nakshabandi F, Cannon B. Cognitive demands and bias: challenges facing clinical competency committees. *J Grad Med Educ*. 2017;9(2):162-164. doi:10.4300/jgme-d-16-00411.1
7. Rosen MA, Bertram AK, Tung M, Desai SV, Garibaldi BT. Use of a real-time locating system to assess internal medicine resident location and movement in the hospital. *JAMA Netw Open*. 2022;5(6):e2215885. doi:10.1001/jamanetworkopen.2022.15885
8. 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(16):1494-1508. doi:10.1056/NEJMoa1800965
9. Block L, Habicht R, Wu A, et al. In the wake of the 2003 and 2011 duty hours regulations, how do internal medicine interns spend their time? *J Gen Int Med*. 2013;28(8):1042-1047. doi:10.1007/s11606-013-2376-6
10. D'Souza T, Rosen M, Bertram AK, Apfel A, Desai SV, Garibaldi BT. Use of a real-time location system to understand resident location in an academic medical center. *J Grad Med Educ*. 2019;11(3):324-327. doi:10.4300/jgme-d-19-00026.1
11. Oman SP, Magdi Y, Simon LV. *Past Present and Future of Simulation in Internal Medicine*. StatPearls Publishing; 2023. https://www.ncbi.nlm.nih.gov/books/NBK549883/
12. Miller C, Toy S, Schwengel D, Isaac G, Schiavi A. Development of a simulated objective structured clinical exam for the APPLIED Certification Exam in Anesthesiology: a two-year experience informed by feedback from exam candidates. *J Educ Perioper Med*. 2019;21(4):e633.
13. Russell SW, Desai SV, O'Rourke P, et al. The genealogy of teaching clinical reasoning and diagnostic skill: the GEL Study. *Diagnosis (Berl)*. 2020;7(3):197-203. doi:10.1515/dx-2019-0107
14. Jacobsen AP, Khiew YC, Murphy SP, Lane CM, Garibaldi BT. The modern physical exam–a transatlantic perspective from the resident level. *Teach Learn Med*. 2020;32(4):442-448. doi:10.1080/10401334.2020.1724792
15. Norcini JJ, Blank LL, Arnold GK, Kimball HR. The mini-CEX (clinical evaluation exercise): a preliminary investigation. *Ann Intern Med*. 1995;123(10):795-799. doi:10.7326/0003-4819-123-10-199511150-00008
16. Clark BW, Lee YZJ, Niessen T, Desai SV, Garibaldi BT. Assessing physical examination skills using direct observation and volunteer patients. *Diagnosis (Berl)*. 2020;8(1):101-110. doi:10.1515/dx-2019-0089
17. Clark BW, Niessen T, Apfel A, et al. Relationship of physical examination technique to associated clinical skills: results from a direct observation assessment. *Am J Med*. 2022;135(6):775-782.e10. doi:10.1016/j.amjmed.2021.11.021
18. Pinilla S, Kyrou A, Klöppel S, Strik W, Nissen C, Huwendiek S. Workplace-based assessments of entrustable professional activities in a psychiatry core clerkship: an observational study. *BMC Med Educ*. 2021;21(1):223. doi:10.1186/s12909-021-02637-4
19. Stahl CC, Collins E, Jung SA, et al. Implementation of entrustable professional activities into a general surgery

residency. *J Surg Educ*. 2020;77(4):739-748. doi:10.1016/j.jsurg.2020.01.012
20. Kitto S, Fantaye AW, Zevin B, Fowler A, Sachdeva AK, Raiche I. A scoping review of the literature on entrustable professional activities in surgery residency programs. *J Surg Educ*. 2024;81(6):823-840. doi:10.1016/j.jsurg.2024.02.011
21. Woodworth GE, Marty AP, Tanaka PP, et al. Development and pilot testing of entrustable professional activities for US anesthesiology residency training. *Anesth Analg*. 2021;132(6):1579-1591. doi:10.1213/ane.0000000000005434
22. Caretta-Weyer HA, Gisondi MA. Design your clinical workplace to facilitate competency-based education. *West J Emerg Med*. 2019;20(4):651-653. doi:10.5811/westjem.2019.4.43216
23. Kinnear B, Santen SA, Schumacher DJ, et al. Using theory-informed group processes to make TIMELESS promotion decisions. *Acad Med*. 2024;99(1):28-34. doi:10.1097/acm.0000000000005445
24. Vennemeyer S, Kinnear B, Gao A, et al. User-centered evaluation and design recommendations for an internal medicine resident competency assessment dashboard. *Appl Clin Inform*. 2023;14(5):996-1007. doi:10.1055/s-0043-1777103
25. Adler-Milstein J, Adelman JS, Tai-Seale M, Patel VL, Dymek C. EHR audit logs: a new goldmine for health services research? *J Biomed Inform*. 2020;101:103343. doi:10.1016/j.jbi.2019.103343
26. Wright A, Pang J, Feblowitz JC, et al. A method and knowledge base for automated inference of patient problems from structured data in an electronic medical record. *J Am Med Inform Assoc*. 2011;18(6):859-867. doi:10.1136/amiajnl-2011-000121
27. Warm EJ, Kinnear B, Knopp MI, et al. Ambulatory long block: a model of precision education and assessment for internal medicine residents. *Acad Med*. 2024;99(4 suppl 1):57-63. doi:10.1097/ACM.0000000000005610
28. Schumacher DJ, Dornoff E, Carraccio C, et al. The power of contribution and attribution in assessing educational outcomes for individuals, teams, and programs. *Acad Med*. 2020;95(7):1014-1019. doi:10.1097/acm.0000000000003121
29. Burk-Rafel J, Sebok-Syer SS, Santen SA, et al. TRainee Attributable & Automatable Care Evaluations in Real-time (TRACERs): a scalable approach for linking education to patient care. *Perspect Med Educ*. 2023;12(1):149-159. doi:10.5334/pme.1013
30. Holmes JH, Beinlich J, Boland MR, et al. Why is the electronic health record so challenging for research and clinical care? *Methods Inf Med*. 2021;60(1-02):32-48. doi:10.1055/s-0041-1731784
31. Rule A, Melnick ER, Apathy NC. Using event logs to observe interactions with electronic health records: an updated scoping review shows increasing use of vendor-derived measures. *J Am Med Inform Assoc*. 2022;30(1):144-154. doi:10.1093/jamia/ocac177
32. Rule A, Kannampallil T, Hribar MR, et al. Guidance for reporting analyses of metadata on electronic health record use. *J Am Med Inform Assoc*. 2024;31(3):784-789. doi:10.1093/jamia/ocad254
33. Xiao X, Mehta HB, Curran J, Garibaldi BT, Alexander GC. Potential drug-drug interactions among U.S. adults treated with nirmatrelvir/ritonavir: a cross-sectional study of the National Covid Cohort Collaborative (N3C). *Pharmacotherapy*. 2023;43(12):1251-1261. doi:10.1002/phar.2860
34. Observational Health Data Sciences and Informatics. Standardized Data: The OMOP Common Data Model. Accessed February 2, 2025. https://www.ohdsi.org/data-standardization/
35. Gordon M, Daniel M, Ajiboye A, et al. A scoping review of artificial intelligence in medical education: BEME guide no. 84. *Med Teach*. 2024;46(4):446-470. doi:10.1080/0142159x.2024.2314198
36. Turner L, Hashimoto DA, Vasisht S, Schaye V. Demystifying AI: current state and future role in medical education assessment. *Acad Med*. 2024;99(4 suppl 1):42-47. doi:10.1097/acm.0000000000005598
37. Spadafore M, Yilmaz Y, Rally V, et al. Using natural language processing to evaluate the quality of supervisor narrative comments in competency-based medical education. *Acad Med*. 2024;99(5):534-540. doi:10.1097/acm.0000000000005634
38. NYU Langone Health. Artificial intelligence supercharges learning for students at NYU Grossman School of Medicine. Accessed February 8, 2025. https://nyulangone.org/news/artificial-intelligence-supercharges-learning-students-nyu-grossman-school-medicine
39. Sardesai N, Russo P, Martin J, Sardesai A. Utilizing generative conversational artificial intelligence to create simulated patient encounters: a pilot study for anaesthesia training. *Postgrad Med J*. 2024;100(1182):237-241. doi:10.1093/postmj/qgad137
40. Knopp MI, Warm EJ, Weber D, et al. AI-enabled medical education: threads of change, promising futures, and risky realities across four potential future worlds. *JMIR Med Educ*. 2023;9:e50373. doi:10.2196/50373
41. Caretta-Weyer HA, Smirnova A, Barone MA, et al. The next era of assessment: building a trustworthy assessment system. *Perspect Med Educ*. 2024;13(1):12-23. doi:10.5334/pme.1110
42. Hauer KE, Park YS, Bullock JL, Tekian A. "My assessments are biased!" Measurement and sociocultural approaches to achieve fairness in assessment in medical education. *Acad Med*. 2023;98(suppl 8):16-27. doi:10.1097/ACM.0000000000005245
43. Burden M, Keniston A, McBeth L, et al. Big brother or big opportunity? Utilization of EHR use metadata

in the inpatient setting: a mixed methods study at 16 organizations [published online ahead of print December 25, 2024]. *J Hosp Med*. doi:10.1002/jhm.13572

44. Palamara K, Chu JT, Chang Y, et al. Who benefits most? A multisite study of coaching and resident well-being. *J Gen Intern Med*. 2022;37(3):539-547. doi:10.1007/s11606-021-06903-5
45. Dyrbye LN, Shanafelt TD, Gill PR, Satele DV, West CP. Effect of a professional coaching intervention on the well-being and distress of physicians: a pilot randomized clinical trial. *JAMA Intern Med*. 2019;179(10):1406-1414. doi:10.1001/jamainternmed.2019.2425
46. Winkel AF, Gillespie C, Park A, et al. Bridging the gap from student to doctor: developing coaches for the transition to residency. *Med Educ Online*. 2023;28(1):2145103. doi:10.1080/10872981.2022.2145103
47. Winkel AF, Chang LY, McGlone P, Gillespie C, Triola M. SMARTer goalsetting: a pilot innovation for coaches during the transition to residency. *Acad Med*. 2023;98(5):585-589. doi:10.1097/acm.0000000000005153

Brian T. Garibaldi, MD, MEHP, is Director, Center for Bedside Medicine, and Charles Horace Mayo Professor of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA; **McKenzie Hollon, MD, FASE,** is an Associate Professor of Medicine and Vice Chair for Education, Department of Anesthesiology, Emory University School of Medicine, Atlanta, Georgia, USA; **Michelle I. Knopp, MD,** is an Assistant Professor, Department of Pediatrics and Department of Biostatistics, Health Informatics, and Data Sciences, University of Cincinnati College of Medicine, and a Member, Divisions of Biomedical Informatics and Hospital Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA; **Abigail Ford Winkel, MD, MHPE,** is a Professor and Vice Chair for Education, Department of Obstetrics and Gynecology, and Assistant Director, Institute for Innovations in Medical Education, New York University (NYU) Grossman School of Medicine, New York, New York, USA; **Jesse Burk-Rafel, MD, MRes,** is an Assistant Professor, Division of Hospital Medicine, and Director, Precision and Translation Medical Education Laboratory, Institute for Innovations in Medical Education, NYU Grossman School of Medicine, New York, New York, USA; and **Holly A. Caretta-Weyer, MD, MHPE,** is a Clinical Associate Professor, Department of Emergency Medicine, and Associate Dean of Admissions and Assessment, Stanford University School of Medicine, Palo Alto, California, USA.

Corresponding author: Brian T. Garibaldi, MD, MEHP, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA, brian.garibaldi@northwestern.edu