Academic Medicine

DOI: 10.1097/ACM.0000000000004165

Longitudinal Milestone Assessment Extending Through Subspecialty Training: The Relationship Between ACGME Internal Medicine Residency Milestones and Subsequent Pulmonary and Critical Care Fellowship Milestones

Janae K. Heath, MD, MS, Tisha Wang, MD, Lekshmi Santhosh, MD, MA, Josh L. Denson, MD, MS, Eric Holmboe, MD, Kenji Yamakazi, PhD, Alison S. Clay, MD, and W. Graham Carlos, MD

J.K. Heath is an assistant professor, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania; ORCID: https://orcid.org/0000-0002-0533-3088.

T. Wang is an associate professor, Department of Medicine, David Geffen School of Medicine at the University of California Los Angeles, Los Angeles, California.

L. Santhosh is an assistant professor, Department of Medicine, University of California San Francisco, San Francisco, California.

J.L. Denson is an assistant professor, Section of Pulmonary, Critical Care, and Environmental Medicine, Tulane University School of Medicine, New Orleans, Louisiana; ORCID: https://orcid.org/0000-0002-8654-7765.

E. Holmboe is an adjunct professor, Department of Medicine, Yale University, New Haven, Connecticut; and the Chief Research, Milestone Development, and Evaluation Officer for the Accreditation Council for Graduate Medical Education, Chicago, Illinois.

K. Yamakazi is the senior analyst for the Accreditation Council for Graduate Medical Education, Chicago, Illinois.

This is the author's manuscript of the article published in final edited form as:

Heath, J. K., Wang, T., Santhosh, L., Denson, J. L., Holmboe, E., Yamazaki, K., Clay, A. S., & Carlos, W. G. (2021). Longitudinal Milestone Assessment Extending Through Subspecialty Training: The Relationship Between ACGME Internal Medicine Residency Milestones and Subsequent Pulmonary and Critical Care Fellowship Milestones. Academic Medicine: Journal of the Association of American Medical Colleges, 96(11), 1603–1608. https://doi.org/10.1097/ACM.0000000000004165

A.S. Clay is an assistant professor, Department of Medicine, Duke University, Durham, North Carolina.

W.G. Carlos is an associate professor, Department of Medicine, Indiana University, Indianapolis, Indiana.

Correspondence should be addressed to Janae K. Heath; 3600 Spruce Street, 822 West Gates Building, Philadelphia, PA, 19104; telephone: (570) 594-1857; email: janae.heath@pennmedicine.upenn.edu; Twitter: @JanaeHeath1.

Supplemental digital content for this article is available at http://links.lww.com/ACADMED/B119, http://links.lww.com/ACADMED/B120, http://links.lww.com/ACADMED/B121, http://links.lww.com/ACADMED/B122, and http://links.lww.com/ACADMED/B123.

Funding/Support: None reported.

Other disclosures: None reported.

Ethical approval: This research was performed with the approval obtained from the institutional review board at the University of Pennsylvania.

Previous presentations: Findings have previously been presented virtually at the Association of Pulmonary and Critical Care Medicine Program Directors annual meeting on March 13, 2020.

Data: The data for the research were from the Accreditation Council for Graduate Medical Education, who performed all analysis for the study.

Abstract

Purpose

The Accreditation Council for Graduate Medical Education (ACGME) milestones were implemented across medical subspecialties in 2015. Although milestones were proposed as a longitudinal assessment tool potentially providing opportunities for early implementation of individualized learning plans in fellowship, the association of subspecialty fellowship ratings with prior residency ratings remains unclear. This study aimed to assess the relationship between internal medicine (IM) residency milestones and pulmonary-critical care medicine (PCCM) fellowship milestones.

Method

A multicenter retrospective cohort analysis was conducted for all PCCM trainees enrolled in ACGME-accredited PCCM fellowship programs in 2017–2018 who had complete prior IM milestone ratings from 2014–2017. Only professionalism and interpersonal and communication skills (ICS) were included based on shared anchors between IM and PCCM milestones. Using a generalized estimating equations model, the association of PCCM milestones ≤ 2.5 during the first year of fellowship with corresponding IM subcompetencies was assessed at each time-point, nested by program. Statistical significance was determined using logistic regression.

Results

The study included 354 unique PCCM fellows. For both ICS and professionalism subcompetencies, fellows with higher IM ratings were less likely to obtain PCCM ratings ≤ 2.5 during the first fellowship year. Each ICS subcompetency was significantly associated with future lapses in fellowship (ICS01: $\beta = -0.67$, $P = 0.003$; ICS02: $\beta = -0.70$, $P = 0.001$; ICS03: $\beta = -0.60$, $P = 0.004$) at various residency timepoints. Similar associations were noted for PROF03 ($\beta = -0.57$, $P = 0.007$).

Conclusions

Findings demonstrated an association between IM milestone ratings and low milestone ratings during PCCM fellowship. IM trainees with low ratings in several professionalism and ICS subcompetencies were more likely to be rated ≤ 2.5 during their first year in PCCM fellowship. This highlights a potential use of longitudinal milestones to target educational gaps at the beginning of PCCM fellowship.

In 2013, an innovative assessment system was developed to align with the framework of 6 general competencies originally developed by the Accreditation Council for Graduate Medical Education (ACGME) and American Board of Medical Specialties.[1–9] This new assessment system incorporated the concept of milestones that use developmental narratives to describe the professional progress of learners. This revised system introduced a criterion-based framework for trainee competency assessment along the 6 core competencies (patient care and procedural skills, medical knowledge, systems-based practice, practice-based learning and improvement, professionalism, and interpersonal and communication skills [ICS]). Overall, this system sought to provide a trajectory-based metric to demonstrate a trainee's progression toward competence and, ultimately, readiness for unsupervised practice. The milestones were introduced in the residency setting in 2013 and expanded to subspecialty training in 2015,[4,10–13] supported by diverse validity evidence across Messick's domain in a variety of general specialties, including correlation to other variables (such as American Board of Internal Medicine certification).[14–17] It has been hypothesized that the milestones may provide an opportunity to improve assessment across the residency to fellowship training periods, potentially allowing for an optimized educational transition and development of individualized learning plans for fellowships. To our knowledge, however, the association of these scores with fellowship training outcomes remains unknown.[18]

For internal medicine (IM) and its subsequent subspecialties, several of the milestones, specifically the domains of professionalism and ICS, consist of shared descriptions and behavioral anchors. Assuming an overall shared mental model of these domains between specialty and subspecialty training, we anticipate that these domains may share relatively consistent trends for an individual learner. More specifically, although the competencies represent context-dependent attributes, there may be context-independent associations between

lapses in residency training and ultimate lapses in fellowship training.[19–21] While it is not known if deficiencies within these milestones during IM training are associated with similar issues in specialty fellowship training, determining these associations could provide crucial data to further operationalize the milestones for individualized development. In fact, if such an association exists, knowledge of any issues during residency could allow for early adoption of an individualized learning plan for the fellow.

Therefore, we aimed to assess the relationship between pulmonary and critical care medicine (PCCM) subspecialty milestones and IM residency milestones. Given the shared descriptions and behavioral anchors used in the professionalism and interpersonal and communication skills domains, we were specifically interested in the association of low ratings within those domains during residency and during fellowship training. In other words, are lower rating in residency on professionalism predictive of future professionalism issues in fellowship training? We hypothesized that although the competencies represent context-dependent attributes, there may be context-independent associations between ratings in residency training and ultimate ratings in fellowship training. Determining such an association could provide additional validity evidence for the relationship of residency milestone ratings to subspecialty ratings, providing educators and training program directors the ability to optimally operationalize these metrics within the subspecialty setting and focus on specific competencies for individual trainees early in their fellowship training.

Method

Setting and participants

We performed a multicenter retrospective cohort analysis of all PCCM trainees enrolled in an ACGME-accredited PCCM fellowship program between 2017–2018. Our cohort was restricted to PCCM trainees (specifically excluding those trainees only specializing in critical care

medicine or pulmonary medicine), given the uncertain correlation of milestone ratings between these other subspecialties. However, the diversity of clinical practice (ambulatory, inpatient, consultant, intensive care), clinical acuity, and procedural opportunities within the field of PCCM should also provide generalizability to a number of other specialties.

Trainees were included only if they had completed a 3-year IM residency program in June 2017 (i.e., a group of residents from the academic year 2014–2017 cohort) to enrich our cohort with individuals with full IM residency milestone data available. We excluded any trainees who were not enrolled in an ACGME-accredited IM residency program during the 2014–2017 academic years. Additionally, trainees in separate pulmonary medicine or critical care medicine only training programs were excluded, as these training programs are distinct from the combined fellowship programs and not all critical care medicine trainees have previously completed an IM residency. Similarly, trainees enrolled in a combined medicine–pediatric training program prior to fellowship training were excluded given the unclear generalizability of the IM milestones in this group.

Complete milestone assessments of each fellow (each consisting of 24 sub-competencies, grouped within the 6 competency domains) were submitted to the ACGME by each institution at biannual intervals (December and June). Similar milestone assessments of these individuals were submitted across the 23 subcompetencies for IM training. Of note, details regarding institutional implementation of milestone ratings and/or rater training were not available. For the purposes of this study, we selected the Professionalism and ICS subcompetencies based on the identical nature of descriptions and behavioral anchors to the IM residency milestones, as each shared similar properties, both in terms of order and content, between the residency and subspecialty disciplines.[22] Furthermore, expertise in both of these domains has been linked to myriad positive patient outcomes, including improved patient satisfaction, adherence, and patient safety

metrics.[23–26] The professionalism competency consists of 4 subcompetencies, and the ICS competency consists of 3 subcompetencies (complete descriptions within each domain are included in Table 1).

After linking residency and subspecialty milestone data for each individual trainee, we de-identified all submitted milestone assessments on the trainee cohort to create the cohort for analysis. We analyzed the subcompetencies in the professionalism and ICS domains using a 6-point scale, consistent with prior literature analyzing milestones, ranging from Level 0 (i.e., critical deficiencies) to Level 5 (i.e., aspirational).

The institutional review board at the University of Pennsylvania approved the study.

Data analysis

We performed summary statistical analyses across all submitted evaluations. To assess the association of IM milestone ratings with PCCM milestone ratings of interest (professionalism and ICS domains), we applied generalized estimating equations (GEE) models[27] to determine the association between the exposure variable of IM residency subcompetency (i.e., PROF01, PROF02, etc.) and ultimate PCCM milestone ratings. In the GEE model, we used an exchangeable working correlation matrix to obtain adjusted regression coefficients and standard errors accounting for the IM by PCCM programs-clustering.

For the initial analysis, each of the PCCM subcompetencies was regressed on the corresponding subcompetency in the IM milestone. For example, we regressed the PROF01 of the PCCM milestone on the PROF01 of the IM milestone. The primary outcome of our study was a binary variable defined as any PCCM trainee with reporting milestone levels ≤ 2.5 for each of the specified subcompetencies at any point in the first year of PCCM subspecialty training (either the mid-year or end-of-year timepoints in fellowship training). The value of 2.5 was used as a cutoff in the fellowship outcomes based on 2 a priori observations. First, the IM milestones have a

unique design compared to other specialties in that level 1 equates to critical deficiencies and level 2 uses descriptors of negative behaviors. Second, we chose this cutoff specifically due to previously published data using positive predictive value cutoffs for IM residency. These prior studies highlighted levels ≤ 2.5 at any point in residency training being associated with at least a 30% risk of not achieving a graduation target in the residency population. Furthermore, a level 3 performance in the current milestone system indicates "competence" (in the Dreyfus model) as well as readiness to care for patients under mostly indirect or reactive supervision. Thus, we utilized a rating of 2.5 or less to capture a group of individuals who were deemed not ready for such indirect supervision. For the analysis, we included only data from the first clinical year of fellowship training, given the heterogeneity of subspecialty training pathways across PCCM fellowship programs following the initial clinical year of training.

Our exposure variable was IM residency milestone ratings. We first assessed these as a continuous variable (ranging from 0 to 5). For the primary analysis, we analyzed the association between IM residency milestone ratings and an outcome of PCCM ratings of ≤ 2.5 (dichotomized) using logistic regression. Statistical significance was determined using a *P* value of .05.

As a secondary analysis, we assessed the association of PCCM ratings of ≤ 2.5 during the 2017–2918 academic year with any IM milestone rating < level 2 at any of the 6 preceding IM assessment periods (dichotomized variable). Statistical significance was determined using logistic regression, with adjustment for multiple comparisons with Bonferroni correction (*P* value of .007).

To address the possibility that observations may have had higher correlation for trainees who were enrolled within the same PCCM program as well as had belonged to common IM program, we used an exchangeable working correlation matrix in each of the GEE analyses. Given the

potential for small numbers of residents in a given cluster, we performed 2 additional analyses to ensure the results were reproducible without evidence of overfitting the model. We first repeated the GEE model, instead clustering only on IM by PCCM programs, and finally repeated the analysis using a logistic regression model without considering data clustering.

We completed all statistical analyses using SAS Enterprise Guide version 7.15 (SAS Institute Inc., Cary, North Carolina).

Results

The cohort consisted of milestone data for 354 unique trainees with complete milestone data from the 2014–2017 academic years (for IM reporting milestones) and from the 2017–2018 academic year (for PCCM reporting milestones), out of 655 enrolled first-year categorical PCCM fellows in 2017–2018. This included data from 198 distinct IM programs, and 143 distinct PCCM fellowship programs. Sixty-five percent (n = 230) of the cohort were men, at a mean age of 30.9 years (+/- 2.3 years). Seventy-two percent of trainees (n = 253) changed institutions between residency and fellowship training. Of note, the majority of the 301 trainees excluded from analysis did not have complete 3-year reporting milestones available from IM training. Consistent with the expected age differences due to the timing of implementation of reporting milestones, the excluded group was older than those included in the analysis (mean age 32.8 versus 30.9, $P = .0001$). There were no other notable demographic differences, as outlined in Supplemental Digital Appendix 1, at http://links.lww.com/ACADMED/B119. In the cohort used for analysis, there were minimal missing values (< 2% of submitted milestones), which were imputed with the last value for analysis.

The milestone ratings across all individuals over the study period exhibited an upward trajectory spanning over the IM residency training (academic years 2014–2017) through the first year of PCCM fellowship (academic year 2017–2018) (Supplemental Digital Appendix 2, at

http://links.lww.com/ACADMED/B120). This trend was present for both those PCCM trainees with reporting milestones ≤ 2.5 as well as those with milestone ratings > 2.5 during the time period of the study.

Within the cohort, approximately one-third of PCCM fellows were rated at or below a level of 2.5 during the first year of subspecialty training across the 4 professionalism subcompetencies and 3 ICS subcompetencies.

IM subcompetency ratings were associated with a future rating at or below 2.5 during the first year of PCCM fellowship training (Supplemental Digital Appendix 3, at http://links.lww.com/ACADMED/B121). Overall, trainees with higher IM ratings were less likely to obtain PCCM ratings at or lower than level 2.5 during the first year in the fellowship training. This trend was apparent in all subcompetencies (across both professionalism and ICS). Within the professionalism competency, each of the 4 subcompetencies was associated with a rating at or below 2.5 during PCCM training (at one of the time points in IM residency). Additional analyses were performed without clustering, which showed similar results (see Supplemental Digital Appendix 4, at http://links.lww.com/ACADMED/B122).

Within the ICS domain, the subcompetency ICS02 ("Communicates effectively in interprofessional teams, e.g., with peers, consultants, nursing, ancillary professionals, and other support personnel") was significantly associated with future deficiencies in fellowship training. The remaining ICS subcompetencies (ICS01, "Communicates effectively with patients and caregivers," and ICS03, "Appropriate utilization and completion of health records") were also associated with future ratings at or below 2.5 in the same subcompetencies during fellowship.

The association between IM residency milestone ratings (dichotomized to identify ratings below level 2 at any time point during residency) and ultimate PCCM ratings is shown in Table 1. Table 1 includes the number of residents who were rated at or below level 2.0 at any point during

the course of IM training period and the associations between IM milestones ratings (specifically receiving a rating of < 2 at any timepoint) and PCCM milestone ratings (specifically rating ≤ 2.5).

These findings suggest that if trainees receive a milestone rating of at or below 2 for the ICS01 subcompetency ("Communicates effectively with patients and caregivers"), they were more likely to be rated at or below level 2.5 during the first year in their PCCM training (odd ratio = 2.50, 99% confidence interval = 1.13, 5.50). Similar trends were noted across the PROF01, PROF02, PROF3, and ICS02 domains (Table 1), although they were not statistically significant. To ensure our findings were not the result of overfitting from clustering, the analyses were performed without clustering using a simple logistic regression, which showed similar results (as outlined in Supplemental Digital Appendix 5, at http://links.lww.com/ACADMED/B123).

Discussion

In this multicenter retrospective cohort study, we found an association between ACGME IM residency milestone ratings and low milestone ratings during PCCM fellowship. Specifically, these data show that IM trainees with low ratings in several professionalism and ICS subcompetencies were more likely to be rated at or below 2.5 (indicating poor achievement) during their first year in a PCCM fellowship. These associations are further highlighted in the upward trend in trajectory when those individuals with low milestone ratings in residency are compared to those without low ratings. Additionally, IM residents rated at or below 2 at any point during residency in the ICS01 subcompetency ("Communicates effectively with patients and caregivers") were more likely to be rated at or below level 2.5 during the first year of their PCCM fellowship. Overall, we believe this study is the first to demonstrate an association between residency and fellowship milestone ratings, adding further validity evidence for the relatively new ACGME milestones. More importantly, this association provides some rationale

for creating individualized learning plans for fellows at the start of training, which may be directed by prior achievements and struggles in residency training.

For this study, we focused on the professionalism and ICS domains. Our findings suggest that low ratings in these areas during residency may indicate a greater propensity for such issues later in training. While these domains are likely interpreted in a context-dependent nature by raters (i.e., different skills within these domains assessed in fellowship as compared to residency), our findings suggest that there are likely context-independent associations between lapses in residency training and ultimate lapses in fellowship training. In fact, while individuals in this cohort achieved minimal competency in residency-specific milestones, there is potential for recurrent lapses in novel environments with unique stressors (in the fellowship context). This is consistent with prior evidence suggesting an association with residency professionalism lapses and future performance (i.e., disciplinary action by state licensing boards).[28–30] The association we noted between IM training and PCCM training highlights potential opportunities for educational handoff between residency and fellowship, as well as potential for the development of individualized learning plans targeting these issues.[31]

These findings not only provide additional validity evidence for the current IM milestones, but also suggest the need to utilize the milestones to implement individualized education plans within subspecialty training as early as possible. Recent data suggest predictable longitudinal trends in milestones assessments in residents within IM, emergency medicine, and family medicine.[20] In fact, early low ratings in residency training were predictive of an increased probability of not achieving designated graduation goals. Our study extends this finding, offering an opportunity for fellowship educators, program directors, and clinical competence committees to consider individualized learning plans early in fellowship using prior IM milestone data.

Currently subspecialty program directors receive residency milestone data from the ACGME for their matriculating fellowship trainees after they have begun the fellowship, however, it is unclear how these data are used in each fellowship program. Given the correlation between residency and fellowship ratings, our study promotes the idea that longitudinal milestones can potentially be used to target educational gaps and individualize training to ensure that fellows continue a trajectory of positive growth over the duration of their graduate medical education. Of note, while the IM milestones cannot be used for selection purposes, they can be used to further individual professional development. In fact, this highlights their utility upon entry to create an individualized learning plan upon matriculation into a PCCM program. Not only can the prior milestones serve as a helpful guide to the PCCM program director (and fellows) to address potential concerns and develop a plan to enhance areas of competency needing attention, they can also play a role in adapting the nature and timing of assessments performed.

Our study has several limitations. First, as each individual training program (both IM and PCCM) has a unique approach to the actions of its clinical competency committee and assignment of milestone ratings, we are unable to comment on the tools and decision making within these assessments. We have also presumed the validity of the IM milestone ratings, which we did not specifically assess with this study, but has been evaluated in other studies.[15,16] Additionally, we are not able to comment on each program's approach to struggling learners and remediation, which may have an impact on future milestone trajectories and potential fellowship lapses. Of note, some residents in this cohort were likely successfully remediated following a low rating in IM training, but this information could still indicate "at risk" individuals. This again would encourage use of these milestones for individualized learning plans for incoming fellows. We also recognize that there may be the risk of heightened scrutiny for fellowship trainees with prior professionalism or communication lapses from residency training. As our

study highlights the significant association with recurrent lapses, we feel such knowledge of prior lapses provides an important opportunity for creation of individualized learning plans (in novel and stressful environments prone to potential lapses) to optimize patient care and safety. Ultimately, further qualitative analysis of how fellowship programs currently use residency milestone data is an important area of future work as it remains unclear how these data are being currently used by fellowship training directors, if at all.

Our study was also limited by the relatively low number of individual trainees included in the final analysis (as we included individuals with complete milestone data across IM residency), limiting power for assessing the ultimate outcome. In fact, we conducted a post hoc power analysis, identifying that our cohort of 354 individuals did not achieve a power of 80% in most subcompetency domains. Despite this limitation, this study is still notable in analyzing the association between IM and PCCM milestones and finding a number of several interesting associations. We believe that the multicenter collection of standardized information by the ACGME significantly enhances the fidelity of the data included in this proof-of-concept study. Larger studies should be pursued in the future once the milestones have been in use for a longer period of time.

Another limitation to our study is generalizability. While we included trainees across all ACGME-accredited programs, our study focused on only 2 competencies (professionalism and ICS) within a single subspecialty, so the relevance to other milestones and other subspecialties is less clear. However, the consistency of the selected milestones between IM and PCCM training programs did allow for clearer and more consistent comparisons, which will remain relevant after the advent of the ACGME Harmonized milestones. Future work in other medical subspecialties is needed to ensure generalizability of our findings.

Our demonstration of an association between IM milestone ratings and subspecialty ratings, specifically within PCCM, warrants further research. Further studies of association between milestones in residency and subspecialty training (across specialties), particularly with the implementation of the upcoming milestones 2.0, are needed. Understanding these associations can help fellowship program leadership operationalize IM milestone data for their learners, perhaps providing targets for meaningful guidance and mentorship within the domains of communication and professionalism, and ultimate longitudinal assessment as fellows progress toward independence.

References

1. Batalden P, Leach D, Swing S, Dreyfus H, Dreyfus S. General competencies and accreditation in graduate medical education. Health Aff Proj Hope. 2002;21:103–111.
2. Holmboe ES, Batalden P. Achieving the desired transformation: Thoughts on next steps for outcomes-based medical education. Acad Med. 2015;90:1215–1223.
3. Swing SR. The ACGME Outcome Project: Retrospective and prospective. Med Teach. 2007;29:648–654.
4. Swing SR, Beeson MS, Carraccio C, et al. Educational milestone development in the first 7 specialties to enter the Next Accreditation System. J Grad Med Educ. 2013;5:98–106.
5. 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;1:5–20.
6. Nasca TJ, Philibert I, Brigham T, Flynn TC. The Next GME Accreditation System: Rationale and benefits. N Engl J Med. 2012;366:1051–1056.
7. Frank JR, Snell LS, ten Cate O, et al. Competency-based medical education: Theory to practice. Med Teach. 2010;32:638–645.
8. Frank JR, Mungroo R, Ahmad Y, Wang M, De Rossi S, Horsley T. Toward a definition of competency-based education in medicine: A systematic review of published definitions. Med Teach. 2010;32:631–637.
9. Iobst WF, Sherbino J, ten Cate O, et al. Competency-based medical education in postgraduate medical education. Med Teach. 2010;32:651–656.
10. Aagaard E, Kane GC, Conforti L, et al. Early feedback on the use of the internal medicine reporting milestones in assessment of resident performance. J Grad Med Educ. 2013;5:433–438.

11. Turner TL, Bhavaraju VL, Luciw-Dubas UA, et al. Validity evidence from ratings of pediatric interns and subinterns on a subset of pediatric milestones. Acad Med. 2017;92:809–819.

12. Warm EJ, Held JD, Hellmann M, et al. Entrusting observable practice activities and milestones over the 36 months of an internal medicine residency. Acad Med. 2016;91:1398–1405.

13. Fessler HE, Addrizzo-Harris D, Beck JM, et al. Entrustable professional activities and curricular milestones for fellowship training in pulmonary and critical care medicine: Report of a multisociety working group. CHEST. 2014;146:813–834.

14. Hauer KE, Clauser J, Lipner RS, et al. The internal medicine reporting milestones: Cross-sectional description of initial implementation in U.S. residency programs. Ann Intern Med. 2016;165:356–362.

15. Hauer KE, Vandergrift J, Lipner RS, Holmboe ES, Hood S, McDonald FS. National internal medicine milestone ratings: Validity evidence from longitudinal three-year follow-up. Acad Med. 2018;93:1189–1204. https://journals.lww.com/academicmedicine/Fulltext/2018/08000/National_Internal_Medicine_Milestone_Ratings_.34.aspx

16. Hauer KE, Vandergrift J, Hess B, et al. Correlations between ratings on the resident annual evaluation summary and the internal medicine milestones and association with ABIM certification examination scores among U.S. internal medicine residents, 2013–2014. JAMA. 2016;316:2253–2262.

17. Beeson MS, Holmboe ES, Korte RC, et al. Initial validity analysis of the emergency medicine milestones. Acad Emerg Med. 2015;22:838–844.

18. Li S-TT. The promise of milestones: Are they living up to our expectations? J Grad Med Educ. 2017;9:54–57.

19. Chan T, Sebok-Syer S, Thoma B, Wise A, Sherbino J, Pusic M. Learning analytics in medical education assessment: The past, the present, and the future. AEM Educ Train. 2018;2:178–187.

20. Holmboe ES, Yamazaki K, Nasca TJ, Hamstra SJ. Using longitudinal milestones data and learning analytics to facilitate the professional development of residents: Early lessons from three specialties. Acad Med. 2020;95:97–103.

21. Wancata LM, Morgan H, Sandhu G, Santen S, Hughes DT. Using the ACMGE milestones as a handover tool from medical school to surgery residency. J Surg Educ. 2017;74:519–529.

22. Holmboe ES, Edgar L, Hamstra SJ. The Milestones Guidebook. Published 2016. Accessed October 17, 2019. https://www.acgme.org/Portals/0/MilestonesGuidebook.pdf?ver=2016-05-31-113245-103. Accessed April 27, 2021.

23. Brennan MD, Monson V. Professionalism: Good for patients and health care organizations. Mayo Clin Proc. 2014;89:644–652.

24. O'Daniel M, Rosenstein AH. Professional communication and team collaboration. In: Hughes RG, ed. Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Advances in Patient Safety. Rockville, MD: Agency for Healthcare Research and Quality; 2008. http://www.ncbi.nlm.nih.gov/books/NBK2637/. Accessed April 27, 2021.

25. Markides M. The importance of good communication between patient and health professionals. J Pediatr Hematol Oncol. 2011;33(Suppl 2):S123–S125.

26. Burgener AM. Enhancing communication to improve patient safety and to increase patient satisfaction. Health Care Manag. 2017;36:238–243.

27. Hardin JW. Generalized estimating equations (GEE). In: Encyclopedia of Statistics in Behavioral Science. Hoboken, NJ: John Wiley & Sons; 2005.

28. Krupat E, Dienstag JL, Padrino SL, et al. Do professionalism lapses in medical school predict problems in residency and clinical practice? Acad Med. 2020;95:888–895.

29. Papadakis MA, Arnold GK, Blank LL, Holmboe ES, Lipner RS. Performance during internal medicine residency training and subsequent disciplinary action by state licensing boards. Ann Intern Med. 2008;148:869–876.

30. Lipner RS, Young A, Chaudhry HJ, Duhigg LM, Papadakis MA. Specialty certification status, performance ratings, and disciplinary actions of internal medicine residents. Acad Med. 2016;91:376–381.

31. Lucey C, Souba W. Perspective: The problem with the problem of professionalism. Acad Med. 2010;85:1018–1024.

Table 1

Association (OR) of Internal Medicine Milestones With Milestone Level 2.5 During the First Year of PCCM Using GEE Adjustment (Clustered on Trainee), From a Multi-Center Retrospective Study of Milestone Assessment During Residency and Subspecialty Training, 2017–2018[a]

Milestone	Subcompetency description	Fellows rated ≤ 2.0 during internal medicine residency, no. (%)	Association (OR) with milestone rating ≤ 2.0 during any point of internal medicine training (95% CI)	*P* value
PROF01	Has professional and respectful interactions with patients, caregivers, and members of the interprofessional team	41 (11.6)	1.48 (0.62, 3.55)	.23
PROF02	Accepts responsibility and follows through on tasks	60 (16.9)	1.47 (0.71, 3.03)	.15
PROF03	Responds to each patient's unique characteristics and needs	60 (16.9)	1.58 (0.87, 2.90)	.04
PROF04	Exhibits integrity and ethical behavior in professional conduct	45 (12.7)	0.96 (0.41, 2.29)	.91
ICS01	Communicates effectively with patients and caregivers	51 (14.4)	2.498 (1.13, 5.50)	.002[b]
ICS02	Communicates effectively in interprofessional teams	48 (13.6)	1.47 (0.65, 3.35)	.21
ICS03	Appropriate utilization and completion of health records	45 (12.7)	1.08 (0.47, 2.47)	.81

Abbreviations: OR, odds ratio; PCCM, pulmonary–critical care medicine; GEE, generalized estimating equations; CI, confidence interval; PROF, professionalism; ICS, interpersonal and communication skills.

[a]Table 1 displays the association between IM milestone rating of ≤ 2.0 (at any time point during residency, dichotomized variable) with a PCCM fellowship milestone rating of ≤ 2.0 during the first year of PCCM training (dichotomized outcome) using a GEE model (clustered on the trainee).

[b]Statistical significant associations as determined by a Bonferonni correction (*P* value = .007).