A collaboration between the *JAOA* and the American Association of Colleges of Osteopathic Medicine (AACOM) to recruit, peer review, publish, and distribute research and other scholarly articles related to osteopathic medical education.

***JAOA*/AACOM**

Assessment Considerations for Core Entrustable Professional Activities for Entering Residency

Machelle Linsenmeyer, EdD; Leslie Wimsatt, PhD; Mark Speicher, PhD; James Powers, DO; Stephen Miller, DO; Emmanuel Katsaros, DO

From the West Virginia School of Osteopathic Medicine in Lewisburg (Dr Linsenmeyer); Des Moines University College of Osteopathic Medicine in Iowa (Dr Wimsatt); Midwestern University Arizona College of Osteopathic Medicine in Glendale (Dr Speicher); Campbell University School of Osteopathic Medicine in Lillington, North Carolina (Dr Powers); Alabama College of Osteopathic Medicine in Dothan (Dr Miller); Western University of the Health Sciences College of Osteopathic Medicine of the Pacific in Pomona, California (Dr Katsaros).

Financial Disclosures: None reported.

Support: None reported.

Address correspondence to Machelle Linsenmeyer, EdD, West Virginia School of Osteopathic Medicine, 400 Lee St N, Lewisburg, WV 24901-1274.

Email: alinsenmeyer@osteo.wvsom.edu

Submitted November 15, 2017; revision received January 11, 2018; accepted January 24, 2018.

Context: In the process of analyzing entrustable professional activities (EPAs) for use in medical education, ten Cate and others identified challenges, including the need for valid and reliable EPA assessment strategies.

Objective: To provide osteopathic medical schools with a database of assessment tools compiled from the literature to assist them with the development and implementation of robust, evidence-based assessment methods.

Methods: MEDLINE, ERIC, PubMed, and other relevant databases were searched using MeSH keywords for articles outlining robust, evidence-based assessment tools that could be used in designing assessments for EPAs 1 through 6.

Results: A total of 55 publications were included in content analysis and reporting. All but 2 of the assessment articles were conducted in an undergraduate or graduate medical education setting. The majority of the 55 articles related to assessment of competencies affiliated with EPA 2 (16 articles) and EPA 4 (15 articles). Four articles focused on EPA 3.

Conclusion: Osteopathic medical schools can use this database of assessment tools to support the development of EPA-specific assessment plans that match the unique context and needs of their institution.

J Am Osteopath Assoc. 2018;118(4):243-251, e16-e21
doi:10.7556/jaoa.2018.049

Keywords: entrustable professional activities, EPA, medical education

In 2015, the American Association of Colleges of Osteopathic Medicine (AACOM) convened a panel of 13 osteopathic educators to serve on an Entrustable Professional Activities (EPAs) Steering Committee. The goal of the committee was to advance the implementation of EPAs within osteopathic medical education. The committee worked in collaboration with EPA liaisons from 43 osteopathic medical schools and branch campuses, a student representative from the Council of Osteopathic Student Government Presidents, and representatives from the Educational Council on Osteopathic Principles (ECOP) to advance its work nationally. Its first publication, *Osteopathic Considerations for Core Entrustable Professional Activities (EPAs) for*

Entering Residency,[1,2] described the distinct osteopathic clinical skills medical students must competently perform on the first day of residency. This reference[2] provided descriptions, guidelines, approaches, and the rationale for implementing the 13 Core EPAs using osteopathic approaches.

As a continuation of AACOM's support of osteopathic medical schools in the EPA movement, the EPA Steering Committee established 4 subcommittees in August 2016 to oversee the development of shared resources in several key areas: (1) faculty development; (2) curriculum; (3) instructional resources; and (4) assessment planning. This article focuses on the work of the EPA Assessment Planning Subcommittee to date.

Overview

Entrustable professional activities were introduced as a novel outcomes-based model for assessing essential, observable work tasks that define a specialty or profession.[3] They are described as tasks observed in daily practice rather than specific knowledge, skills, and attitudes assessed in isolation. For example, in an oral case presentation, a trainee may require excellent skills in eliciting a history, performing a physical examination, determining a differential diagnosis, developing a plan, and communicating with others. While each of these individual skills can be assessed, the task and how the various knowledge, skills, and so on are used can vary depending on the workplace context (eg, outpatient vs emergency department) and a number of other factors. The overall performance on the task and all affiliated competencies measured requires documentation in a variety of patient contexts and settings to qualify as an entrustable activity.

The key principles underlying EPA assessment and decision making (ie, workplace learning and trust) are generalizable to the continuum of physician training.[4] EPAs offer medical educators a workplace-relevant approach to the assessment of physicians-in-training. At the same time, the structuring of EPAs requires the development of new measures capable of addressing the underlying elements of entrustment decision making. Entrustment decisions must be based on clear definitions of trust, specific types and levels of targeted entrustments, and explicit definitions of what each development stage means in terms of moving students from novice learners to physicians capable of providing safe, unsupervised, and professional care.[4,5]

In the process of analyzing EPAs for use in medical education, researchers have identified several challenges, including the need for valid and reliable EPA assessment strategies.[6,7] Assessment guidelines outlined by ten Cate et al[5] call for the use of targeted assessment activities that are consistently observable, measurable, time-limited, and geared toward practice-relevant tasks suitable for entrustment. If assessments are not designed in this manner, validity and reliability may prove challenging.[8]

In this article, we provide an overview of existing assessment strategies that can be used as a basis for conceptualizing potential approaches to the assessment of EPAs in osteopathic medical education. The aim is to assist osteopathic medical schools with the development of robust, evidence-based assessment methods and encourage the design and implementation of innovations in EPA assessment.

Methods

During an initial meeting of the AACOM EPA Assessment Planning Subcommittee in August 2016, the group identified the lack of tools to support assessment planning as a critical gap in resources. Members felt that several types of information could be assembled to assist osteopathic medical schools with the development of robust, evidence-based assessment tools. The group agreed that some of the existing research on clinical assessment could prove valuable in supporting the design and customization of EPA assessment methods initiated across campuses. A decision was made to research the evidence in support of EPA-related assessment tool development and

implementation with a focus on EPA 1 through EPA 6 because a higher level of entrustment should be targeted for these EPAs before entry into clinical training for all medical students. The other EPAs, although important, may require a lower level of entrustment or be assessed on a smaller scale before entry into clinical training. Each subcommittee member investigated assessment approaches affiliated with competencies embedded within 2 different EPAs and worked with his or her campus librarian to conduct the search of major publication databases.

Three overarching questions were used to guide the literature review:

1. What research evidence exists to support the assessment of core competencies embedded within each of the first 6 EPAs?
2. What are the identified strengths and weaknesses of assessment instruments currently used to assess the competencies embedded in EPA 1 through EPA 6?
3. What opportunities exist for the future development of assessment tools that specifically target the measurement of EPA-related competency attainment?

Using advice from the academic librarians, the research team developed a strategy to search MEDLINE, ERIC, PubMed, and other relevant databases. MeSH keywords included sets of EPA-related terms such as "history," "medical history taking," "students, medical," and "educational measurement," among others (**eAppendix 1**). Articles with "Entrustable Professional Activities," "EPAs," or EPA-related competency wordings in the title were also included. All searches were conducted between August 2016 and March 2017. While the scope of the search was primarily focused on the most recent 10-year timeframe, several relevant articles that fell outside of this timeframe provided useful information related to EPA 2 and EPA 4 and were, therefore, included in the results. Articles were excluded from the review if the following was true: the article originated from a nonmedical field (eg, veterinary, pharmacy, dentistry), the full text was not accessible, the focus of the article was on how to design and implement an EPA rather than on how to assess learner entrustability on EPA-related competencies, the article assessed curricula or practitioners rather than learners, an English version of the article could not be obtained, it described a self-assessment solely (not including self-assessments used as part of a multisource feedback or objective structured clinical examination), or the article described the process used to develop the tool only. Also excluded from the review were editorials, commentaries, interviews, debates, and book reviews. The researchers then revisited the searches in January 2018 to update the findings before final publication.

The group systematically searched the key journals and publication databases. To ensure consistency and accuracy during the data collection process, the subcommittee developed a classifying system to extract the most relevant information from each article. The categories and categorical elements were selected from analysis of pertinent literature, conference presentations, committee discussion, etc. The classification system ultimately included the following categories: assessment category, EPA relevance, affiliated competency domains, level of learner, targeted specialty area(s), skills assessed, assessment type, feedback mechanisms, content validity, internal structure, and type of internal structure (**Figure**). Articles were not placed into the classification system/database until they had met the selection criteria.

Although most journals examined in this review clearly stated each article's primary topic in the title, abstract, or key words, every article was reviewed in detail by subcommittee members with expertise in medical education research. Each reviewer (M.L., L.W., M.S., J.P., S.M., and E.K.) undertook a primary review of competencies embedded in 2 different EPAs and independently coded the data using the classification categories. Each reviewer also served as a secondary reviewer and cross-checked the codings of the primary reviewer. Any disagreements were resolved by discussion, consensus, and consultation with a third

1. Assessment category (checklist, rating sheet, OSCE, 360° evaluation, portfolio, formative, summative)
2. EPA relevance (EPA 1, EPA 2, EPA 3, EPA 4, EPA 5, EPA 6, EPA 7, EPA 8, EPA 9, EPA 10, EPA 11, EPA 12, EPA 13)
3. Competency domains (osteopathic principles and practice, medical knowledge, patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice)
4. Level of learner (UME, GME, other)
5. Targeted specialty area(s) (anesthesia, cardiology, emergency medicine, family medicine, geriatrics, intensive care, internal medicine, neurology, obstetrics/gynecology, OMM, ophthalmology, orthopedics, pediatrics, plastic surgery, psychiatry, quality improvement, sports medicine, surgery, urology, other)
6. Skill(s) assessed (clinical examination skills; critical appraisal of literature; data gathering, diagnostic reasoning, decision making, clinical judgement; handover skill; humanism [eg, empathy]; interviewing skills; medical record keeping; oral case presentation; patient admission; postoperative management; preoperative management; procedural skill [other than surgical, eg, thoracoscopy]; skill in communication, counseling, obtaining consent; surgical skills and performance [micro, robotic, laparoscopic]; team work [multidisciplinary, multiprofessional]; time management; other)
7. Type of assessment tool (case-based discussion [eg, exploring adaptive competence]; knowledge tests [eg, written, multiple choice tests, oral examinations]; long practice observation [including multisource feedback]; procedural short practice observation; product evaluation [such as review of medical records]; rubric; short practice observation [including mini-clinical evaluation exercises]; simulation [including standardized patients and robotic or technological simulations])
8. Feedback mechanism (quantitative written, qualitative written, qualitative oral)
9. Content validity (yes, no)
10. Internal structure (yes, no)
11. Type of internal structure (Cronbach α, factor analysis, generalizability, interitem correlation, interrater reliability, Kuder-Richardson, variance, other)

Figure.
Classification system to extract the most relevant information from each article on entrustable professional activity (EPA) assessment. *Abbreviations:* GME, graduate medical education; OMM, osteopathic manipulative medicine; OSCE, objective structured clinical examination; UME, undergraduate medical education.

member of the review team. Finally, the researchers perused the Association of American Medical Colleges' *Toolkits for the 13 Core EPAs*[9] to ensure that key assessment tools were not missed in this compilation.

Results

A total of 55 articles were included in the content analyses and reporting. All but 1 of the articles were conducted in an undergraduate or graduate medical education setting. Although our review focused on empirical articles, important conceptual and theoretical findings were included in the reporting to enhance understanding of the theoretical and conceptual frameworks that form the basis for the development of EPA assessment tools.

The majority of articles were related to the assessment of competencies affiliated with EPA 2 (16 articles [29.1%]) and EPA 4 (15 articles [27.3%]). Four articles (7.3%) focused on EPA 3–related competency assessment, 6 (10.9%) on EPA 1, 8 (14.5%) on EPA 5, and 6 (10.9%) on EPA 6. A majority of the articles targeted the assessment of medical students (34 [61.8%]). In contrast, 12 (21.8%) related to the assessment of residents, 7 (12.7%) of both medical students and residents, and 2 (3.6%) did not specify the learner level.

All but 3 articles focused on 2 or more competency areas. Learner competency related to patient care was

Table 1.
Entrustment Scales for UME, GME, and Beyond[a]

Level of Supervision	UME	GME and Beyond	
	Chen et al[4]	ten Cate et al[5]	Ottawa Scale[10]
1	Not allowed to practice EPA	Is present and observes	"I had to do"
2	Allowed to practice EPA only under proactive, full supervision	Acts with direct supervision	"I had to talk them through"
3	Allowed to practice EPA only under reactive/on-demand supervision	Acts with indirect supervision	"I had to prompt them from time to time"
4	Allowed to practice EPA unsupervised	Acts without supervision	"I needed to be in the room just in case"
5	Allowed to supervise others in practice of EPA	Provides supervision	"I did not need to be there"

[a] The UME levels were proposed by Chen et al,[4] whereas the scales for GME and beyond are currently used in evaluating learners.

Abbreviations: EPA, entrustable professional activity; GME, graduate medical education; UME, undergraduate medical education.

the focus of assessment in all of the articles except 3. At least half of the articles targeted interpersonal and communication skills and/or medical knowledge, and slightly fewer than half targeted practice-based learning and improvement. The 2 competencies addressed the least were systems-based practice and professionalism.

All but 3 articles identified the type of assessment tool used to measure learner competency. Assessment tools most frequently reported in the literature were knowledge tests (n=21), simulation (n=16), and assessment rubrics (n=11). In contrast, product evaluations (n=7), long practice observations (n=3), short practice observations (n=2), and case-based discussions (n=2) were least often identified. The reviewers noted inconsistencies in the reporting of feedback to learners and inconsistencies in the provision of reliability and validity evidence across articles. In most instances, it was not clear how (or whether) postassessment feedback was provided to learners. When the feedback method was mentioned, written feedback was the most commonly used approach. Twenty-three articles included a description of how instrument reliability was determined, and 15 provided validity evidence.

Although the type of response scale was not itemized in our final report, we noted that response scales embedded with the various assessment tools varied considerably, within a wide range of behavioral anchors, dichotomous measures (eg, yes/no; performed/didn't perform), and agreement, satisfaction, and/or confidence scales represented. This finding highlighted to the subcommittee the importance of selecting appropriate response scales when the focus of assessment is intended as an entrustment decision. Entrustment scales will be needed to frame assessment results in terms of the type and designated level of supervision required. In this sense, the language of entrustment scale construction represents a shift in approach from competency-based "person-descriptors" to EPA-based "work-descriptors" as described by ten Cate et al,[6] who suggested using 5 levels of supervision to reflect increasing trust in trainee autonomy.

The subcommittee encountered several entrustment and co-activity scales commonly found in the undergraduate and graduate medical education literature (Table 1) but not yet widely reported in research on EPA-related competency assessment.[4,5,10] Additionally, the toolkits document[9] provides a thorough review of

scales and suggested modifications to scales to better fit the undergraduate medical education system. As mentioned previously, the content and format of EPAs represent a distinct shift toward more authentic assessment of work-based skills. During the design and development phase, the existing literature can be used to help map out EPAs as shown in **Table 2** using an EPA description and EPA matrix.[6,11] A detailed description of all 55 articles included in this review are shown in **eAppendix 2**.

Discussion

This literature review provides insight into key structural elements and processes identified as important in the assessment of EPA-related competencies. Taken together, the results indicate that several challenges should be taken into account when designing and assessing EPAs:

- **Rater variability and reliability:** There are many factors that can influence faculty ratings of learner performance (eg, comparisons with other learners, different ideas as to what a numerical value represents, faculty member's own skills and experiences, context, time, limitations in cognitive processes).[7,12-19] These challenges are not unique to the assessment of EPAs. However, educators must remain mindful that EPA assessment scales need to be designed as descriptors, narratives, or "entrustability scales" that help guide evaluators and lead to improvements in interrater agreement.[20] We acknowledge that sometimes variability and different perspectives can be deemed positive.
- **Clearly defined faculty roles:** Defining "appropriate" levels of learner supervision, ensuring quality of care delivered to patients, and safeguarding learner development can pose challenges if faculty roles are not clearly defined. Faculty must understand and be able to ensure learner competence while also ensuring patient safety. Faculty must also be able to address the fine line between quality of care and challenging learners to improve.[5,21,22] It is highly dependent on faculty skills in assessment and feedback.[7]
- **Assessment should be holistic:** When assessing an EPA, faculty need to have a broad picture of the learner that includes competencies and milestones. This approach requires mapping EPAs to these broader elements to develop a shared mental model for entrustment. This approach takes time and buy-in, and it can vary by institutional mission, resource capacity, curricular structure, and programming support for faculty development, among other factors.[7]
- **Time for assessment:** The time available to faculty for engagement in assessment planning, development, and implementation is limited. It can affect the quality of assessment, feedback, role modeling, and education provided to students.[23-27] To avoid delay in the delivery of feedback to learners, EPA assessment methods must be both feasible and useful to the stakeholders involved.[23,28]
- **Big data and longitudinal assessment:** Effective EPA assessment requires the scheduling of numerous observations conducted by multiple raters over time, with follow-up triangulation of the findings undertaken by an individual or group (eg, assessment committee).[29] To this end, a large volume of data must be collected across the learning continuum to limit the impact of contextual variations on overall trainee assessment.[20]
- **Assessment cannot be standardized:** Effective EPA assessment relies on the independent observations of numerous medical professionals under widely varying conditions. The aim is to assess how learners can perform in a variety of educational environments/clinical settings as they work with diverse patient populations and fluctuating supervisory expectations.
- **Assessments should collectively measure all aspects of "entrustability":** Assessment outcomes will be used to confirm not only the ability, but also the right and duty, of a trainee to act unsupervised.[5]

Many experts advocate for matrix-type assessment planning that maps EPAs to competency frameworks

Table 2.
Components of a Fully Described EPA[6]

Component	Description
1. Title of the EPA	Should be concise and informative (ie, readily understood). As it only reflects work, it should not be stated as a learning objective or skill, merely as an activity. Limit to ≤10 words. Use neutral infinitive tense to avoid the association with individuals (eg, "discharging patients" instead of "discharges a patient").
2. Specification and limitations	Should clearly list what is and is not included, given the level of the intended trainees. Include the context and targeted transition (eg, entering residency, fellowship, autonomous practice).
3. Most relevant domains of competence	Should relate the EPA to the competency framework used. Those domains of competencies or competencies of the framework that are most applicable may be mentioned.
4. Required experience, knowledge, skills, attitude, and behavior	Trainees should be aware what knowledge, skills, and attitudes are expected before they can be trusted to carry out the EPA to help them prepare for entrustment. It may also be helpful to understand which workplace experiences are considered necessary before entrustment (type of rotation, type of patients, number of procedures).
5. Assessment information sources to assess progress and ground a summative entrustment decision	Supervisors should be aware of which sources of information should be used to determine progress. Sources can be observed behavior or skill at the bedside or at morning report meetings; a skills test; information from colleagues, nurses, and patients; a double-checked procedure; a case-based discussion; and other sources. For trainees as well as supervisors, it is important to state how many times an EPA or its constituent parts must have been observed to enable taking a summative entrustment decision, and to state who takes the decision. It is highly recommended that multiple staff members sign off on such decisions. Supervisors should feel personal responsibility for these important decisions.
6. Entrustment for which level of supervision is to be reached at which stage of training?	The consequence of an entrustment decision is stated as the permission to act under a designated level of supervision (eg, indirect supervision, distant supervision) not generally permitted before that time. Next, it is necessary to state at which transition of training trainees must ultimately master the EPA at the designated level. Graduation should require that all core EPAs of the program be mastered. When building an individual workplace curriculum, it is useful to estimate when this trainee is expected to receive the entrustment decision, based on prior training and expected rotations and experiences.
7. Expiration date	Optional but recommended. Entrustment should drop if no maintenance of competence for this EPA happens, for example, over 1 to 5 years, depending on the EPA. Revalidation may require marginal or a more substantive check.

Source: Reproduced with permission from the Association for Medical Education in Europe. Modified from ten Cate O, Chen HC, Hoff RG, Peters H, Bok H, van der Schaaf M. Curriculum development for the workplace using Entrustable Professional Activities (EPAs): AMEE Guide No. 99. *Med Teach.* 2015;37 (11):983-1002. doi:10.3109/0142159X.2015.1060308

and milestones (matched to appropriate levels of entrustment).[6,30] Narrative statements included as descriptors can also be used to guide decisions about entrustment level. These approaches can serve as a mechanism for feedback and reporting by providing a common language or map for discussing learner performance and progress over time.

We noted several limitations related to this review. For example, literature related to the assessment of EPAs is still in its infancy, making it difficult to support findings with previously published research. The lack of an established set of EPA-related search terms mandated the use of a relatively complex set of search algorithms, which may require refinement as research in this area advances. While a wide range of articles were located, the diversity of study populations and methodologies, represented in the various articles, made the meaningful comparison of instrument psychometric properties difficult. The results of this review may be outdated quickly as assessment of EPAs progress and more articles are published. Therefore, additional searches will be necessary to

locate emergent findings. It is also possible that some exclusion criteria (especially limiting to the medical field) may have resulted in oversight of relevant assessments.

Future EPA assessment subcommittee research will focus on assessment tools affiliated with EPA 7 through EPA 13 and refining search terms to better identify articles (eg, filtering out articles related to specific tests and finding articles related to the ability to recommend or interpret appropriate tests for EPA 3). The committee will develop similar MeSH terms for EPA 7 through EPA 13. The MeSH terms developed and used in this review (**eAppendix 1**) can be applicable to institutions interested in continuing to search and develop EPA-related assessment tools going forward. A continued goal of the subcommittee research will be to create practical tools for use by medical schools interested in developing assessments that uniquely align with their organizational contexts and settings.

This article provides a unique frame of reference by summarizing the limited research literature to date. Given that the core set of medical schools charged with EPA piloting have not yet completed their initial piloting process nor extensively published on EPA assessment tool development, it is the hope of the subcommittee that this information will provide a useful benchmark against which future EPA assessment research and reporting can be measured.

Conclusion

A review of the extant literature over the past year allowed the AACOM EPA Assessment Planning Subcommittee to denote key gaps in the existing literature, as well as identify several considerations worth keeping in mind when designing assessments for use in entrustment decision making. In addition to the current findings, educators should note that institutional context will be an important covariant in determining the effectiveness of EPA assessment in osteopathic medical education because each institution has a different mission, curricular resources, and support structures. Therefore, this article may assist institutions by providing a baseline set of assessment tools that can be used to guide the development of EPA-specific assessment plans. Each medical school is urged to consider how the tools presented in this article can be adapted to meet the needs of its learners. The featured articles can also be reframed to address key entrustment decisions using deliberately assigned levels of supervision. Measures should address not only ability, but also integrity, reliability, and humility.[31] Medical educators may find it useful to implement several of the entrustment scales (**Table 1**)[4,5] and EPA planning tools (**Table 2**)[6,11] as a way of ensuring that all major decision points and assessment components are taken into consideration. By developing multiple measures and making consistent use of entrustment scales, instruments will be better able to document evidence of progress toward summative entrustment decisions.

Acknowledgment

We acknowledge Stephen C. Shannon, DO, MPH, president of the American Association of Colleges of Osteopathic Medicine, for his vision and support in advancing the EPA initiative, as well as Mary Essig, MLS, library director at the West Virginia School of Osteopathic Medicine in Lewisburg; Rebecca Hines, MLKS, education librarian and assistant professor at Des Moines University in Iowa; and Kelli Hines, MLIS, librarian at Western University of Health Sciences in Pomona, California, for their insights, advice, and efforts in the literature search process.

Author Contributions

All authors provided substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; all authors drafted the article or revised it critically for important intellectual content; all authors gave final approval of the version of the article to be published; and all authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. American Association of Colleges of Osteopathic Medicine *Osteopathic Considerations for Core Entrustable Professional Activities (EPAs) for Entering Residency*. Chevy Chase, MD: American Association of Colleges of Osteopathic Medicine; 2016. http://www.aacom.org/docs/default-source/med-ed-presentations/core-epas.pdf?sfvrsn=20. Accessed February 27, 2018.

2. Basehore PM, Mortensen LH, Katsaros E, et al. Entrustable Professional Activities for entering residency: establishing common osteopathic performance standards in the transition from medical school to residency. *J Am Osteopath Assoc*. 2017;117(11):712-718. doi:10.7556/jaoa.2017.137

3. ten Cate O. Entrustability of professional activities and competency-based training. *Med Educ*. 2005;39(12):1176-1177. doi:10.1111/j.1365-2929.2005.02341.x

4. Chen HC, van den Broek WE, ten Cate O. The case for use of entrustable professional activities in undergraduate medical education. *Acad Med*. 2015;90(4):431-436. doi:10.1097/ACM.0000000000000586

5. ten Cate O Hart D, Ankel F, et al. Entrustment decision making in clinical training. *Acad Med*. 2016;91(2):191-198. doi:10.1097/ACM.0000000000001044

6. ten Cate O, Chen HC, Hoff RG, Peters H, Bok H, van der Schaaf M. Curriculum development for the workplace using entrustable professional activities (EPAs): AMEE Guide No. 99. *Med Teach*. 2015;37(11):983-1002. doi:10.3109/0142159X.2015.1060308

7. Carraccio C, Englander R, Holmboe ES, Kogan JR. Driving care quality: aligning trainee assessment and supervision through practical application of entrustable professional activities, competencies, and milestones. *Acad Med*. 2016;91(2):199-203. doi:10.1097/ACM.0000000000000985

8. ten Cate O. Trusting graduates to enter residency: what does it take? *J Grad Med Educ*. 2014;6(1):7-10. doi:10.4300/JGME-D-13-00436.1

9. Obeso V, Brown D, Phillipi C, et al, eds. *Core Entrustable Professional Activities for Entering Residency: Toolkits for the 13 Core EPAs*. Washington, DC: Association of American Medical Colleges; 2017. https://www.aamc.org/download/482214/data/epa13toolkit.pdf. Accessed February 26, 2018.

10. Rekman J, Hamstra SJ, Dudek N, Wood T, Seabrook C, Gofton W. A new instrument for assessing resident competence in surgical clinic: the Ottawa Clinic assessment tool. *J Surg Ed*. 2016;73(4):575-582. doi:10.1016/j.jsurg.2016.02.003

11. Chen HC, McNamara M, Teherani A, ten Cate O, O'Sullivan P. Developing entrustable professional activities for entry into clerkship. *Acad Med*. 2016;91(12):247-255. doi:10.1097/ACM.0000000000000988

12. Govaerts MJB, van der Vleuten CPM, Schuwirth LWT, Muijtjens AMM. Broadening perspectives on clinical performance assessment: Rethinking the nature of in-training assessment. *Adv Health Sci Educ Theory Pract*. 2007;12(2):239-260. doi:10.1007/s10459-006-9043-1

13. Govaerts MJ, Van de Wiel MW, Schuwirth LW, Van der Vleuten CP, Muijtjens AM. Workplace-based assessment: raters' performance theories and constructs. *Adv Health Sci Educ Theory Pract*. 2013;18:375-396. doi:10.1007/s10459-012-9376-x

14. Kogan JR, Conforti L, Bernabeo E, Iobst W, Holmboe E. Opening the black box of clinical skills assessment via observation: a conceptual model. *Med Educ*. 2011;45(10):1048-1060. doi:10.1111/j.1365-2923.2011.04025.x

15. Gingerich A, Regehr G, Eva KW. Rater-based assessments as social judgments: rethinking the etiology of rater errors. *Acad Med*. 2011;86(10 suppl):S1-S7. doi:10.1097/ACM.0b013e31822a6cf8

16. Gingerich A, Kogan J, Yeates P, Govaerts M, Holmboe E. Seeing the "black box" differently: assessor cognition from three research perspectives. *Med Educ*. 2014;48(11):1055-1068. doi:10.1111/medu.12546

17. Yeates P, O'Neill P, Mann K, Eva KW. Effect of exposure to good vs poor medical trainee performance on attending physician ratings of subsequent performances [published correction appears in *JAMA*. 2013;309(3):237]. *JAMA*. 2012;308:2226-2232. doi:10.1001/jama.2012.36515

18. Govaerts MJ, Schuwirth LW, Van der Vleuten CP, Muijtjens AM. Workplace-based assessment: effects of rater expertise. *Adv Health Sci Educ Theory Pract*. 2011;16(2):151-165. doi:10.1007/s10459-010-9250-7

19. van der Vleuten CP. When I say … context specificity. *Med Educ*. 2014;48(3):234-235. doi:10.1111/medu.12263

20. Rekman J, Gofton W, Dudek N, Goften T, Hamstra SJ. Entrustability scales: outlining their usefulness for competency-based clinical assessment. *Acad Med*. 2016;91(2):186-190. doi:10.1097/ACM.0000000000001045

21. Kennedy TJT, Lingard L, Baker GR, Kitchen L, Regehr G. Clinical oversight: conceptualizing the relationship between supervision and safety. *J Gen Intern Med*. 2007;22(6):1080-1085. doi:10.1007/s11606-007-0179-3

22. Teman NR, Gauger PG, Mullan PB, Tarpley JL, Minter RM. Entrustment of general surgery residents in the operating room: factors contributing to provision of resident autonomy. *J Am Coll Surg*. 2014;219(4):778-787. doi:10.1016/j.jamcollsurg.2014.04.019

23. Klamen DL, Williams RG, Roberts N, Cianciolo AT. Competencies, milestones, and EPAs—are those who ignore the past condemned to repeat it? *Med Teach*. 2016;28(9):904-910. doi:10.3109/0142159X.2015.1132831

24. Chisholm CD, Whenmouth LF, Daly EA, Cordell WH, Giles BK, Brizendine EJ. An evaluation of emergency medicine resident interaction time with faculty in different teaching venues. *Acad Emerg Med*. 2004;11(2):149-155.

25. Han H, Roberts NK, Korte R. Learning in the real place: medical students' learning and socialization in clerkships at one medical school. *Acad Med*. 2015; 90(2):231-239. doi:10.1097/ACM.0000000000000544

26. Williams RG, Dunnington GL. Assessing the ACGME competencies with methods that improve the quality of evidence and adequacy of sampling. *ACGME Bull*. 2006;19:38-42.

27. van Loon KA, Teunissen PW, Driessen EW, Scheele F. The role of generic competencies in the entrustment of professional activities: a nationwide competency-based curriculum assessed. *J Grad Med Educ*. 2016;8(4):546-552. doi:10.4300/JGME-D-15-00321.1

28. Williams RG, Chen X, Sanfey H, Markwell S, Mellinger JD, Dunnington G. The measured effect of delay in completing operative performance ratings on clarity and detail of ratings assigned. *J Surg Edu*. 2014;71(6):e132-e138.

29. Williams RG, Dunnington GL, Mellinger JD, Klamen DL. Placing constraints on the use of the ACGME milestones: a commentary on the limitations of global performance ratings. *Acad Med*. 2015;90(4):404-407. doi:10.1097/ACM.0000000000000507

30. Pangaro L, ten Cate O. Frameworks for learner assessment in medicine: AMEE Guide No. 78. *Med Teach*. 2013;35:e1197-e1210. doi:10.3109/0142159X.2013.788789

31. ten Cate O. Entrustment as assessment: recognizing the ability, the right, and the duty to act. *J Grad Med Educ*. 2016;8(2):261-262. doi:10.4300/JGME-D-16-00097.1

eAppendix 1.

MeSH (Medical Subject Heading) terms: strategies to guide entrustable professional activity (EPA) assessment tool searches for EPAs 1 through 6. [a]EPA 3 will take more time because diagnostic tests comprise a broad category.

EPA 1

("history and physical"[tw] OR "physical exam"[tw] OR "physical examination"[tw] OR "physical examination"[MeSH Terms] OR "history taking"[tw] OR "medical history taking"[MeSH Terms] OR "interviews as topic"[MeSH Terms]) AND ("Schools, Medical"[MeSH] OR "Students, Medical"[MeSH] OR "Education, Medical, Undergraduate"[MeSH] OR "Clinical Clerkship"[MeSH]) AND ("Curriculum/standards"[MeSH] OR "Educational Measurement"[MeSH] OR instrument[tw] OR rubric[All Fields] OR tool[tw] OR "checklist"[MeSH Terms] OR "checklist"[All Fields]) AND ("Validation Studies"[Publication Type] OR reliability[tw] OR validate[tw] OR validity[tw] OR validation[tw] OR "Reproducibility of Results"[MeSH Terms])

EPA 2

("diagnosis"[MeSH] OR "Diagnosis, Differential"[MeSH] OR diagnosis[tw] OR "Clinical Examination") AND ("Schools, Medical"[MeSH] OR "Students, Medical"[MeSH] OR "Education, Medical, Undergraduate"[MeSH] OR "Clinical Clerkship"[MeSH]) AND ("Curriculum/standards"[MeSH] OR "Educational Measurement"[MeSH] OR instrument[tw] OR rubric OR tool[tw] OR checklist) AND ("Validation Studies" [Publication Type] OR reliability[tw] OR validate[tw] OR Validity[tw] OR validation[tw] OR "Reproducibility of Results"[MeSH Terms])

("Diagnosis, Differential"[MeSH] OR diagnosis[tw] OR "Clinical Examination") AND ("Schools, Medical"[MeSH] OR "Students, Medical"[MeSH] OR "Education, Medical, Undergraduate"[MeSH] OR "Clinical Clerkship"[MeSH]) AND ("Curriculum/standards"[MeSH] OR "Educational Measurement"[MeSH] OR instrument[tw] OR rubric OR tool[tw] OR checklist) AND ("Validation Studies" [Publication Type] OR reliability[tw] OR validate[tw] OR validity[tw] OR validation[tw] OR "Reproducibility of Results"[MeSH Terms])

EPA 3[a]

"Diagnostic Techniques and Procedures"[MeSH] AND ("Schools, Medical"[MeSH] OR "Students, Medical"[MeSH] OR "Education, Medical, Undergraduate"[MeSH] OR "Clinical Clerkship"[MeSH]) AND ("Curriculum/standards"[MeSH] OR "Educational Measurement"[MeSH] OR instrument[tw] OR rubric OR tool[tw] OR checklist) AND ("Validation Studies" [Publication Type] OR reliability[tw] OR validate[tw] OR validity[tw] OR validation[tw] OR "Reproducibility of Results"[MeSH Terms]) AND "humans"[MeSH Terms]

EPA 4

(prescribing[tw] OR "Drug Therapy"[MeSH] OR "Medication Errors"[MeSH] or pharmacology[tw] OR "Pharmacology, Clinical"[MeSH] OR "Pharmacology/education"[MeSH]) AND ("Schools, Medical"[MeSH] OR "Students, Medical"[MeSH] OR "Education, Medical, Undergraduate"[MeSH] OR "Clinical Clerkship"[MeSH]) AND ("Curriculum/standards"[MeSH] OR "Educational Measurement"[MeSH] OR instrument[tw] OR rubric OR tool[tw] OR checklist) AND ("Validation Studies" [Publication Type] OR validate[tw] OR validity[tw] OR validation[tw] OR reliability[tw] OR "Reproducibility of Results"[MeSH Terms])

EPA 5

(Documentation[tw] OR "Health record"[tw] OR "Medical record"[tw] OR "soap note"[tw] OR "computerized provider order entry") AND ("Schools, Medical"[MeSH] OR "Students, Medical"[MeSH] OR "Education, Medical, Undergraduate"[MeSH] OR "Clinical Clerkship"[MeSH]) AND ("Curriculum/standards"[MeSH] OR "Educational Measurement"[MeSH] OR instrument[tw] OR rubric OR tool[tw OR checklist]) AND ("Validation Studies" [Publication Type] OR validate[tw] OR validity[tw] OR validation[tw] OR reliability[tw] OR "Reproducibility of Results"[MeSH Terms])

EPA 6

("case presentation" OR "oral presentation" OR "patient presentation") AND ("Schools, Medical"[MeSH] OR "Students, Medical"[MeSH] OR "Education, Medical, Undergraduate"[MeSH] OR "Clinical Clerkship"[MeSH]) AND ("Curriculum/standards"[MeSH] OR "Educational Measurement"[MeSH] OR instrument[tw] OR rubric OR tool[tw] OR checklist) AND ("Validation Studies" [Publication Type] OR validate[tw] OR validity[tw] OR validation[tw] OR reliability[tw] OR "Reproducibility of Results"[MeSH Terms])

eAppendix 2.

Articles to support entrustable professional activity (EPA) assessment tool development (as of January 2018) by EPA. [a]Brinkman et al mention these measures but no results were shown to verify reliability. *Abbreviations:* GME, graduate medical education; NA, not available/applicable; MSF, multisource feedback; NI, not included; OSCE, objective structured clinical examination; SOAP, subjective, objective, assessment, plan; SP, standardized patient; UME, undergraduate medical education.

EPA 1: Gather a history and perform a physical examination											
Assessment Category	Citation	EPA Relevance	Competency Domains	Level of Learner	Targeted Specialty Area(s)	Skill(s) Assessed	Type of Assessment Tool	Feedback Mechanism	Content Validity	Internal Structure	Type of Internal Structure
Rating sheet	Bynum et al. Teaching medical students the art of the 'write-up.' *Clin Teach*. 2015;12(4):246-249.	1	Patient care, interpersonal and communication skills	UME	NI	Data gathering, diagnostic reasoning, decision making, clinical judgement, medical record keeping	Rubric	NI	No	Yes	Cronbach α, interrater reliability
Multiple/ systematic review	Keifenheim et al. Teaching history taking to medical students: a systematic review. *BMC Med Educ*. 2015;15:159.	1	Patient care, interpersonal and communication skills, practice-based learning	UME	NI	Humanism, interviewing, communication, counseling, obtaining consent	Knowledge tests (written, multiple choice, oral), short practice observation (including mini-clinical evaluation exercises), simulation (including SPs and robotic or technological simulations)	Quantitative written, qualitative written, qualitative oral	NA	NA	NA
Rating sheet	King et al. Developing validity evidence for the written pediatric history and physical exam evaluation rubric. *Acad Pediatr*. 2017;17(1):68-73.	1, 2, 3, 5	Medical knowledge, patient care, interpersonal and communication skills, professionalism	UME	Pediatrics	Data gathering, diagnostic reasoning, decision making, clinical judgement, medical record keeping, communication, counseling, obtaining consent	Rubric	Quantitative written	Yes	Yes	Cronbach α, Factor analysis, inter-item correlation
Rating sheet	Middleman et al. Reliability of the history and physical assessment (HAPA) form. *Clin Teach*. 2011;8(3):192-195.	1	Patient care, interpersonal and communication skills	UME	Pediatrics	Data gathering, diagnostic reasoning, decision making, clinical judgement, medical record keeping, communication, counseling, obtaining consent	Rubric	NI	No	Yes	Interrater reliability
Rating sheet, OSCE, formative	Perkowski. Critical synthesis package: mini-clinical evaluation exercise (mCEX). *MedEdPORTAL Publications*. 2014;10:9793.	1, 2	Patient care, interpersonal and communication skills	UME GME	Internal medicine	Clinical examination, data gathering, diagnostic reasoning, decision making, clinical judgement, humanism, interviewing, communication, counseling, obtaining consent, other (organization/efficiency; clinical competence)	Rubric	Quantitative written	Yes	Yes	Cronbach α, factor analysis, inter-item correlation, interrater reliability
Multiple/ systematic review	Setyonugroho et al. Reliability and validity of OSCE checklists used to assess the communication skills of undergraduate medical students: a systematic review. *Patient Educ Couns*. 2015;98(12):1482-1491.	1, 2, 3, 5	Patient care, interpersonal and communication skills, practice-based learning	UME	NI	Interviewing, communication, counseling, obtaining consent	Rubric, simulation (including SPs and robotic or technological simulations)	NI	Yes	Yes	NI

EPA 2: Prioritize a differential diagnosis following a clinical encounter											
Assessment Category	Citation	EPA Relevance	Competency Domains	Level of Learner	Targeted Specialty Area(s)	Skill(s) Assessed	Type of Assessment Tool	Feedback Mechanism	Content Validity	Internal Structure	Type of Internal Structure
Rating sheet, OSCE	Berger et al. Assessment of medical student clinical reasoning by "lay" vs physician raters: inter-rater reliability using a scoring guide in a multidisciplinary objective structured clinical examination. *Am J Surg*. 2012;203(1):81-86.	2	Medical knowledge, patient care	UME	NI	Clinical examination skills, diagnostic reasoning, decision making, clinical judgement	Simulation (including SPs and robotic or technological simulations)	NI	No	Yes	Cronbach α, interrater reliability
Rating sheet, 360° evaluation, formative	Borman et al. Initial performance of a modified milestones global evaluation tool for semiannual evaluation of residents by faculty. *J Surg Educ*. 2013;70(6):739-749.	1, 2, 8, 9, 12, 13	Medical knowledge, patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice	GME	Surgery	Critical appraisal of the literature, diagnostic reasoning, decision making, clinical judgment, handover, pre- and postoperative management, procedural (other than surgical, eg thoracoscopy), surgical skills and performance (micro, robotic, laparoscopic), teamwork	Long practice observation (including MSF)	NI	No	No	NA
Rating sheet, summative	Carrière et al. Assessing clinical reasoning in pediatric emergency medicine: validity evidence for a Script Concordance test. *Ann Emerg Med*. 2009;53(5):647-652.	2	Medical knowledge, patient care	GME	Pediatrics, emergency medicine	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice oral)	NI	Yes	Yes	Cronbach α, variance
Review paper, Rating sheet, formative, summative	Charlin et al. Standardized assessment of reasoning in contexts of uncertainty: the Script Concordance approach. *Eval Health Prof*. 2004;27(3):304-319.	2	Medical knowledge, patient care	Not specified	Family medicine, gynecology, radiology, rheumatology, urology	Clinical examination, diagnostic reasoning, decision making, clinical judgement	Short practice observation (including mini-clinical evaluation exercises)	NA	NA	NA	NA
OSCE, summative	Falcone et al. Using elements from an acute abdominal pain Objective Structured Clinical Examination (OSCE) leads to more standardized grading in the surgical clerkship for third-year medical students. *J Surg Educ*. 2011;68(5):408-413.	2	Medical knowledge, patient care	UME	Surgery	Clinical examination, diagnostic reasoning, decision making, clinical judgement, interviewing	Simulation (including SPs and robotic or technological simulations)	NI	No	No	NA

EPA 2 (continued): Prioritize a differential diagnosis following a clinical encounter											
Assessment Category	Citation	EPA Relevance	Competency Domains	Level of Learner	Targeted Specialty Area(s)	Skill(s) Assessed	Type of Assessment Tool	Feedback Mechanism	Content Validity	Internal Structure	Type of Internal Structure
OSCE, summative	Falcone et al. Differential diagnosis in a 3-station acute abdominal pain objective structured clinical examination (OSCE): a needs assessment in third-year medical student performance and summative evaluation in the surgical clerkship. *J Surg Educ*. 2011;68(4):266-269.	2	Medical knowledge, patient care	UME	Surgery	Diagnostic reasoning, decision making, clinical judgement	Simulation (including SPs and robotic or technological simulations)	NI	No	No	NA
Review Paper, formative	Farmer et al. A practical guide to assessing clinical decision-making skills using the key features approach. *Med Educ*. 2005;39(12):1188-1194.	2	Medical knowledge, patient care	UME	NI	Diagnostic reasoning, decision making, clinical judgement	Case-based discussion (eg, exploring adaptive competence), knowledge tests (written, multiple choice, oral)	NA	NA	NA	NA
Rating sheet	Humbert et al. Assessing clinical reasoning skills in scenarios of uncertainty: convergent validity for a Script Concordance test in an emergency medicine clerkship and residency. *Acad Emerg Med*. 2011;18(6):627-634.	2	Medical knowledge, patient care	UME, GME	Emergency medicine	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice, oral)	NI	Yes	Yes	Cronbach α, variance
Rating sheet	Humbert et al. Assessment of clinical reasoning: a Script Concordance test designed for pre-clinical medical students. *Med Teach*. 2011;33(6):472-477.	2	Medical knowledge, patient care	UME	NI	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice, oral)	NI	Yes	Yes	Cronbach α, variance
Assessment Guide, Rating sheet	Lubarsky et al. Script Concordance testing: from theory to practice: AMEE guide no. 75. *Med Teach*. 2013;35(3):184-193.	2	Medical knowledge, patient care	UME	NI	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice, oral)	NA	NA	NA	NA
Rating sheet	Meterissian et al. Is the Script Concordance test a valid instrument for assessment of intraoperative decision-making skills? *Am J Surg*. 2007;193(2):248-251.	2	Medical knowledge, patient care	GME	Surgery	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice, oral)	NI	Yes	Yes	Cronbach α, variance

EPA 2 (continued): Prioritize a differential diagnosis following a clinical encounter											
Assessment Category	Citation	EPA Relevance	Competency Domains	Level of Learner	Targeted Specialty Area(s)	Skill(s) Assessed	Type of Assessment Tool	Feedback Mechanism	Content Validity	Internal Structure	Type of Internal Structure
Rating sheet, summative	Page et al. Developing key-feature problems and examinations to assess clinical decision-making skills. *Acad Med*. 1995;70(3):194-201.	2	Medical knowledge, patient care	UME	NI	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice, oral)	NI	No	No	NA
Rating sheet	Petrucci et al. Assessing clinical judgment using the Script Concordance test: the importance of using specialty-specific experts to develop the scoring key. *Am J Surg*. 2013;205(2):137-140.	2	Medical knowledge, patient care	GME	Surgery	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice, oral)	NI	Yes	Yes	Cronbach α, variance
Review Paper, Formative, summative	Russell. Critical synthesis package: script concordance testing (SCT). *MedEdPORTAL Publications*. 2013;9:9492.	2	Medical knowledge, patient care	UME GME	Gynecology, radiology, surgery	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice, oral)	NA	NA	NA	NA
Review Paper, Formative, summative	Schultz et al. The application of entrustable professional activities to inform competency decisions in a family medicine residency program. *Acad Med*. 2015;90(7):888-897.	2	Medical knowledge, patient care	GME	Family medicine	Clinical examination, data gathering, diagnostic reasoning, decision making, clinical judgement, handover, medical record keeping, procedural skill (other than surgical, eg thoracoscopy), communication, counseling, obtaining consent	Rubric	NA	NA	NA	NA
Rating sheet	Sibert et al. Online clinical reasoning assessment with the Script Concordance test: a feasibility study. *BMC Med Inform Decis Mak*. 2005;5:18.	2	Medical knowledge, patient care	UME GME	Urology	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice, oral)	NI	No	No	NA

EPA 3: Recommend and interpret common diagnostic and screening tests											
Assessment Category	Citation	EPA Relevance	Competency Domains	Level of Learner	Targeted Specialty Area(s)	Skill(s) Assessed	Type of Assessment Tool	Feedback Mechanism	Content Validity	Internal Structure	Type of Internal Structure
Formative	Hartman et al. A novel tool for assessment of emergency medicine resident skill in determining diagnosis and management for emergent electrocardiograms: a multicenter study. *J Emerg Med*. 2016;51(6):697-704.	3	Medical knowledge	GME	Emergency medicine	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice oral)	NI	Yes	No	NA
NA	Jablonover et al. ECG as an Entrustable Professional Activity: CDIM Survey Results, ECG teaching and assessment in the third year. *Am J Med*. 2016:129(2):226-230.	3	Medical knowledge, patient care	UME	Internal medicine	Diagnostic reasoning, decision making, clinical judgement	NA	NI	No	No	NA
Checklist	Kreiter et al. A report on the piloting of a novel computer-based medical case simulation for teaching and formative assessment of diagnostic laboratory testing. *Med Educ Online*. 2011;16:5646.	3	Medical knowledge	UME	NI	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice oral)	Quantitative written	No	Yes	Cronbach α, classical test theory, generaliz-ability study, decision study
Checklist, OSCE, Formative	Schmidt et al. Competency assessment in senior emergency medicine residents for core ultrasound skills. *Western J Emer Med*. 2015;16(6):923-926.	3	Medical knowledge, patient care	GME	Emergency medicine	Clinical examination skills, Data gathering, Diagnostic reasoning, decision making	Knowledge tests (written, multiple choice, oral), simulation (including SPs and robotic or technological simulations)	NI	No	No	NA

EPA 4: Enter and discuss orders and prescriptions											
Assessment Category	Citation	EPA Relevance	Competency Domains	Level of Learner	Targeted Specialty Area(s)	Skill(s) Assessed	Type of Assessment Tool	Feedback Mechanism	Content Validity	Internal Structure	Type of Internal Structure
OSCE, Formative	Brinkman et al. Self-reported confidence in prescribing skills correlates poorly with assessed competence in fourth year medical students. *J Clin Pharmacol.* 2015;55(7):825-830.[a]	4	Medical knowledge, patient care, interpersonal and communication skills, practice-based learning and improvement, systems-based practice	UME	NI	Diagnostic reasoning; decision making; clinical judgement; Skill in communication, counseling, obtaining consent	Simulation (including SPs and robotic or technological simulations)	NI	No	Yes	Cronbach α, inter-item correlation, interrater reliability[a]
360° Evaluation, summative, multiple types and areas of assessment	British Pharmacological Society. MSC assessment. prescribing safety assessment: assessment blueprint. 2015;December 10. https://prescribingsafetyassessment.ac.uk/ resource/PSA-Blueprint-November-2017-final/pdf	4	Patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice	UME	Geriatrics, obstetrics/ gynecology, pediatrics, psychiatry, surgery, medicine + general practice	Clinical examination, critical appraisal of literature, data gathering, diagnostic reasoning, decision making, clinical judgement, medical record keeping, pre- and postoperative management, communication, counseling, obtaining consent	Knowledge tests (written, multiple choice, oral)	NI	No	No	NA
360° Evaluation, formative, new patient forms	Davis et al. Assessment of medical student clinical competencies in the neurology clinic. *Neurology.* 2007;68:597-599.	4	Medical knowledge, patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice	UME	Neurology	Clinical examination skills, diagnostic reasoning, decision making, clinical judgement, medical record keeping, oral case presentation	Product Evaluation (such as review of medical records)	Qualitative written qualitative oral	No	No	NA
Checklist, Rating sheet, OSCE, formative	Dwyer et al. Cognitive and psychomotor entrustable professional activities: can simulators help assess competency in trainees? *Clin Orthop Relat Res.* 2016;474(4):926-934.	4	Patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice	GME	Orthopedics	Pre- and postoperative management; diagnostic reasoning, decision making, clinical judgment; surgical skills and performance (micro, robotic, laparoscopic), communication, counseling, obtaining consent	Simulation (including SPs and robotic or technological simulations)	Qualitative written qualitative oral	Yes	Yes	Cronbach α, inter-item correlation, interrater reliability
Checklist, formative, written OSCE + extended matching questions	Harding et al. The performance of junior doctors in applying clinical pharmacology knowledge and prescribing skills to standardized clinical cases. *Br J Clin Pharmacol.* 2010;69:6):598-606.	4	Medical knowledge	GME	NI	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice oral), short-practice observation (including mini-clinical evaluation exercises)	NI	Yes	Yes	Cronbach α

EPA 4 (continued): Enter and discuss orders and prescriptions											
NA	Harendza et al. Ordering patterns for laboratory and radiology tests by students from different undergraduate medical curricula. *BMC Med Educ.* 2013;13:109-113.	4	Patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice	UME	NI	Diagnostic reasoning, decision making, clinical judgement	NI	NI	Yes	No	NA
summative	Jamshed et al. Understanding of antibiotic use and resistance among final-year pharmacy and medical students: a pilot study. *J Infect Dev Ctries.* 2014;8(6):780-785.	4	Medical knowledge, patient care, professionalism, practice-based learning and improvement, systems-based practice	UME, other, pharm-acy students	NI	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (such as written, multiple choice tests, or oral examinations)	NI	Yes	NA	NA
Checklist, formative, clinical vignettes (Fig 1) that required the students to complete simulated online order forms	Marshall et al. Informatics in radiology: evaluation of an e-learning platform for teaching medical students competency in ordering radiologic examinations. *Radiographics.* 2011;31(5):1463-1474.	4	Patient care, professionalism, practice-based learning and improvement, systems-based practice	UME	Radiology	Diagnostic reasoning, decision making, clinical judgement, medical record keeping	Product evaluation (such as review of medical records), simulation (including SPs and robotic or technological simulations)	Quantitative written	No	No	NA
Checklist, OSCE, formative	McEvoy et al. Are fourth-year medical students as prepared to manage unstable patients as they are to manage stable patients? *Acad Med.* 2014;89(4):618-624.	4, 10	Patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice	UME	Emergency medicine	Diagnostic reasoning, decision making, clinical judgement, clinical examination skills, interviewing skills, data gathering, teamwork (multi-disciplinary, multi-professional)	Simulation (including SPs and robotic or technological simulations)	Quantitative written	No	Yes	Variance, compared SP's grading results of the practice simulations with videos of those sessions in order
Checklist, Rating sheet, OSCE, portfolio, formative, multi-source assessment	Murphy et al. Insightful Practice: a robust measure of medical students' professional response to feedback on their performance. *BMC Med Educ.* 2015;15:125.	4	Patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice,	UME	NI	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice oral), simulation (including SPs and robotic or technological simulations)	Quantitative written qualitative oral	No	Yes	Cronbach α, inter-item correlation, Interrater reliability, variance, multiple sources of assessment data used

EPA 4 (continued): Enter and discuss orders and prescriptions											
OSCE, Difficult to tell whether Rating sheet or checklist	Scobie et al. Meeting the challenge of prescribing and administering medicines safely: structured teaching and assessment for final year medical students. *Med Educ.* 2003;37:434-437.	4	Patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice	UME	NI	Diagnostic reasoning, decision making, clinical judgement, counseling, communication, obtaining consent	Product evaluation (eg, review of medical records), simulation (including SPs and robotic or technological simulations)	NI	No	No	NA
Rating sheet, formative	Thomas et al. A simple intervention raised resident-physician willingness to assist transgender patients seeking hormone therapy. *Endocr Pract.* 2015;21(10):1134-1142.	4	Patient care, professionalism, practice-based learning and improvement, systems-based practice	GME	Family medicine, internal medicine	Diagnostic reasoning, decision making, clinical judgement	Rubric	NI	No	No	NA
Rating sheet, formative, summative	Wan. Using the script concordance test to assess clinical reasoning skills in undergraduate and postgraduate medicine. *Hong Kong Med J.* 2015;21:455-461.	4	Patient care, practice-based learning and improvement, systems-based practice	UME, GME	NI	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice, oral)	Quantitative written qualitative oral	No	No	NA
Rating sheet, summative	Wu et al. Development and validation of the McMaster Prescribing Competency Assessment for Medical Trainees (MacPCA). *J Popul Ther Clin Pharmacol.* 2015;22(2):e173-178.	4	Patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice	UME	NI	Diagnostic reasoning, decision making, clinical judgement	Knowledge tests (written, multiple choice, oral)	NI	Yes	No	NA
Summative, case write-ups - clerkship	Yaman et al. Evaluation of case write-up: assessment of prescription writing skills of fifth year medical students at UKM Medical Centre. *Procedia. Social Behav Sci.* 2012;60:249-253.	4	Patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, systems-based practice	UME	Obstetrics/ gynecology, pediatrics, surgery	Diagnostic reasoning, decision making, clinical judgement	Case-based discussion (eg, exploring adaptive competence), product evaluation (eg, review of medical records)	NI	No	No	NA

EPA 5: Document a clinical encounter in the patient record											
Assessment Category	**Citation**	**EPA Relevance**	**Competency Domains**	**Level of Learner**	**Targeted Specialty Area(s)**	**Skill(s) Assessed**	**Type of Assessment Tool**	**Feedback Mechanism**	**Content Validity**	**Internal Structure**	**Type of Internal Structure**
Formative, other, chart recall + case-based discussions	Al-Wassia et al. Using patients' charts to assess medical trainees in the workplace: a systematic review. *Med Teach.* 2015;37:S82-S87.	5	NI	GME	Psychiatry, surgery, primary care + oncology	Medical record keeping, oral case presentation	NI	NI	No	Yes	Cronbach α, interrater reliability (types of reporting varied across the 7 articles included)
Rating sheet, formative	Baker et al. The IDEA assessment tool: assessing the reporting, diagnostic reasoning, and decision-making skills demonstrated in medical students' hospital admission notes. *Teach Learn Med.* 2015;27(2):163-173.	5	Medical knowledge, patient care, interpersonal and communication skills	UME	Internal medicine	Clinical examination, diagnostic reasoning, decision making, clinical judgement, interviewing, medical record keeping, patient admission	Product evaluation (eg, review of medical records)	Quantitative Written	Yes	Yes	Cronbach α, factor analysis, inter-item correlation, interrater reliability, variance
Formative, summative, assessment rubric	Carter et al. Document a clinical encounter in the patient record development of an assessment rubric.	5	Patient care, interpersonal and communication skills, professionalism, systems based practice	UME	NI	Data gathering, diagnostic reasoning, decision making, clinical judgement, medical record keeping, communication, counseling, obtaining consent	Rubric	NI	No	No	NA
Rating sheet, portfolio	Dean et al. The effectiveness of a bundled intervention to improve resident progress notes in an electronic health record. *J Hosp Med.* 2015;10(2):104-107.	5	Patient care, interpersonal and communication skills	GME	Pediatrics	Medical record keeping	Product evaluation (eg, review of medical records)	NI	No	Yes	Inter-item correlation, interrater reliability
Checklist, formative	Milano et al. Simulated electronic health record (Sim-EHR) curriculum: teaching EHR skills and use of the EHR for disease management and prevention. *Acad Med.* 2014;89(3):399-403.	5	Medical knowledge, patient care, practice-based learning and improvement, systems-based practice	UME, GME	Family medicine, internal medicine	Medical record keeping	Simulation (including SPs and robotic or technological simulations)	Quantitative written	No	No	NA

EPA 5 (continued): Document a clinical encounter in the patient record											
Formative, other, SOAP note review	Seo et al. A pilot study on the evaluation of medical student documentation: assessment of SOAP notes. *Korean J Med Educ.* 2016;28(2):237-241.	5	Patient care, interpersonal and communication skills	UME	NI	Medical record keeping	Simulation (including SPs and robotic or technological simulations), product evaluation (eg, review of medical records)	NI	No	Yes	Cronbach α, inter-item correlation, Interrater reliability
Formative; other, SOAP note review	Stephens et al. Teaching principles of practice management and electronic medical record clinical documentation to third-year medical students. *J Med Pract Manage.* 2010;25(4):222-225.	5	Patient care, interpersonal and communication skills, practice based learning and improvement	UME	Family medicine	Medical record keeping	Simulation (including SPs and robotic or technological simulations)	NI	No	No	NA
Checklist; formative; summative; Other, student peer assessment + faculty raters	Storjohann et al. Assessment of a revised method for evaluating peer-graded assignments in a skills-based course sequence. *Am J Pharm Educ.* 2015;79(8):123.	5	Patient care, interpersonal and communication skills, practice based learning and improvement	Other, pharm-acy students	Other, pharmacy	Medical record keeping	Knowledge tests (written, multiple choice, oral)	Quantitative written	No	No	NA

EPA 6: Provide an oral presentation of a clinical encounter											
Assessment Category	Citation	EPA Relevance	Competency Domains	Level of Learner	Targeted Specialty Area(s)	Skill(s) Assessed	Type of Assessment Tool	Feedback Mechanism	Content Validity	Internal Structure	Type of Internal Structure
OSCE, formative, summative, curriculum	Daniel et al. Preparing for clerkships: learning to deliver specialty-specific oral presentations. *MedEdPORTAL Publications*. 2015;11:10261.	6	Patient care, interpersonal and communication skills	UME	Internal, family, and emergency medicine, pediatrics, surgery, obstetrics, psychiatry, neurology	Data gathering, interviewing, oral case presentation	Simulation (including SPs and robotic or technological simulations)	Qualitative written	No	No	NA
OSCE, formative, summative	Daniel et al. Teaching oral presentation skills to second-year medical students. *MedEdPORTAL Publications*. 2015;11:10017.	6	Medical knowledge, patient care, interpersonal and communication skills	UME	Emergency, family, and internal medicine, neurology, obstetrics, pediatrics, psychiatry, surgery	Clinical examination, data gathering, interviewing, oral case presentation	MSF	Qualitative written	No	No	NA
Curricular evaluation	Duong et al. Medical student perceptions on the instruction of the emergency medicine oral case presentation. *J Emerg Med*. 2015;48(3):337-343.	6	Patient care, interpersonal and communication skills	UME	Emergency medicine	Oral case presentation	Rubric	NI	No	No	NA
Rating sheet, formative	Lewin et al. The patient presentation rating tool for oral case presentations. *MedEdPORTAL Publications*. 2014;10:9659.	6	Patient care, interpersonal and communication skills	UME, GME	Internal medicine, neurology, pediatrics, psychiatry, surgery	Clinical examination, data gathering, diagnostic reasoning, decision making, clinical judgement, oral case presentation, combination (general aspects)	Rubric	Quantitative written, qualitative written, qualitative oral	No	Yes	Interitem correlation
Rating sheet, formative	Sox et al. Feedback on oral presentations during pediatric clerkships: a randomized controlled trial. *Pediatrics*. 2014;134(5):965-971.	6	Patient care, interpersonal and communication skills	UME	Pediatrics	Clinical examination, data gathering, diagnostic reasoning, decision making, clinical judgement, oral case presentation, combination (general aspects)	Rubric	Quantitative written, qualitative written, qualitative oral	No	Yes	Interitem correlation
Checklist, Rating sheet, 360° evaluation	Williams et al. Developing oral case presentation skills: peer and self-evaluations as instructional tools. *Ochsner J*. 2016;16(1):65-69.	6	Patient care, interpersonal and communication skills	UME	Internal medicine	Data gathering, interviewing, oral case presentation, communication, counseling, obtaining consent	Simulation (including SPs and robotic or technological simulations)	Quantitative qualitative written	No	No	NA