Mapping Cognitive Overlaps Between Practice-Based Learning and Improvement and Evidence-Based Medicine: An Operational Definition for Assessing Resident Physician Competence

Madhabi Chatterji, PhD
Mark J. Graham, PhD
Peter C. Wyer, MD

Abstract

Purpose The complex competency labeled practice-based learning and improvement (PBLI) by the Accreditation Council for Graduate Medical Education (ACGME) incorporates core knowledge in evidence-based medicine (EBM). The purpose of this study was to operationally define a "PBLI-EBM" domain for assessing resident physician competence.

Method The authors used an iterative design process to first content analyze and map correspondences between ACGME and EBM literature sources. The project team, including content and measurement experts and residents/fellows, parsed, classified, and hierarchically organized embedded learning outcomes using a literature-supported cognitive taxonomy. A pool of 141 items was produced from the domain and assessment specifications. The PBLI-EBM domain and resulting items were content validated through formal reviews by a national panel of experts.

Results The final domain represents overlapping PBLI and EBM cognitive dimensions measurable through written, multiple-choice assessments. It is organized as 4 subdomains of clinical action: Therapy, Prognosis, Diagnosis, and Harm. Four broad cognitive skill branches (Ask, Acquire, Appraise, and Apply) are subsumed under each subdomain. Each skill branch is defined by enabling skills that specify the cognitive processes, content, and conditions pertinent to demonstrable competence. Most items passed content validity screening criteria and were prepared for test form assembly and administration.

Conclusions The operational definition of PBLI-EBM competence is based on a rigorously developed and validated domain and item pool, and substantially expands conventional understandings of EBM. The domain, assessment specifications, and procedures outlined may be used to design written assessments to tap important cognitive dimensions of the overall PBLI competency, as given by ACGME. For more comprehensive coverage of the PBLI competency, such instruments need to be complemented with performance assessments.

Madhabi Chatterji, PhD, is Associate Professor Measurement and Evaluation at Teachers College, Columbia University; **Mark J. Graham, PhD,** is Director Education Research for the Center for Education Research and Evaluation, Columbia University; and **Peter C. Wyer, MD,** is Associate Clinical Professor of Medicine, Columbia University College of Physicians & Surgeons.

This project was funded in part by a National Board of Medical Examiners (NBME) Edward J. Stemmler, MD, Medical Education Research Fund grant. The project does not necessarily reflect NBME policy, and NBME support provides no official endorsement.

The authors thank Dr. Suzana Alves Silva, of the Federal University of the State of Rio de Janeiro, for her graphic assistance with the Figure.

Corresponding author: Madhabi Chatterji, PhD, Teachers College, Columbia University, 525 W 120th Street, New York, NY 10027, 212.678.3357, mb1434@columbia.edu

Editor's Note: The online version of this article contains examples of items developed using the PBLI-EBM domain specifications (Appendix).

DOI: 10.4300/JGME-D-09-00029.1

Introduction

The Accreditation Council for Graduate Medical Education (ACGME)[1,2] views competency in practice-based learning and improvement (PBLI) to broadly embody the goals of lifelong learning and continuous improvement of both patient care and clinical-teaching practices. This competency is 1 of 6 competencies embodied in the ACGME Outcome Project,[3] which seeks to adapt medical education to the rapidly evolving body of knowledge and organizational frameworks involved in the practice of medicine today. Together with systems-based practice, PBLI constitutes a cognitive and practice-related behavioral construct not covered in traditional models of medical education.[4,5]

The development of tools for the assessment of ACGME competencies has recently emerged as a major

priority[2,4,6] for the Outcome Project. In particular, attention is now being directed toward assessment of the systems-based practice and PBLI competency areas.[4,7] Currently, there are no developmentally informative or psychometrically validated instruments for assessing resident physician competency in PBLI. This competency is complex and includes both improvement learning and evidence-based medicine (EBM), which were previously treated as distinct and separate domains of knowledge, skill, and behavior.[1]

Background

Competency expectations for PBLI (BOX 1) could be manifested through attitudinal (values), behavioral (performance in practice contexts), and cognitive (knowledge and thinking skill) dimensions of resident physician practice. Because of the complexity and layered composition of PBLI, the ACGME acknowledges that multiple assessment modalities may be necessary to comprehensively assess resident competence in this area.[2] Cognitive dimensions of PBLI performance would call for intellectual processing of relevant information,[8] such as retrieving, accessing, reading, interpreting, and applying evidence from the scientific literature on the effectiveness of particular medical therapies for the care of patients. Some of these areas could be efficiently and validly measured with structured, written modes of assessment.

Performance dimensions of PBLI, on the other hand, are manifested through actual demonstrations of patient care behaviors in clinical-practice contexts. While these demonstrations would also draw on requisite cognitive capacities of physicians, they would demand qualitatively different skill and knowledge sets. For example, physicians could make decisions using the best evidence available on a given therapy, while weighing local conditions relevant to patient care. Simultaneously, they could pursue personal-development goals pertinent to PBLI's lifelong learning expectations.

As evident in BOX 1, keeping up-to-date with practice-relevant medical literature and drawing on it for the purpose of continuously improving practice constitutes a principal goal of PBLI competence. This articulated expectation of resident proficiency also links with the practice of EBM.[1] A superficial examination of BOX 1 shows at least 4 expected outcomes to speak to a body of general knowledge and skills that ties directly to the literature on EBM.[9]

Evidence-based medicine potentially serves to advance PBLI goals insofar as these capacities involve caring for individual patients, keeping up-to-date on practice-relevant medical literature, and applying it routinely to practice.[1,2] Although behaviors such as "use information technology to manage information…" may be construed to apply to a broad range of potential applications, they nonetheless constitute an explicit and required element of the process of bringing new information from clinical research to bear on practice-related decisions.

For PBLI and the other competencies described in the Outcome Project to constitute independently assessable constructs, they must be formally and consistently defined in operational terms.[4] Skill categories within the competency framework must be defined in a way that renders them meaningful with respect to the intended and underlying behaviors in the competency description. Following a methodology[10] previously used in connection with the competency of systems-based practice, we approached the task of defining selected but key cognitive dimensions embedded within PBLI that overlap with EBM. This research and development effort serves as a first step in a more complete and comprehensive specification of the larger PBLI domain and the design of appropriately aligned assessment tools.

Purpose

A necessary first step in sound assessment design is a specification of the domain in terms of observable responses, behaviors, tasks, or performances.[11] In this study, we sought to remove some of the ambiguities in the broadly defined PBLI competency area by specifying selected cognitive aspects that overlap with EBM, in more explicit terms. Our objective was to operationally define a "PBLI-EBM" domain by identifying, parsing, and organizing the common cognitive proficiencies through a systematic content analysis of relevant documentary sources, and to then validate it through review by established experts.

Our work was guided by the following questions:

- What should residents *know* and *be able to do* when they are "experts" in PBLI? Under what contexts or clinical *conditions* would experts demonstrate their knowledge and performance skills?
- How would one differentiate between *expert* and *nonexpert* performance in terms of the embedded

BOX 1 **The Accreditation Council for Graduate Medical Education Definition of Practice-Based Learning and Improvement**[3]

Residents must be able to investigate and evaluate their patient care practices, appraise and assimilate scientific evidence, and improve their patient care practices. Residents are expected to:

1. analyze practice experience and perform practice-based improvement activities using a systematic methodology.
2. locate, appraise, and assimilate evidence from scientific studies related to their patients' health problems.
3. obtain and use information about their own population of patients and the larger population from which their patients are drawn.
4. apply knowledge of study designs and statistical methods to the appraisal of clinical studies and other information on diagnostic and therapeutic effectiveness.
5. use information technology to manage information, access online medical information, and support their education.
6. facilitate the learning of students and other health care professionals.

(but yet unidentified) concept knowledge or thinking skills relevant to the various cognitive or more-practice-related, performance dimensions of PBLI?

- Will the necessary knowledge and skill sets be the same when residents tackle a *therapy* versus a *harm*, *prognosis*, or *diagnosis* issue with a patient?[9]
- What *tasks* or *assessment item types* would best tap into the wide range of underlying PBLI capacities of residents?

Our point of departure for the project was a pilot project,[12] which found that cognitive skills essential to a practice-based definition of EBM are frequently underemphasized by teachers and insufficiently mastered by learners in standard EBM instructional settings, such as workshops. Much of the available EBM training for physicians emphasizes critical appraisal of the scientific literature through journal club experiences, rather than contextualized aspects of patient care. The current research project addresses this critical issue by bridging common elements of PBLI and EBM within the context of the ACGME competency framework.

Method

Iterative Design Process

We applied an iterative design and validation process to operationally define a PBLI-EBM domain and develop an item pool,[13] where results from our earlier work phases guided the later phases of the assessment design process. Application of the approach has been demonstrated[10,14,15] in projects involving both ACGME competencies and with general education constructs. Core elements of our methodology draw from traditions in outcomes-based curriculum and assessment design typically employed in elementary and secondary education and in applications of domain sampling theory in the assessment literature.[11,16–20]

We began by specifying the "construct domain," or identifying a theoretically defensible but observable set of indicators (actions, words, responses, behaviors) representative of the competency area to be measured. The domain specification procedure involved a review and content analysis of curriculum-relevant documentary sources, expert consensus building, and the use of both qualitative reviews and survey research methods to gather support for the resulting organizational framework. As illustrated in Boxes 1–4, Tables 1–2 and the Figure, we then employed 3 iterative task cycles to arrive at a final domain, an item pool, and assessment specifications that we subsequently used to assemble parallel PBLI-EBM test forms.

Iteration 1: Preliminary Specification of the PBLI-EBM Domain

Content Analysis and Mapping of PBLI and EBM Skills The first iteration focused on content analysis and mapping of the agreement between ACGME and EBM literature. Our aim was to identify the overlapping versus unique skills in the PBLI and EBM areas (expectations 2–5 in BOX 1). We began with overall semantic comparisons of the statements by extracting key words or themes in context, using methods in qualitative text coding from linguistics.[21] To test for potential ambiguities or conceptual deficiencies in domain definition, a small set of items and item scenarios were developed, tied to the preliminary domain.

We drew upon 2 categories of literature: the traditional "educational literature" pertinent to EBM and "instructional literature" on EBM (papers pertinent to elaborating the concepts of clinical epidemiology relevant to EBM). From the EBM educational literature,[22,23] we recognized skill categories defined as Ask, Acquire, Appraise, and Apply as relevant to the aforementioned PBLI task prescriptions (BOX 1). To this, we added an independent set of categories derived uniquely from the instructional literature of EBM,[9,24] characteristically defined as Therapy, Diagnosis, Prognosis, and Harm, that represented the different types of clinical actions that physicians are called upon to take with patients. Systematic parsing methods typically employed in educational curriculum and test development were also applied to the material.[13,17–19] The method is illustrated with two excerpted statements from ACGME and EBM sources, respectively, in TABLE 1. We emphasized the skill categories Ask, Acquire, Appraise, and Apply in order to be consistent with the format reflected in the EBM educational literature.[23] As elaborated in the results section (FIGURE), our preliminary domain was subsequently transformed into a domain dominated by categories defined as Therapy, Diagnosis, Prognosis, and Harm (TDPH).

Preliminary Item Development and Pilot Testing of Items To test the viability of the above definition of subdomains by Ask, Acquire, Appraise, and Apply areas, we explored the process of generating test items using clinical scenarios. Resident volunteers from 3 target specialties—emergency medicine, internal medicine, and pediatrics—participated in the process of drafting clinical scenarios corresponding to resident-level clinical responsibilities and experience and developing candidate test items. The items were largely concentrated within the Ask subdomain. Our aim was to test respondent ability to classify clinical questions as pertaining to *one* of the areas under TDPH, as well as attempt appropriate formulation of questions for the Ask and Acquire skill sets. The items from this round were reviewed and validated internally by the item-writing and measurement teams of the research project, who noted and recorded difficulties faced in detail.

Iteration 2: Procedures for Respecification of the PBLI-EBM Domain

In the second iteration, we engaged in a more formal respecification of the domain with a particular focus on

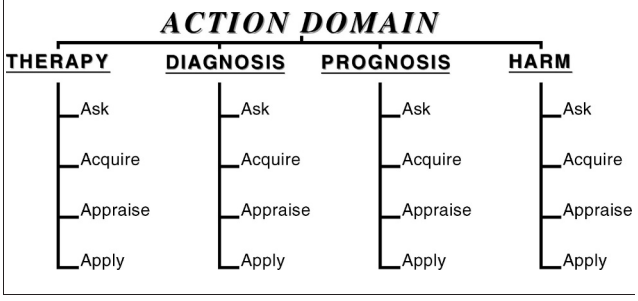
TABLE 1 **Content Analysis Procedure: Parsing and Classifying Learning Outcomes Using a Cognitive Taxonomy**

Literature Source	Excerpt From Source Document	Breakdown	Cognitive Classification per Taxonomy
Accreditation Council for Graduate Medical Education[3]	"Residents are expected to locate, appraise, and assimilate evidence from scientific studies related to their patients' health problems."	Residents are expected to:	
	Cognitive Processes:	Locate evidence	Application (use guidelines)
		Appraise evidence	Higher-order thinking (make judgments using learned criteria)
		Assimilate evidence	Higher-order thinking (synthesize information)
	Content:	From scientific studies related to their patients' health problems	NA
	Conditions:	Absent or unspecified	NA
American Medical Association *Users' Guides to the Medical Literature*[9]	"Residents (should be able to develop/conduct)... a clear definition of relevant questions; a thorough search of the literature relating to the questions; a critical appraisal of the evidence and of the applicability of the evidence to the clinical situation; and a balanced application of the conclusions to the clinical problem."	Residents should:	
	Cognitive Processes:	Develop questions	Application (use guidelines)
		Conduct a thorough search	Application (use guidelines)
		Critically appraise the evidence	Higher-order thinking (make judgements using learned criteria)
		Apply the evidence	Higher-order thinking (make decisions)
	Content:	Absent or unspecified	NA
	Conditions:	A given clinical situation or problem	NA

Abbreviation: NA, not applicable.

filling in the missing content in terms of medical knowledge and knowledge of clinical-research methodology. We started with a content analysis of literature relevant to the selected dimensions of PBLI and EBM to articulate a culminating or "end" outcome representing expert performance in PBLI-EBM. Starting with this end outcome, we employed a "design downwards" process[18] to systematically map backwards all the embedded skills and subject matter knowledge expected to lead to the specified culminating performance (see BOX 2).

FIGURE **Graphic Representation of PBLI-EBM Domain as a Tree Diagram**



Backward Analysis Based on substantive overlaps identified between PBLI and EBM in the previous iteration of content analysis (see TABLE 1), we formulated the end outcome for expert residents in PBLI-EBM as follows: "Guided by an appropriately formulated clinical question (demonstrating *ask skills*), resident physicians will *acquire*, *appraise*, and *apply* research-based evidence to make clinical decisions pertinent to specific patient types in each of the 4 subdomains of clinical action or advisement: *therapy*, *diagnosis*, *prognosis*, and *harm*."

We began development of a revised domain tree by backward analysis from this outcome, with the separate TDPH subdomains as dominant strands requiring specific kinds of concept knowledge. The steps in the procedure are summarized in BOX 2.

BOX 2 **Steps in Domain Specification Procedure**

1. Using a content analysis of salient curricular materials (text, audio, video, or literature sources), identify the culminating performance outcome expected in a competent or expert learner. This should be set as a long-term, broad goal.
2. Perform a backward analysis from the culminating performance expectation (see FIGURE). Subdomains with embedded. Subdomains with embedded skills should branch out in the form of a tree diagram from the end outcome.
3. Parse and identify the full-range content, conditions, and cognitive processes underlying broad outcomes as well as possible (see TABLE 1).
4. Formulate or restate embedded competencies and skills as outcome statements. Each statement should have a clear content element and cognitive process. If applicable, some might have a condition specified.
5. Use a taxonomy of learning outcomes (from suitable educational resources) to classify cognitive demands implied in outcome statements.
6. Organize the learning outcomes from general to more specific, mapping reasonable pathways from novice to expert development on a continuum.??

Example of outcome statement: Given clinical-patient scenarios pertaining to a therapy problem (condition), the resident should be able to use published evidence-based medicine guidelines (content) to craft a focused question to guide an evidence search (cognitive process—application).
Guiding questions:

1. What are the different types of subject matter and cognitive processes implied or articulated explicitly in the material?
2. Under what conditions would expert physicians typically demonstrate the cognitive processes and mastery of content elements, based on the literature sources?
3. What is a reasonable progression from novice to greater levels of expertise?

(Adapted from Chatterji,[13] chapter 6.)

The backward analysis from the culminating outcome required a continuation of the same process of content analysis shown in TABLE 1 to identify the key embedded content, cognitive processes, and conditions for observable competence. Relevant curricular literature served as the sources for formulating the general and more specific outcome statements. At follow-up meetings, the project's medical and measurement teams reviewed the drafts of the reconfigured domain and arrived at a consensus on its substantive utility.

Execution of the above procedure resulted in a tree diagram of hierarchically organized competencies and enabling skills leading to a culminating performance outcome. The broader outcome, and the more specific embedded competencies and enabling skills, were next classified with a literature-supported taxonomy of learning outcomes adapted from the educational literature.

Several cognitive taxonomic tools[17,25] are available to help classification of different types and levels of cognitive competence and skills embedded in curricular domains. We applied an adapted version of the Functional Taxonomy of Learning Outcomes[13] (hereafter referred to as the "taxonomy"), permitting recognition of the following types of cognition:

- Concept knowledge and understanding: retrieval of discipline-related concepts, definitions, terms, or principles, or a demonstration of basic levels of concept understanding
- Application: skills in applying discipline-based concepts, principles, rules, algorithms, or guidelines while solving a problem or performing a task
- Higher-order thinking and problem-solving skills: analysis (breaking a problem into parts), synthesis (putting together a whole from the parts), evaluation (making decisions or judgments using learned

TABLE 2 **Table of Specifications for Item Pool Generation and Test Form Assembly in Practice-Based Learning and Improvement-Evidence-Based Medicine**[a]

	Therapy, %	Diagnosis, %	Prognosis, %	Harm, %	Total Questions, N (Target %)
Weights Allocated to Subdomains	30	30	10–20	20–30	100 (100)
Ask Skills—Situated in Specific Subdomains					20 (20)
Ask Skills—Crosses Subdomains					10 (10)
Acquire Skills—Crosses Subdomains					20 (20)
Appraise Skills—Situated in Specific Subdomains					35 (35)
Apply Skills—Situated in Specific Subdomains					15 (15)

[a] Final test form with roughly equal distribution of questions in pediatrics, emergency medicine, and internal medicine. Cells contain items that fit each subdomain crossed with cognitive-skill areas. The "Total Questions" cell shows targeted number of questions in a hypothetical item pool of 100.

criteria), or some combination of these types of cognition in open-ended, creative, or structured problem settings

- Complex procedural skills: cognitive processes required to tackle multistep, complex, and integrative tasks, calling for use of relevant concept knowledge, application of principles and rules, procedures inherent in a discipline, or higher-order thinking processes in open-ended, creative, or structured problem settings

The domain and pool of items produced from the second iteration were validated and refined through an internal review by medical and educational-measurement specialists on the project's research team.

Developing a Table of Assessment Specifications We also generated a formal Table of Specifications, or a test blueprint to guide item and test development (TABLE 2). The process involved facilitated discussions and consensus seeking among the medical and measurement teams in a workshop. Starting with a 2-dimensional matrix representing cognitive-skill branches and the clinical action subdomains, we arrived at a table specifying the distribution of items in each cell and the weight allocations based on relative number and percent of items. TABLE 2 reflects several levels of priority set by the clinical and assessment design teams. Among the subdomains, Therapy and Diagnosis were prioritized over Prognosis and Harm and given more item and task representation, based on their greater importance to practitioners.[12] Additionally, items were intended to be distributed equally in 3 medical-specialty areas: pediatrics, internal medicine, and emergency medicine.

Item Pool Generation Three workshops were dedicated to item production, guided by the internally validated Table of Specifications and the more specific item-level specifications (BOX 3). A decision was made to use multiple-choice items because the format is more versatile for tapping varying levels of cognition represented in the domain. The format also allows for efficient large-group test administrations necessary for program-wide monitoring of resident outcomes. Two types of multiple-choice items were produced: stand-alone and scenario-dependent. The medical item-writing team comprised 6 individuals from internal medicine, pediatrics, and emergency medicine, in concordance with our 3-specialty focus for the present study. To start, measurement specialists on the research team provided general training in assessment design and item writing aligned with the domain. Following this, the principal investigator (a specialist in emergency medicine) led the writing team in item pool generation. We followed the Table of Specifications (TABLE 2) closely to track item production in each cell.

As demonstrated through the 2 examples in BOX 3, the assessment design training provided experiences and examples in parsing and content analyzing individual and related sets of skills and competencies from the domain, and in applying standard item-writing rules to generate items matched to the skill specifications.[13,16,19] Following item-writing rounds, we conducted content validation exercises to review, revise, and improve quality of the item pool. Guiding questions for these reviews included the following:

- Does the item match the content, cognitive processes, and conditions specified in the competency or skill statements?
- Is the item clear and easy to read?
- Is there any bias (related to ethnicity, culture, gender, disability, or medical specialty) indicated in the language, content, or scenarios?

BOX 3 **Item Development Procedures: Examples Tapping "Concept Knowledge" Competencies**

Competency in Ask Branch
Given a mix of clinical questions to guide an evidence search, identify a Harm question and distinguish it from Therapy, Diagnosis, and Prognosis questions.

Analysis of Competency
Condition: Given a mix of clinical questions to guide an evidence search
Content: Definitions of Harm, Therapy, Diagnosis, and Prognosis (TDPH); examples of Harm and Therapy, Diagnosis, and Prognosis questions and clinical scenarios
Cognitive process(es): Identify and distinguish (concept knowledge).

Matching Ask Item in Multiple-choice Format in Harm Subdomain
Evidence searches are guided by well-formulated questions. Which of the following questions about newborn fever pertains to the concept of *harm* as opposed to *therapy*, *prognosis*, or *diagnosis*? Select the best answer.

a. Does the height of fever in a baby correlate with bacterial load?
b. Does immediate treatment, as opposed to delayed treatment with antibiotics, decrease morbidity?
c. Does hospitalization of the febrile infant lead to increased risk of hospital-acquired illness?
d. Does a high white cell count correlate with a high risk of bacteremia?

Competency in Acquire Branch
With reference to a clinical scenario, identify the defining characteristics of classes of research designs and evidence synthesis methods that yield the best available evidence relevant to answering therapy, diagnosis, prognosis, or harm questions and that may be obtained through syntheses or empirical studies of various designs.

Analysis of Competency
With reference to a clinical scenario, define characteristics of classes of research designs and evidence synthesis methods that yield the best available evidence-evidence standards and hierarchies of evidence in TDPH action paths.
Cognitive process(es): Identify (concept knowledge)

Matching Acquire Item
A patient presents with progressive and recurring symptoms of asthma. In terms of strength of evidence, which of the following would yield the *best* available evidence regarding a question on the most suitable *therapy* for asthma? Select the best answer.

a. Narrative reviews
b. Systematic reviews of randomized controlled trials
c. Primary studies of randomized controlled trials
d. Systematic reviews

- Is there a clear correct or best answer?
- Are the distracters (wrong answers) written to tap common errors and misunderstandings, and are they clearly incorrect?
- Does the overall item distribution comply with the intended weights in the Table of Specifications?

Iteration 3: Procedures for External Validation

The third and final iteration involved obtaining external, expert review and validation of all the products—namely, the refined PBLI-EBM domain, the item pool, and the Table of Specifications. Once that feedback was received and incorporated into the products, PBLI-EBM test forms were assembled using the externally validated blueprint and prepared for a field test. Specific types of feedback were solicited from the external evaluators, with the help of a brief questionnaire and cover letter explaining the goals of the project (BOX 4). The 6-member validation team included editors and authors of the two leading textbook sources on EBM, an author of original work related to EBM, and an experienced teacher of EBM.

BOX 4 **Content Validation Survey**

Name:
Phone number:
Area of expertise:
Questions for review of the practice-based learning and improvement (PBLI)–evidence-based medicine (EBM) domain and Table of Specifications:
Please indicate "Yes" or "No" in the space provided. If "No," please clarify what you would like changed, added, or deleted using the "Track Changes" option in the Microsoft Word document.

1. _____Overall, the content and organization of the domain seems an appropriate fit for the Accreditation Council for Graduate Medical Education standards and existing literature on EBM (please also see references and the Table of Definitions).
2. _____The subdomains appropriately measure PBLI-EBM skills and concept knowledge.
3. _____The Ask, Acquire, Appraise, and Apply subdomains are appropriate.
4. _____The enabling skills encompass all the key elements that need to be tested.

Other comments on additions/changes: [space provided]

Please review the following items, written to tap into the given skill or skill set from the PBLI-EBM domain. Correct answers, when given by item writers, are shown in italics.
[Item and skill specifications provided]
Item #____
Indicate with a "Yes" or "No" your judgment of item quality.

1. The item matches the competencies and/or skills listed. ___
2. The item is clear and easy to read and understand. ___
3. There is no evident or inherent bias in the language, content, or scenarios (bias related to ethnicity, culture, gender, disability, or medical specialty). ___
4. There is clearly one correct or best answer. ___
5. The distracters (wrong answers) are reasonable but incorrect. ___

Please recommend any changes you would make to the item using the "Track Changes" option.

Results

Iteration 1: Mapping Correspondences Between PBLI and EBM

Results from the first iteration are provided in TABLE 3. A sample of the results may also be viewed in TABLE 1, which shows the analysis procedure and parsed expected outcomes for residents. As seen in TABLE 3, the semantic mapping showed some clear correspondences between 3 dimensions of PBLI and EBM outcome statements. Our review allowed us to begin matching some of the identified cognitive dimensions with appropriate assessment methods. The implied concept knowledge and higher-order thinking for appraising medical research, for example, could potentially be tapped through written assessments with structured response items (eg, with multiple-choice items and patient scenario–based tasks calling for simulated applications of the implied EBM principles). Other cognitive dimensions involving individualized patient care decisions characteristic of PBLI, would demand contextualized observations of resident performance in actual practice contexts.

The preliminary content analysis suggested that there were 4 requisite skill areas that could potentially serve as the principal subdomains (broad learning outcomes) around which resident actions would revolve. Our starting definitions were as follows:

- Ask skills: Residents should be able to frame clinical-research questions to guide searches of the medical literature (research evidence).
- Acquire skills: Residents should be able to seek and access appropriate research resources.
- Appraise skills: Residents should be able to evaluate the quality of available research evidence drawing on knowledge of research methods and data analysis.
- Apply skills: Residents should be able to use the evidence, along with other relevant information, for making specific patient care decisions.

Several insights emerged from the exploratory item-writing work using this preliminary domain specification. Firstly, it became evident that our resident volunteers (item writers) were significantly challenged by the task of classifying questions into TDPH categories, more so than anticipated. This led to the realization that the TDPH terms themselves needed to be rigorously defined as part of the construct definition process.

More importantly, it became evident that the information literacy skill categories of Ask, Acquire, Appraise, and Apply had been defined in a content-free mode (see TABLE 1 for the analysis that revealed this finding). The skills were still too broadly configured to allow for item design that would ensure adequate levels of content-based validity (sufficiently tight alignment of items with the content, cognitive processes, or conditions). Inadequate domain specification poses barriers to both item generation and evaluation of how representative the item pool is across content and cognitive-process categories; this

TABLE 3 **Correspondence Between Practice-Based Learning and Improvement (PBLI) Cognitive-Skill Components and Categories Characterizing the Evidence-Based Medicine (EBM) Education Literature**

PBLI[3]		EBM[26]		
Performance-Based Skills	**Cognitive Skills**	**Cognitive Skills Measurable With Written Assessments**	**Cognitive Skills Measurable With Performance Assessments**	**Comments**
Perform practice-based improvement activities using a systematic methodology	**Analyze practice experience**	**Identify the information need and formulate the question**	NA	The cognitive component of "analyzing practice experience" is common to PBLI and EBM.
Assimilate evidence from scientific studies related to patients' health problems	**Locate, appraise, and *assimilate evidence from scientific studies* related to patients' health problems**	***Integrate the evidence*, once appraised, into clinical decisions in given patient scenarios**	Integrate the evidence, once appraised, into clinical decisions while taking individual patient values and circumstances into account	The cognitive skills pertaining to "assimilation" or "integration" of information from clinical research into decision making for patients are shared by PBLI and EBM.
Obtain and use information about *one's* own population of patients and the larger population from which *one's* patients are drawn.	NA	Unmatched		This skill domain within PBLI is not included in the standard definition of EBM or EBM curricula. This PBLI skill set is considered to be largely performance based.[27]
NA	**Apply knowledge of study designs and statistical methods to the appraisal of clinical studies and other information on diagnostic and therapeutic effectiveness**	**Critically appraise the evidence for internal validity, external validity, issues of bias, research design limitations, and applicability to patient scenarios.**	NA	Both descriptions pertain to the use of standardized criteria for evaluating the internal and external validity of studies relevant to a clinical question and involve cognitive skills.
Use information technology to support one's own education	**Use information technology to manage information and access online medical information**	**Search for and select the best evidence from clinical research**	NA	The ability to accurately and efficiently search for clinical evidence within the broader framework of using online resources involves cognitive skills common to PBLI and EBM.
Facilitate the learning of students and other health care professionals	NA	NA		This performance-based skill domain within PBLI is not an explicit part of the core skill set of EBM.
Unmatched		Unmatched	Evaluate the effectiveness and efficiency of performing the "Ask," "Acquire," "Appraise" and "Apply" steps and seek ways to continuously improve this process.	This skill domain within EBM is not explicitly a part of the ACGME definition of PBLI but is central to the performance-based domains of improvement learning.

Abbreviations: ACGME, Accreditation Council for Graduate Medical Education; NA, not applicable.

principle is critical for assuring construct validity of assessment results.[11]

Iteration 2: A Reconfigured Domain and Item Pool

In the reconfigured PBLI-EBM domain shown in FIGURE we separated qualitatively different types of Ask, Acquire, Appraise, and Apply skills that were unique to, and thus nested within, each of the TDPH subdomains of clinical action. TABLE 4 provides an excerpt of the PBLI-EBM domain in the Ask skill branch of the Therapy subdomain. BOX 5 presents an example of a scenario-based item set from the item pool.

As shown in TABLE 4, assuming no prior exposure to PBLI-EBM material, the culminating performance outcome

TABLE 4 **Domain Specifications Showing an Excerpt of the Therapy Subdomain in Practice-Based Learning and Improvement (PBLI)–Evidence-Based Medicine (EBM)**

Culminating Performance Outcome
Guided by an appropriately formulated clinical question, resident physicians will acquire, appraise, and apply research-based evidence to make clinical decisions pertinent to specific patient types, in 4 subdomains of clinical action: Therapy, Diagnosis, Prognosis, and Harm. (complex procedural skills)
Subdomain 1.0: Therapy
Ask Skill Branch
Embedded competency 1.1: To begin an evidence search on a therapy, resident physicians will be able to ask a well-formulated clinical question, identifying the specific patient population (P), the type of intervention in question (I), a comparison/control condition (C), and the outcomes for the patient population (O), or PICO.[9]
Enabling Skills
1.1-a. 1 Define the concept of "therapy" as a form of clinical action or advisory, and distinguish it from "diagnosis," "prognosis," and "harm" as applicable to PBLI-EBM. (concept knowledge)
1.1-a. 2 Given examples of clinical questions pertinent to the 4 subdomains of clinical action and advisement, identify a Therapy question and distinguish it from Diagnosis-, Prognosis-, and Harm-related questions. (concept knowledge)
1.1-a. 3 Given clinical scenarios corresponding to issues of therapy, diagnosis, prognosis, or harm, identify the questions that correspond to issues of therapy. (concept knowledge)
1.1-b. Rationalize the need for inclusion of all PICO elements to frame a focused clinical question about therapy before beginning an evidence search. (higher-order thinking)
1.1-c. Given clinical scenarios pertaining to therapy that are ill formulated, identify missing PICO elements to reframe the question before beginning an evidence search. (application skills)
1.1-d. Identify/define/distinguish between the meaning of terms such as *outcome*, *intervention*, *exposure*, and *comparison/control condition* when asking a clinical question to guide an evidence search pertaining to an issue of therapy. (concept knowledge)

was expected in expert resident physicians who had completed requisite medical training or were approaching the end of their residency training. It was classified as a complex procedural skill based on the taxonomy of learning outcomes. The hierarchical organization meant that in contrast to the culminating outcome, an enabling skill embedded within was expected to carry a relatively lower cognitive load. For example, the following embedded competency tapping the recall of PBLI-EBM concept knowledge fell at a lower cognitive level (concept knowledge) per the taxonomy: "Define the concept of 'therapy' as a form of clinical action or advisory, and distinguish it from 'diagnosis,' 'prognosis,' and 'harm.' (concept knowledge)"

Concept knowledge was a necessary prerequisite to application or higher-order thinking. In contrast, the following embedded competencies were categorized as application skill based on the taxonomy. In the first, examinees would be called upon to apply accepted guidelines and conventions in PBLI-EBM to frame a clinical question to guide an evidence search. In both, they demonstrate the skill in the context of a clinical scenario, such as treatment options for a patient presenting with asthma symptoms:

- Given clinical scenarios pertaining to therapy, use EBM guidelines to craft a focused question to guide an evidence search. (application skill)
- Given questions situated in a clinical scenario pertaining to therapy that are poorly formulated, identify missing population, intervention, control or comparison condition, and outcome elements and reframe the question before beginning an evidence search. (application skill)

As formulated, the embedded competencies or enablers could be assessed, via individual items or in unison with others as a component of a more complex task, during the assessment design phase. Some PBLI-EBM tasks would call for *integrative* utilization of embedded knowledge and skills; others might call for a *separate* use or even a *step-by-step*, *sequential* use.

The tree organization was also *not* intended to imply a firm or fixed path for skill acquisition. We accepted that different physicians could draw on the delineated concept knowledge and skills in different ways to show expertise. Rather, the tree helped map out a literature-supported and more complete range of embedded concept knowledge and cognitive skills, organized by level, that could now guide instruction and assessment design in a more-informed way.

BOX 5 A Scenario-Based Item Set From the Therapy Subdomain: Appraise Skill Branch Item Specifications

Higher-Order Thinking Skills
Given excerpted data and evidence tables/graphs from therapy-related studies, interpret and draw conclusions from the evidence for a specific patient type, reported using odds ratios, risk reduction, relative risk reduction, or number needed to treat.
Concept Knowledge
Embedded Concept Knowledge
Recall/recognize information on relevant measures of comparative outcomes, such as odds ratios, risk reduction, relative risk reduction, and number needed to treat.
Application Skills
Embedded Application Skill
Calculate relevant measures of comparative outcomes, such as odds ratios, risk reduction, relative risk reduction, and number needed to treat.
Matching Item Set
The data below summarizes the 1-year mortality outcomes from a randomized trial comparing endoscopic ligation with endoscopic sclerotherapy for bleeding esophageal varices. Calculate and interpret the relative risk reduction of ligation compared to sclerotherapy, based on this data:

- Following 64 total ligation procedures, there were 18 deaths and 46 survivals.
- Following 65 total sclerotherapy procedures, there were 29 deaths and 36 survivals

Item 1: Relative risk reduction for death in the above trial is:

a. Risk in the ligation group (18/64) compared to risk in the sclerotherapy group (29/65).
b. Risk in the ligation group (18/64) minus the risk in the sclerotherapy group (29/65).
c. The difference in risk between the two groups [(29/65) – (18/64)] compared to the risk in the ligation group (18/64).
d. The difference in risk between the two groups [(29/65) – (18/64)] compared to the risk in the sclerotherapy group (29/65).

Item 2: Based on the data, which of the following is the most reasonable conclusion pertinent to patient care?

a. Endoscopic sclerotherapy places patients at lower risk for mortality.
b. Most patients will prefer to receive endoscopic ligation instead of endoscopic sclerotherapy.
c. The number of patients who would need to be switched from endoscopic sclerotherapy to endoscopic ligation to save 1 life could be as few as 6.
d. Reasonable patients would prefer endoscopic sclerotherapy over endoscopic ligation.

Source for data: Guyatt and Rennie[9(pp352–362)]

The process of item development led to an item pool of 119, 7 of which comprised a series of related questions pertaining to a single scenario. Treated individually, the resulting item pool is composed of 141 items that can be scored. The final test forms (A and B) had a proportionate distribution of items across the subdomains and skill categories based on the Table of Specifications (TABLE 2). With respect to clinical emphasis, the apportionment of the independent items across target specialties was 63 of 119 (52%) in internal medicine, 26 of 119 (22%) in emergency medicine, and 19 of 119 (15%) in pediatrics, with 13 of 119 (11%) not classifiable by specialty.

Using the cognitive taxonomy to classify each outcome, competency, and enabling skill in the domain allowed us to write items to elicit different types of skill or knowledge usage, as shown in BOX 5. Items and item sets tied to different cognitive specifications were expected to place different cognitive demands on test takers. Such assumptions about cognitive demand could subsequently be empirically tested through item statistics during field tests of assessment instruments.

Iteration 3: External Content Validation

The PBLI-EBM domain and item pool review resulted in several edits to items and the language in the final domain and item pool. Overall there was consensus among reviewers on the basic framework and breadth of content and cognitive skills represented in the items and domain. Of the larger pool, most items passed content validity screening criteria and were prepared for test form assembly and administration. Field testing and psychometric refinement of the item pool and PBLI-EBM assessments is continuing.

Discussion

The purpose of the ACGME Outcome Project is to objectify educational outcomes as the basis for accreditation of US residency programs.[6] To achieve this objective, the performance descriptors and outcomes need to be rendered observable and measurable.[4] We employed an iterative methodology previously successfully used to derive an operational definition of systems-based practice[10] to define the critical cognitive skills within the PBLI competency.

The ACGME[1] recognizes a close correspondence between the skills and attributes included in its description of PBLI and those associated with EBM. However, our project is the first to advance a fully defined construct, mapping the overlapping dimensions of PBLI and EBM. Because the ACGME identifies EBM as a component of other competencies outside PBLI, such as medical knowledge,[1] our task was to define and conceptualize skill sets common to both EBM and PBLI in a way that uniquely represents the stated intent of the ACGME and enhances valid assessment. We believe that our effort has produced a differentiated elaboration of the PBLI-EBM domain, a corresponding table of specifications, and an item pool that taps necessary cognitive dimensions of the ACGME practice-based learning competency that are measurable through written assessments. These formats are suitable for large-scale administration in medical education programs. The systematic-design approach enhanced our understandings of EBM itself as an integrated informant of clinical practice.

The ACGME formulation of PBLI may be summarized as a model for reflective, self-directed learning and for practice improvement. Within our domain definition, the relative emphasis on skills pertaining to asking questions based on practice experience and applying the results of researching them serves to integrate the emphasis on individual practice improvement with the skills required to keep practice up-to-date with scientific literature. We

believe that this conceptual integration is crucial to the integrity of PBLI as a construct.

Improvement learning is a concept that evolved independently of other dimensions of medical education.[28] Self-improvement in response to experience is a general concept that, like EBM, may be considered applicable to all 6 ACGME competencies. Correspondingly, it can be incorporated into a competency-based curriculum in a fashion that has little to do with concern for bringing clinical decisions in line with current scientific evidence. Varkey and colleagues[7] describe an Objective Structured Clinical Examination model in which improvement learning and critical appraisal of the medical literature constitute parallel and nonintersecting components. Such an approach, in which the behavioral, performance-based, and cognitive dimensions of what is intended to constitute a single construct are divorced from each other, appears to us to be unsatisfying as a solution for guiding assessment efforts.

In our pilot study, we assessed the ability of participants in a standard EBM workshop to respond to a simulated clinical encounter by identifying, correctly classifying, and researching a question requiring knowledge of current clinical evidence.[12] This identified a category of EBM skills, characterized as ''initiation'' skills, which were inconsistently mastered by workshop attendees and which are crucial to a practice-based understanding and competency like EBM. These skills include an understanding of the nature of clinical questions being asked and the ability to connect them to appropriate electronic resources and study designs. These skill sets are, in fact, frequently de-emphasized in standard EBM workshops and curricula, in favor of concentration on appraisal of research reports.[29] Existing EBM textbooks[9] and published efforts at developing EBM assessment tools[30,31] are similarly weighted toward critical appraisal of evidence and away from practice-based initiation skill sets.

The EBM instructional literature is also heavily concentrated on issues of therapy, in contrast to those pertaining to diagnosis, prognosis, and harm.[32] One reason for the weighting on critical appraisal within EBM literature and instructional experiences is that, within residency programs, these skills are largely sequestered in the framework of ''journal club'' activities.[33] The PBLI-EBM domain definition and specifications developed in our project provide a conceptual framework that should remedy such sequestration and serve to advance the common objectives of EBM and the development of competence in PBLI.

The establishment of therapy, diagnosis, prognosis, and harm as ''action domains'' governing the content of the enabling-skill categories (ask, acquire, appraise, and apply) constitutes an original feature of our elaboration of the PBLI-EBM cognitive domain. It is also uniquely appropriate to a practice-based conception of these skills. These categories are drawn from EBM-related literature dating from the early 1980s.[32] However, they have not previously been consistently and comprehensively defined as categories of clinical action, nor have they been systematically applied across the range of knowledge and skills required to inform clinical practice through the results of clinical research. The implications of recognizing the action domains as the primary determinants of the content of EBM skills are potentially far reaching and point to approaches to integration of EBM with other dimensions of clinical practice. The implications of this are undergoing further examination by our research team.

We have presented the iterative process to facilitate replication of the procedures in other ACGME curricular areas and medical education domains. Specifically, TABLES 1 and 2 and BOXES 1 through 5 provide demonstrations and concrete examples to facilitate future applications. The methodology can be generalized, and it is supported by established methods of curriculum-based assessment design documented in standards-based curriculum projects in general education (see, for example, the assessment of the mathematics standards of the National Council of Teachers of Mathematics).

Limitations

The cognitive skills embedded in PBLI-EBM do not encompass the entire PBLI competency given by the ACGME, but they serve as a good beginning. Added cognitive capacities required for performance-based skills and behavioral attributes, including general knowledge of ''improvement learning'' principles, need to be addressed and elaborated through future research. There is a growing literature on the use of portfolios as vehicles for performance-based and behavioral aspects of competency assessment. However, a systematic review[34] found that published efforts reflect only preliminary attention to rigorous assessment. Further work on the development of performance-based assessment is needed and under consideration by our research team.

Conclusions

This paper reports the results of the construct domain specification phase for selected cognitive dimensions of a PBLI-EBM domain. The nationally validated PBLI-EBM domain could contribute practical guidance to residency program directors in understanding this crucial component of the Outcome Project framework and provides a foundation for local assessment development efforts.

Our research project demonstrates that multidisciplinary forms of knowledge are necessary to accomplish objectives of a project such as this. Medical and educational-measurement faculty and specialists contributed complementary types of expertise to the effort.

Our project shows a process for effectively coordinating interdisciplinary work across campuses.

We reiterate that the current domain represents a start in bringing clarity to the PBLI competency, and it speaks only to the overlapping PBLI and EBM cognitive dimensions that are measurable through written, multiple-choice assessments. The need for further research to clarify and assess remaining cognitive and practice-based dimensions of PBLI continues.

References

1 Joyce B. Practical implementation of the competencies. Available at: http://www.acgme.org/outcome/e-learn/Module2.ppt#256,1,Practical Implementation of the Competencies. Accessed November 17, 2009.
2 Joyce B. Developing an assessment system: facilitator's guide. Available at: http://www.acgme.org/outcome/e-learn/Module3.ppt. Accessed November 17, 2009.
3 Accreditation Council for Graduate Medical Education. Outcome Project: General Competencies. Available at: http://www.acgme.org/outcome/comp/compFull.asp#3. Accessed November 19, 2009.
4 Lurie SJ, Mooney CJ, Lyness JM. Measurement of the general competencies of the Accreditation Council for Graduate Medical Education: a systematic review. *Acad Med.* 2009;84:301–309.
5 Varkey P, Karlapudi S, Rose S, Nelson R, Warner M. A systems approach for implementing practice-based learning and improvement and systems-based practice in graduate medical education. *Acad Med.* 2009;84:335–339.
6 Swing S. The ACGME Outcome Project: retrospective and prospective. *Med Teach.* 2007;29:648–654.
7 Varkey P, Natt N, Lesnick T, Downing S, Yudkowsky R. Validity evidence for an OSCE to assess competency in systems-based practice and practice-based learning and improvement: a preliminary investigation. *Acad Med.* 2008;83:775–780.
8 Pellegrino JW, Chudowsky N, Glazer R, eds. *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press; 2001.
9 Guyatt G, Rennie D. *Users' Guides to the Medical Literature: A Manual For Evidence-Based Clinical Practice*. Chicago, IL: AMA Press; 2002.
10 Graham MJ, Naqvi Z, Encandela J, Harding KJ, Chatterji M. Systems-based practice defined: taxonomy development and role identification for competency assessment of residents. *J Graduate Med Educ.* 2009. In press.
11 American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: APA, 1999.
12 Wyer PC, Naqvi Z, Dayan PS, Celentano JJ, Eskin B, Graham M. Do workshops in evidence-based practice equip participants to identify and answer questions requiring consideration of clinical research? A diagnostic skill assessment. *Adv Health Sci Educ.* 2009;14:515–533.
13 Chatterji M. *Designing and Using Tools for Educational Assessment*. Boston, MA: Allyn & Bacon; 2003.
14 Chatterji M. Measuring leader perceptions of school readiness for standards-based reforms and accountability. *J Appl Meas.* 2002;3:455–485.
15 Chatterji M, Sentovich C, Ferron J, Rendina-Gobioff G. Using an iterative validation model to conceptualize, pilot-test, and validate scores from an instrument measuring Teacher Readiness for Educational Reforms. *Educational and Psychological Measurement* 2002;62:442–463.
16 Crocker LM, Algina J. *Introduction to Classical and Modern Test Theory.* Orlando, FL: Harcourt Brace Jovanovich; 1986.
17 Marzano RJ, Pickering DJ, McTighe J. *Assessing Student Outcomes: Performance Assessment Using the Dimensions of Learning Model*. Alexandria, VA: Association for Supervision and Curriculum Development; 1993.
18 McTighe J, Thomas R. Backward design for forward action. *Educational Leadership* 2003;60:52–55.
19 Millman J, Greene J. The specification and development of tests of achievement and ability. In: Linn RL, ed. *Educational Measurement*. 3rd ed. New York, NY: Macmillan; 1989:335–366.
20 Nunnally JC, Bernstein IH. *Psychometric Theory.* 3rd ed. New York, NY: McGraw-Hill; 1994.
21 Ryan GW, Bernard HR. Data management and analysis methods. In: Denzin N, Lincoln Y, eds. *Handbook of Qualitative Research*. 2nd ed. Thousand Oaks, CA: Sage Publications; 2000:769–802.
22 Shaneyfelt T, Baum KD, Bell D, et al. Instruments for evaluating education in evidence-based practice: a systematic review. *JAMA.* 2006;296:1116–1127.
23 Straus SE, Green ML, Bell DS, et al. Evaluating the teaching of evidence based medicine: conceptual framework. *BMJ.* 2004;329:1029–1032.
24 Guyatt GH, Rennie D. User's guides to the medical literature. *JAMA.* 1993;270:2096–2097.
25 Bloom BS, Englehart MD, Furst EJ, Hill WH, Krathwoohl DR. *Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I. Cognitive Domain*. White Plains, NY: Longman; 1956.
26 Straus SE, Richardson WS, Glasziou P, Haynes RB. *Evidence-Based Medicine: How to Practice and Teach EBM*. 3rd ed. Edinburgh, UK: Elsevier Churchill Livingstone; 2005.
27 Manning PR. Practice-based learning and improvement: a dream that can become a reality. *J Contin Educ Health Prof.* 2003;23(suppl 1):S6–S9.
28 Headrick LA, Richardson A, Priebe GP. Continuous improvement learning for residents. *Pediatrics.* 1998;101:768–774.
29 Thomas KG, Thomas MR, York EB, Dupras DM, Schultz HJ, Kolars JC. Teaching evidence-based medicine to internal medicine residents: the efficacy of conferences versus small-group discussion. *Teach Learn Med.* 2005;17:130–135.
30 Fritsche L, Greenhalgh T, Falck-Ytter Y, Neumayer HH, Kunz R. Do short courses in evidence based medicine improve knowledge and skills? Validation of Berlin questionnaire and before and after study of courses in evidence based medicine. *BMJ.* 2002;325:1338–1341.
31 Ramos KD, Schafer S, Tracz SM. Validation of the Fresno test of competence in evidence based medicine. *BMJ.* 2003;326:319–321.
32 Wyer PC, Silva SA. Where is the wisdom? I. A conceptual history of evidence-based medicine. *J Eval Clin Prac.* 2009. In press.
33 Kuhn GJ, Wyer PC, Cordell WH, Rowe BH. A survey to determine the prevalence and characteristics of training in evidence-based medicine in emergency medicine residency programs. *J Emerg Med.* 2005;28:353–359.
34 Tochel C, Haig A, Hesketh A, et al. The effectiveness of portfolios for post-graduate assessment and education: BEME Guide No 12. *Med Teach.* 2009;31:299–318.