Transforming Medical Assessment: Integrating Uncertainty Into the Evaluation of Clinical Reasoning in Medical Education

Suzette Cooke, MD, PhD, and Jean-Francois Lemay, MD

Abstract

In an age where practicing physicians have access to an overwhelming volume of clinical information and are faced with increasingly complex medical decisions, the ability to execute sound clinical reasoning is essential to optimal patient care. The authors propose two concepts that are philosophically paramount to the future assessment of clinical reasoning in medicine: assessment in the context of "uncertainty" (when, despite all of the information that is available, there is still significant doubt as to the best diagnosis, investigation, or treatment), and acknowledging that it is entirely possible (and reasonable) to have more than "one correct answer." The purpose of this article is to highlight key elements related to these two core concepts and discuss genuine barriers that currently exist on the pathway to creating such assessments. These include acknowledging situations of uncertainty, creating clear frameworks that define progressive levels of clinical reasoning skills, providing validity evidence to increase the defensibility of such assessments, considering the comparative feasibility with other forms of assessment, and developing strategies to evaluate the impact of these assessment methods on future learning and practice. The authors recommend that concerted efforts be directed toward these key areas to help advance the field of clinical reasoning assessment, improve the clinical care decisions made by current and future physicians, and have positive outcomes for patients. It is anticipated that these and subsequent efforts will aid in reaching the goal of making future assessment in medical education more representative of current-day clinical reasoning and decision making.

In an age where practicing physicians have access to an overwhelming volume of clinical information and are faced with increasingly complex medical decisions, the ability to execute sound clinical reasoning is essential to optimal patient care. If the skill of clinical reasoning is so essential to physician performance then surely this core competency receives targeted attention in formative and summative evaluations in medical education. But does it?

To date, evaluation of medical trainees is based primarily on the foundational elements of knowledge and comprehension, especially in the context of certifying examinations. Much less attention is placed on the formative or summative assessment of essential, higher-order functions associated with the skill of clinical reasoning. As such, a gap exists between what we recognize is required for clinical performance and what is currently being assessed. Recognizing this deficiency, the Royal College of Physicians and Surgeons of Canada (RCPSC)[1] and the Accreditation Council for Graduate Medical Education in the United States[2] have requested that specific methods be developed to assess the clinical reasoning competency of medical trainees in formative and summative assessments. The first and most critical step in the development of any new method of assessment is to have a clear understanding of the specific construct being assessed and the concepts that characterize that construct.

We propose two concepts that are philosophically paramount to the future assessment of clinical reasoning in medicine: assessment in the context of "uncertainty" (when, despite all of the information that is available, there is some doubt as to the best diagnosis, investigation, or treatment), and acknowledging that it is entirely possible (and reasonable) to have more than "one right answer." Attempting to bring each of these realities of clinical medicine to the field of *assessment* of medical trainees creates genuine challenges. Our purpose in this article is to highlight key elements related to these two core concepts and identify current barriers in the field of assessment related to the adoption of these concepts. We anticipate that these and subsequent efforts will aid in reaching the goal of making future assessment in medical education more representative of current-day clinical reasoning and decision making.

Assessment in the Context of Uncertainty

Traditionally, medical *assessment* has taken place in the context of *certainty*. Trainees are primarily tested on their ability to remember a significant number of facts, patterns, associations, and algorithms. For example, multiple-choice questions (MCQs) have a single correct answer, and short-answer questions (SAQs) usually have a series of acceptable answers that are interpreted as having a single meaning. Even in objective structured clinical examinations (OSCEs), the scoring rubrics are often associated with either giving points for the initial steps of information gathering, generating a shortlist of investigations, pattern recognition, and the final (and often straightforward algorithm driven) steps of management; or awarding

S. Cooke is clinical associate professor, Department of Paediatrics, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada.

J.F. Lemay is professor, Department of Paediatrics, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada.

Correspondence should be addressed to Suzette Cooke, Department of Paediatrics, Alberta Children's Hospital, 2888 Shaganappi Trail NW, Calgary, AB, Canada, T3B 6A8; telephone: (403) 955-3017; e-mail: suzette.cooke@ahs.ca.

Acad Med. 2017;92:746–751.
First published online January 31, 2017
doi: 10.1097/ACM.0000000000001559

points associated with performing a focused physical examination. In OSCEs associated with standardized patients, if candidates remember to ask the seven cardinal questions of a patient's history and can recall simple characteristics associated with a diagnosis or basic elements of a counseling session, they will often be able to secure enough points to pass the station. Candidates in OSCEs are rarely asked how they are applying that information, including which key features they are using to develop a working diagnosis, what they are attempting to rule in and out from requested investigations, and how they have synthesized their decision about a particular diagnosis, investigation, or treatment. Similarly, candidates are not routinely expected to show how they have integrated the principles of Bayes' theorem, which considers disease prevalence, probability, sensitivity, and specificity.[3] Furthermore, candidates' thinking processes are rarely probed with respect to contextual features that may have affected their decision making and how they intend to address or manage atypical aspects of the case. In addition, current scoring keys for most formats of assessment (MCQ, SAQ, OSCE, Medical Council of Canada Qualifying Examination Part 1 and 2, United States Medical Licensing Examination) are typically derived from "gold standard" textbooks or journals or other evidence-based sources such as national medical association statements and specialty-specific guidelines. Put simply—it is much easier to create assessments for knowledge-based tests for which there is certainty and clear evidence. So, why should assessment change now?

Current methods of assessment should change now because they do not sufficiently reflect the realities of clinical reasoning processes required in applied medicine. Durning et al[4] have observed that "our current models of assessing clinical reasoning often fall short of what is actually occurring." The RCPSC CanMeds 2015 framework for physician competencies also recognizes this deficit.[5] New competencies, including the ability to use clinical reasoning to manage clinical complexity and clinical uncertainty, have been acknowledged as being central to the medical expert role. These new inclusions recognize that while some real clinical cases are "clear-cut" and that basic knowledge, comprehension, and application will suffice, there are many cases that contain multiple variables and a significant degree of *uncertainty*. Fargason et al[6] have stated that "medicine is full of uncertainty." "For many students, the development of the ability to accept uncertainty and to deal with it effectively is 'the most difficult adaptational task confronting them.'"[7] In 2012, Hull et al[8] suggested, "Training novices for competence in routine situations is insufficient; we also need to know they are equipped to contend with the myriad, unpredictable, non-routine situations they will confront in clinical practice."

A physician's tolerance of uncertainty influences her or his clinical practice. Physicians who are less tolerant of uncertainty are more likely to order excessive diagnostic testing and additional empiric treatment.[9] This behavior increases health care costs and places patients at risk for experiencing adverse events.[10,11] In addition, physicians who are less comfortable with uncertainty are less likely to discuss this uncertainty with their patients and are less likely to engage in shared decision making.[12] We suggest that many common situations in clinical medicine, alone or in combination, contribute to this *uncertainty*. These include circumstances when:

- The amount of "currently available" information for the case is limited.
- Key (critical) information is missing.
- Some information is contributory but not "discriminating."
- Some information is unexpected and/or may be contradictory.
- The information available may be weighted differentially.
- Key features available do not readily "fit" a diagnosis.
- There is either no evidence or poor evidence for a particular path or course of action.

Furthermore, we propose that contextual features of the case and human factors that are unique to each patient be routinely considered and made explicit. These details may help create increased "definition" within zones of uncertainty and support more optimal decision making:

- Contextual factors that may be highly relevant to decision making must enter the equation for successful patient care. These include age, gender, medical acuity and complexity, geographical location of the health care team and of the patient and family, available resources (including expertise, time, and physical and financial resources), health care system factors, system language and communication issues, cultural beliefs, and social circumstances.
- There is inherent variability in human beings. No two patients with the same condition are likely to present in exactly the same way; nor will patients respond to treatment (i.e., albuterol or chemotherapy) in identical fashion. Physicians must learn to recognize patient-specific factors and when it is appropriate to adjust the diagnosis, investigation, or therapeutic regimen based on the integration of these factors.

Therefore, while knowledge and comprehension continue to be necessary in present-day clinical reasoning (and assessment), they are not sufficient. Clinicians must be able to discern what information is available, what is most relevant, and what key pieces are missing. They must *apply* medical data as well as contextual features. Some of these factors have been revealed in the clinical decision-making processes of resident physicians.[13] Clinicians must also appreciate potential patient-specific factors. At any given point in time, a clinician must be able to *analyze* the case, *synthesize* a working diagnosis (even if it is not the final diagnosis), and make active decisions regarding investigations and treatment. Although clinicians are always working toward the "goal" of *increased confidence*, they must develop the clinical reasoning processes necessary to manage the initial (and intermittent) situations of *uncertainty*. Recognizing and accepting these situations of uncertainty (after collecting and analyzing the best available knowledge, evidence, and experience) is essential to a physician's ability to move forward, communicate honestly with patients, and effectively manage their care.

When these cognitive processing skills receive primary focus, there is improved discernment of information, better decision making, and a predictable improvement in the desired outcomes of safe, effective, and efficient patient care. Medical education research is revealing that medical expertise is closely linked to an individual's performance in clinical reasoning and diagnostic performance.[14] In addition, Fargason et al[6] have concluded that in order to remain viable, academic health centers must broaden their educational goals so that trainees can learn to manage uncertainty, and especially uncertainty associated with managing care delivery. Therefore, medical educators have a moral responsibility to embed these higher-order clinical reasoning skills of *application*, *analysis*, and *synthesis* into both teaching and assessment and to do this within real-world contexts of *uncertainty*.

Can There Really Be More Than One Correct Answer?

In some situations in clinical medicine, there is enough information about the patient and sufficient evidence in the literature (or via expert consensus) to adhere to a single, clear path of diagnosis, investigation, or treatment. In other situations, the information available at the time may be weak, or, given the information available and the context of the case, there may be insufficient evidence to dictate one specific choice. Depending on the case and the context, two (or more) approaches to establishing the diagnosis, creating a plan for investigation, or prescribing treatment may be safe and reasonable.

The Script Concordance Test (SCT)[15] is an interesting testing method that supports two of the core principles we feel are paramount to the assessment of modern-day clinical reasoning: testing in contexts of uncertainty and respecting the possibility that there may be more than one acceptable path (e.g., correct answer).

A clinical example is illustrated in Appendix 1. The base scenario describes an 11-month-old infant who has a three-day history of fever, irritability, decreased oral intake, and one episode of vomiting. On the basis of an interpretation of the clinical history, physical examination, and the results of the laboratory investigations, this infant most likely has a urinary tract infection (pyelonephritis) and secondary complications of dehydration, metabolic acidosis, and hypernatremia. However, at this stage, the evidence for the *underlying diagnosis* is not conclusive. The infant could have a primary infection located elsewhere in the body including meningitis, appendicitis, myocarditis, or an occult abscess. The physician must decide on a working diagnosis and a treatment plan. Most prudent physicians would characterize this infant as "sick" and, at minimum, prescribe intravenous (IV) or intramuscular antibiotics. They would also seek to ensure close follow-up of clinical status and laboratory results. Several management options are provided in Part One of this case (see Appendix 1). On the basis of the reasoning and rationale provided above, it is inappropriate to either discharge this patient without further treatment (option a) or discharge and provide oral antibiotics (option b). Management options c and d offer reasonable antibiotic treatment, with option d being more conservative (admission to hospital and broad-spectrum coverage). Most would agree, however, that both options (c and d) are safe and acceptable.

As the case evolves and more information becomes available (Part Two), pyelonephritis is confirmed; however, after 48 hours the child still has a fever. Some children (especially infants) may still have fever on day 2 of this type of infection, but for the majority of children, the fever will have resolved by day 3. Therefore, the decision to treat with IV ampicillin and observe for another 24 hours (option c) is reasonable. It is equally acceptable to conduct further tests (option d) to rule out other possible diagnoses. This may include performing a lumbar puncture to assess for the possibility of meningitis or obtaining an abdominal ultrasound to assess for appendicitis or an occult abdominal abscess. This is a classic example where a clinician at each stage has two or more safe and reasonable options and he/she may select either one (as did the SCT panel of experts for this case).[16] Such situations occur frequently in clinical medicine. Durning et al[4] state that "there are often multiple correct paths and what defines expertise is performance within a set of boundary conditions." We suggest taking this one step further—rather than avoiding these clinical situations in the domain of assessment, we have an opportunity to embrace them and to explore options for presenting these equally valid choices to the health care team as well as the patient and family. Building alignment with each of these stakeholders is key to creating a supportive and unified health care plan.

In summary, we propose that respecting the possibility of the existence of "more than one correct answer" reflects *clinical reality* and will ultimately make the assessment of clinical reasoning more valid and reliable.

The Context of Uncertainty: Current Barriers

There are four different barriers to creating assessments that target decision making in the context of uncertainty.

The first barrier is *acknowledging situations of uncertainty* in clinical medicine by all stakeholders involved with assessment. This includes candidates, examiners, test developers, medical schools, program directors, clinical preceptors, certifying bodies, and the general public. By exposing and highlighting these situations, we come closer to accepting that gray zones will always exist, and it is better to acknowledge, confront, and engage these situations in formative and summative assessments rather than steering away from them.

The second barrier is that a *clear framework* must be developed to identify the progressive levels of clinical reasoning skills needed to manage varying degrees of uncertainty in clinical decision making. In the early phases, trainees need to be able to identify "key features" embedded within a case. Then they need to apply appropriate "weight" to each of these features within the context of the particular case (as some features will more heavily influence a case at any given point in time). That weighting may be influenced by the quality and the reliability of the information being received. It may include a sign or symptom or other element (i.e., allergy) on history taking, a laboratory investigation, or a treatment option. This information may also be interpreted and prioritized by applying the principles of Bayes' theorem.[3] Weighting of case features may be further enhanced by an awareness of possible cognitive biases (i.e., anchoring, confirmation bias, premature closure).[17] The consideration of cognitive biases refers to creating a habit of personal reflection to help appreciate the personal, system, and environmental factors that may be contributing to a physician's thought process.

The next step is often absent or, when present, underestimated. The trainee must evaluate not only the patient's medical data but also the contextual factors that "color the picture" of every clinical case. The textbook case is the "prototype" case, but each real case is an "exemplar" with typical and atypical features, as well as additional contextual variables that are entrenched and must be considered.[18,19] Teasing out the specific patient case factors within the milieu of specific contextual factors requires the clinical reasoning skill of *analysis*: being able to select the most relevant medical and contextual aspects of a case, determining which ones will predominate, and synthesizing the implications for clinical decision making. Using such a framework allows those creating assessments to build valid cases and questions that help test and measure the progression of the skill of clinical reasoning in contexts of uncertainty.

A third barrier, that is real and extremely relevant to certification bodies, relates to the *defensibility of the assessment method*. In each instance of assessment, "an interpretative argument must be built for which evidence is collected in support of proposed inferences."[20] More specifically, "the interpretation of scores must be linked to a network of theory, hypotheses, and logic which are presented to support or refute the reasonableness of the desired interpretations."[21] On this point, there appears to be some significant debate in the clinical reasoning literature. Traditionally, a consensus response has been sought, producing a single answer. More recently, SCT has demonstrated the potential to consider the "cumulative response" of a panel of experts. Charlin and van der Vleuten[15] contend that, particularly in areas of uncertainty, the cumulative response of the expert panel helps to illuminate a variety of potential acceptable responses including modal responses and responses near to the modal response.[15] We support this general concept. However, a valid criticism of this scoring method is the absence of even a limited peer review of the responses provided by the experts.[22] Experts can also misinterpret a question or make an error. Because their responses ultimately form the scoring key, it is vital to screen that key prior to the assessment. We propose that the diversity of expert responses in clinical reasoning assessment be supported within a context of peer oversight to ensure that all responses reflect plausible, rational responses and safe patient care.

A fourth barrier, facing relevant assessment of clinical reasoning in contexts of uncertainty, relates to *feasibility*. In both training assessments and summative evaluations, training program directors and specialty examination board members must be mindful of a multitude of variables[23–25]:

- What will the format of the assessment be? What resources (i.e., computers, standardized patients, mannequins, etc.) does the specific format require? Is there a reliance on computer technology? Is it reliable? Is it secure?
- Who will develop the assessment? What expertise is required? Where and how will this occur? What is the cost?
- Where will the assessment take place? Is the location available, easily accessible, secure, and appropriate for the assessment method? Can the location accommodate a large number of candidates? What is the cost of the facility?
- When will the assessment take place? How long does it take? Does this fit reasonably well with the availability of candidates, examiners, and administrators?
- Does the administration of the assessment require candidates or examiners to incur significant costs related to travel and accommodation?
- Does the timing, location, number of candidates, and number of examiners necessitate the creation of more than one examination (i.e., Exam A and Exam B)? Does this threaten/create risk for unintended test exposure/contamination? Is sequestering of candidates necessary?
- What is the time and financial cost to develop, administer, score, distribute, and securely store results of the assessment?

Presently there are few targeted methods for formative or summative assessment of clinical reasoning in graduate training. For instance, efforts are being taken in some local and national training programs to embed elements of clinical reasoning into OSCE examinations. While the OSCE can be a very natural and appropriate place to conduct some aspects of clinical reasoning assessment, it is difficult to conduct sufficient sampling due to other competencies also tested by the OSCE method (i.e., other medical expert, communication, collaboration, and manager skills), as well as the limited number of stations that can be administered in any one OSCE examination. In addition, OSCEs require on-site presence of both candidates and examiners, are time consuming to develop and administer, and are expensive.[26] The SCT method offers several advantages from the perspective of feasibility. The SCT may take place at any time and in any location that can offer secure and reliable Internet access. SCT can be administered in 60 to 90 minutes and produces instantaneous scoring. The SCT format supports the integration of visual images (i.e., x-rays, rashes, an ECG), video (i.e., seizures), and audio (i.e., heart or lung sounds); this format helps to create clinical realism. Assuming all candidates take the test at the same time, only one form of the test is required. Candidates incur minimal costs. Content expertise and test design skills are required to develop the SCT; however, these are no more difficult or costly than those associated with developing MCQ or SAQ assessments. We suggest the SCT can be best introduced at the formative stages of trainee assessment while undergoing piloting within summative trainee assessments. To assist with faculty development, SCT could simultaneously be introduced to preceptors to help support clinical reasoning teaching while concurrently preparing faculty members in their own professional development and maintenance of certification (summative assessments).

Another criterion to consider, when developing a robust assessment, is *the impact on future learning and practice.*[4] Medical educators are well aware that assessment drives learning.[27] New assessment methods stimulate new foci of learning. We support efforts made to improve and refine current methods of assessment (i.e., the OSCE and SCT), as well as the development of other methods for clinical reasoning assessment. Why? Because we recognize that as these methods are integrated into formative and summative assessments, trainees will be motivated to learn and excel in this critical expert skill. We anticipate that these effects will aid trainees in developing clinical reasoning skills, managing uncertainty, making better clinical decisions for their patients during their training, and with increasing experience over the course of their career.

Concluding Remarks

With advancements in health care worldwide, infants, children, and adults are surviving illness and living with increasingly complex health care conditions. Concurrently, medical research continues to push the boundaries of our clinical knowledge, and as we blaze "new trails" we naturally encounter uncharted territory and uncertainty. During this process, however, we often discover more than one safe and acceptable approach to managing that uncertainty.

In this article, we have attempted to highlight the need to incorporate two principles: embracing uncertainty, and acknowledging the potential for more than one acceptable path into formative and summative assessments of clinical reasoning. It is important to recognize that genuine barriers exist on the pathway to creating such assessments. These include acknowledging situations of uncertainty, creating clear frameworks that define progressive levels of clinical reasoning skills, providing validity evidence to increase the defensibility of such assessments, considering the comparative feasibility with other forms of assessment, and developing strategies to evaluate the impact of these assessment methods on the future learning and practice. We propose that it is important that concerted efforts be directed toward these key areas to help advance the field of clinical reasoning assessment, improve the clinical care decisions made by current and future physicians, and have positive outcomes for patients and families.

Acknowledgments: The authors wish to thank Tanya Beran, Harish Amin, and Amonpreet Sandhu for their contributions to this manuscript. The authors also wish to acknowledge Linda Beatty for her administrative support in preparing this manuscript.

Funding/Support: This research was supported by a grant from the Royal College of Physicians and Surgeons of Canada: Fellowship Studies in Medical Education (Dr. Suzette Cooke, recipient).

Other disclosures: None reported.

Ethical approval: Reported as not applicable.

References

1 Royal College of Physicians and Surgeons of Canada. CanMEDS 2005 physician competency framework. www.royalcollege.ca/portal/page/portal/rc/common/documents/canmeds/resources/publications/framework_full_e.pdf. Published 2005. Accessed March 10, 2012.
2 Swing SR. The ACGME outcome project: Retrospective and prospective. Med Teach. 2007;29:648–654.
3 Medow MA, Lucey CR. A qualitative approach to Bayes' theorem. Evid Based Med. 2011;16:163–167.
4 Durning SJ, Artino AR Jr, Schuwirth L, van der Vleuten C. Clarifying assumptions to enhance our understanding and assessment of clinical reasoning. Acad Med. 2013;88:442–448.
5 Frank JR, Snell L, Sherbino J, ed. CanMEDS 2015 physician competency framework. Ottawa, Ontario, Canada: Royal College of Physicians and Surgeons of Canada; 2015. www.royalcollege.ca/rcsite/documents/canmeds/canmeds-full-framework-e.pdf. Accessed November 19, 2016.
6 Fargason CA Jr, Evans HH, Ashworth CS, Capper SA. The importance of preparing medical students to manage different types of uncertainty. Acad Med. 1997;72:688–692.
7 Knight J. Medical Student: Doctor in the Making. New York, NY: Appleton-Century-Crofts; 1973.
8 Hull L, Arora S, Aggarwal R, Darzi A, Vincent C, Sevdalis N. The impact of nontechnical skills on technical performance in surgery: A systematic review. J Am Coll Surg. 2012;214:214–230.
9 Luther VP, Crandall SJ. Ambiguity and uncertainty: Neglected elements of curricula? Acad Med. 2011;86:799–800.
10 McGaghie WC, Thompson JA. America's best medical schools: A critique of the U.S. News & World Report rankings. Acad Med. 2001;76:985–992.
11 Page S. The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies. Princeton, NJ: Princeton University Press; 2007.
12 Tancredi DJ, Bertakis KD, Jerant A. Short-term stability and spread of the U.S. News & World Report primary care medical school rankings. Acad Med. 2013;88:1107–1115.
13 McBee E, Ratcliffe T, Picho K, et al. Consequences of contextual factors on clinical reasoning in resident physicians. Adv Health Sci Educ Theory Pract. 2015;20:1225–1236.
14 Schuwirth LW, van der Vleuten CP. General overview of the theories used in assessment: AMEE guide no. 57. Med Teach. 2011;33:783–797.
15 Charlin B, van der Vleuten C. Standardized assessment of reasoning in contexts of uncertainty: The script concordance approach. Eval Health Prof. 2004;27:304–319.
16 Cooke S, Lemay JF, Beran T, Sandhu A, Amin H. Development of a method to measure clinical reasoning in pediatric residents: The Pediatric Script Concordance Test. Creat Educ. 2016;7:814–823.
17 Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Acad Med. 2003;78:775–780.
18 Bordage G. Prototypes and semantic qualifiers: From past to present. Med Educ. 2007;41:1117–1121.
19 Norman G, Young M, Brooks L. Non-analytical models of clinical reasoning: The role of experience. Med Educ. 2007;41:1140–1145.
20 Kane MT. An argument-based approach to validity. Psychol Bull. 1992;112:527–535.
21 Downing SM. Validity: On meaningful interpretation of assessment data. Med Educ. 2003;37:830–837.
22 Lineberry M, Kreiter CD, Bordage G. Threats to validity in the use and interpretation of Script Concordance Test scores. Med Educ. 2013;47:1175–1183.
23 McAleer S. Choosing assessment instruments. In: Dent JA, Harden RM, eds. A Practical Guide for Medical Teachers. 3rd ed. Edinburgh, UK: Elsevier; 2009:318–324.
24 Schuwirth LWT, van der Vleuten CPM. How to design a useful test: The principles of assessment. In: Swanick T, ed. Understanding Medical Education: Evidence, Theory and Practice. Association for the Study of Medical Education. West Sussex, UK: Wiley-Blackwell; 2010:195–206.
25 Norcini J, Anderson B, Bollela V, et al. Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 Conference. Med Teach. 2011;33:206–214.
26 Brown C, Ross S, Cleland J, Walsh K. Money makes the (medical assessment) world go round: The cost of components of a summative final year objective structured clinical examination (OSCE). Med Teach. 2015;37:1–7.
27 van der Vleuten CP. The assessment of professional competence: Developments, research and practical implications. Adv Health Sci Educ Theory Pract. 1996;1:41–67.

Appendix 1

Script Concordance Test Scenario: A Method for Assessment of Clinical Reasoning in Contexts of Uncertainty

Base Scenario: An 11-month-old infant is seen in the emergency department with a three-day history of fever (up to 102.2°F) and irritability. He had one episode of nonbilious vomiting yesterday. He has had no diarrhea. The mother has not noticed any problems with his breathing. She has bathed him daily and not noticed a rash. She reports he has been taking ~16 ounces of formula daily and 8 ounces of water or juice. His intake of solids has been a little less than usual. On physical examination you note the following:

Vital Signs:

Heart rate + 134 beats/minute	Temperature = 103.3°F
Respiratory rate = 28 breaths/minute	Blood pressure = 90/55 mm Hg
O_2 saturation = 97% in room air	

He is inactive during your physical examination. He has dry mucous membranes. His capillary refill time is three seconds centrally and peripherally. His physical examination is otherwise unremarkable. Blood work and a capillary blood gas are drawn. A peripheral IV is inserted. He receives a bolus of normal saline (20 mL/kg) after which time he voids. Repeat vital signs show a heart rate of 115 bpm, respiratory rate of 24/min, blood pressure of 92/56 mm Hg, and temperature of 103.1°F. He is now a bit more interactive. His investigations reveal the following:

CBCd:	Chemistry Panel:	Other Labs:
Hgb = 13 g/dL	Na = 148 mEq/L	C-reactive protein = 36 mg/L
WBCs = 18,600 mm^3	K = 5.5 mEq/L	CBG: pH = 7.29, CO_2 = 42 mm Hg
Neutrophils = 10,400 mm^3	BUN = 8.1 mEq/L	(Base deficit = −7)
Lymphocytes = 4,800 mm^3	HCO_3 = 14 mEq/L	
Bands = 200 mm^3	Creatinine = 0.7 mg/dL	

Urinalysis = 10–20 WBCs, positive for nitrites, no RBCs, no protein, 2+ ketones.
A blood culture is pending. A lumbar puncture was not performed.

Part One

Considering the treatment plan of...	This treatment plan becomes...				
a) Discharge home, parents to return if they have concerns	−2	−1	0	+1	+2
b) Discharge home, prescribe high-dose oral amoxicillin and arrange to call family once culture results are available	−2	−1	0	+1	+2
c) Give one dose of IM ceftriaxone, discharge home, and request family return to the Emergency Department in 24 hours	−2	−1	0	+1	+2
d) Admit, treat with IV ampicillin and cefotaxime (pending 48-hour cultures)	−2	−1	0	+1	+2

Part Two

Case Evolution: After 48 hours of treatment with IV ampicillin and cefotaxime, his blood culture is negative. His urine culture is positive for *Escherichia coli* sensitive to amoxicillin, nitrofurantoin, and sulfamethoxazole-trimethoprim. His oral intake (fluids and solids) is improving. His parents report his mood is better although his energy is still low.

Vital Signs:

Heart rate = 92 beats/minute	Blood pressure = 96/56 mm Hg
Respiratory rate = 18 breaths/minute	Temperature = 101.5°F
O_2 saturation = 98% in room air	

His physical examination today is unremarkable.

Considering the treatment plan of...	This treatment plan becomes...				
a) Discharge home, parents to return if they have concerns	−2	−1	0	+1	+2
b) Discharge home, prescribe high-dose oral amoxicillin × 10 days and follow up in one week with the family doctor or pediatrician	−2	−1	0	+1	+2
c) Remain in hospital, continue IV ampicillin and reassess in 24 hours	−2	−1	0	+1	+2
d) Remain in hospital, continue IV ampicillin and pursue further investigations	−2	−1	0	+1	+2

Scale −2 = Contraindicated, −1 = Not indicated, 0 = Neither contraindicated nor indicated, +1 = Indicated, +2 = Strongly indicated (essential)