REACT: Rapid Evaluation Assessment of Clinical Reasoning Tool

Brian D. Peterson, MD[1], Charles D. Magee, MD, MPH[2], James R. Martindale, Ph.D.[2], Jessica J. Dreicer, MD[2], M. Kathryn Mutter, MD, MPH[2], Gregory Young, MD[2], Melissa Jerdonek Sacco, MD, MS[2], Laura C. Parsons, MD, MPH[2], Stephen R. Collins, MD, MSc[2], Karen M. Warburton, MD[2], and Andrew S. Parsons, MD, MPH[2]

[1]University of Virginia Health, Charlottesville, VA, USA; [2]University of Virginia School of Medicine, Charlottesville, VA, USA.

INTRODUCTION: Clinical reasoning encompasses the process of data collection, synthesis, and interpretation to generate a working diagnosis and make management decisions. Situated cognition theory suggests that knowledge is relative to contextual factors, and clinical reasoning in urgent situations is framed by pressure of consequential, time-sensitive decision-making for diagnosis and management. These unique aspects of urgent clinical care may limit the effectiveness of traditional tools to assess, teach, and remediate clinical reasoning.

METHODS: Using two validated frameworks, a multidisciplinary group of clinicians trained to remediate clinical reasoning and with experience in urgent clinical care encounters designed the novel Rapid Evaluation Assessment of Clinical Reasoning Tool (REACT). REACT is a behaviorally anchored assessment tool scoring five domains used to provide formative feedback to learners evaluating patients during urgent clinical situations. A pilot study was performed to assess fourth-year medical students during simulated urgent clinical scenarios. Learners were scored using REACT by a separate, multidisciplinary group of clinician educators with no additional training in the clinical reasoning process. REACT scores were analyzed for internal consistency across raters and observations.

RESULTS: Overall internal consistency for the 41 patient simulations as measured by Cronbach's alpha was 0.86. A weighted kappa statistic was used to assess the overall score inter-rater reliability. Moderate reliability was observed at 0.56.

DISCUSSION: To our knowledge, REACT is the first tool designed specifically for formative assessment of a learner's clinical reasoning performance during simulated urgent clinical situations. With evidence of reliability and content validity, this tool guides feedback to learners during high-risk urgent clinical scenarios, with the goal of reducing diagnostic and management errors to limit patient harm.

J Gen Intern Med 37(9):2224–9
DOI: 10.1007/s11606-022-07513-5

Received September 15, 2021
Accepted March 25, 2022
Published online June 16, 2022

INTRODUCTION

Clinical reasoning encompasses the process of data collection, synthesis, and interpretation to generate a working diagnosis, facilitating management decisions. The bulk of research on teaching and assessment focuses on diagnosis, namely the process by which a differential diagnosis is generated and narrowed through data gathering.[1–5] Recently, Cook et al.[6] described management reasoning as a necessary companion to diagnostic reasoning, accounting for patient preferences, societal values, logistical constraints, and resource availability when making testing and treatment decisions for patients. Urgent clinical situations, those in which the patient's clinical condition is rapidly declining, require accelerated decision-making with respect to both diagnosis and management.[7,8]

Errors in clinical reasoning among practicing clinicians are common, estimated to occur in up to 10–15% of hospitalized patient encounters.[9] Learners who struggle with urgent clinical situations may be labeled as not recognizing "sick vs not sick" or as lacking in communication skills or clinical knowledge. Clinical reasoning deficits have been commonly identified among struggling medical trainees described in single-center remediation programs.[10,11] The University of Colorado reports that clinical reasoning was the primary deficit in 25–30% of residents and 40–45% of medical students referred to their remediation program over a 6-year period.[12] Over a 4-year period, the University of Virginia identified that 34% of learners referred to a Graduate Medical Education (GME) remediation program struggled with clinical reasoning.[13] A true estimate of prevalence data is difficult to establish as validated clinical reasoning assessment tools have limitations.[14]

Dual-process theory is a commonly understood cognitive model for clinical reasoning wherein decision-making occurs through a combination of system 1 (heuristic processes) and system 2 (analytical processes).[15] Urgent clinical situations are contextualized well with dual-process theory, requiring heuristics and efficient analytic reasoning for time-sensitive diagnosis as well as rapid assessment, stabilization, and management prior to the determination of a diagnosis. To improve clinical reasoning in urgent situations, algorithms for specific

clinical scenarios have been developed to facilitate management reasoning and improve patient outcomes. For example, the Advanced Cardiovascular Life Support algorithms guide management in "code" situations and offer a method for analytical diagnosis while the patient is being resuscitated (e.g., the H's and T's).[16] However, the majority of urgent clinical situations do not reach this final common pathway, and as such, require nuanced clinical reasoning without aid from established algorithms for management and diagnosis.

Situated cognition theory (SCT) provides an attractive optic for the assessment of clinical reasoning performance relative to the myriad interacting factors impacting formative evaluation in urgent clinical situations. Rencic et al.[17] proposed a conceptual framework that considers six clinical reasoning performance assessment elements: the clinician or assessee, patient, rater, assessment method, task, and environment. Through this conceptual lens, direct observation of clinical reasoning performance offers the most authentic assessment, but requires rigor to identify and manage the numerous interacting factors that influence clinical reasoning performance. Assessment of clinical reasoning in urgent clinical situations, characterized by high acuity or decompensation of a patient's clinical status, is particularly challenging due to the unplanned nature and multitude of uncontrollable factors that may have distracting or detrimental effects, including the subjugation of educational goals for the urgent provision of care. A simulated patient encounter that offers an urgent patient care situation in a high-fidelity simulation environment therefore offers an ideal surrogate model, allowing for control of many factors across each of the six assessment elements.

Herein, we describe a novel tool for formative assessment and feedback of learner performance during urgent clinical situations that require rapid, time-sensitive diagnostic and management reasoning. This behaviorally anchored tool, known as REACT (Rapid Evaluation Assessment of Clinical Reasoning Tool), was designed by content experts based on domain-specific frameworks to guide feedback to learners during high-risk urgent clinical scenarios, with the goal of reducing diagnostic and management errors.

METHODS

In 2016, the Committee on Seeking Competence through Help (COACH) was formed. COACH is a unique peer support program at the University of Virginia (UVA) aimed to help medical learners who are referred for, or who request, help with clinical performance. Since its creation, COACH has worked with more than 100 trainees in 14 different departments with generally positive outcomes. Subsequently, UVA School of Medicine implemented a clinical remediation program employing the same framework and much of the same personnel. In 2018, a subcommittee was formed to focus on strategies to identify and coach learners who struggle with clinical reasoning. This group, composed of primary care and specialist clinician educators from internal and hospital medicine, emergency medicine, pediatrics, anesthesiology, family medicine, critical care, and obstetrics and gynecology, met monthly to review struggling learners, discuss best practices in clinical reasoning assessment and remediation, and review the clinical reasoning literature.[18,19] The subcommittee serves as a pool of clinical reasoning coaches, available as needed, to provide one-on-one coaching to struggling trainees and students. To address an identified need, the group began a joint effort in late 2020 to design an evidence-based tool to assess and provide formative feedback to learners during urgent patient care situations.

The REACT tool (Fig. 1) was designed by a multidisciplinary group of clinician educators from the COACH subcommittee with expertise in teaching and assessing clinical reasoning across both undergraduate medical education (UME) and GME. REACT was named to represent the rapid patient evaluation required during urgent clinical scenarios. The group met on four occasions to design the tool, first identifying evidence-based domains of diagnosis and management reasoning specific to urgent patient care situations and then associating a range of behavioral anchors with each domain. The group began with two validated frameworks: (1) the Society to Improve Diagnosis in Medicine's (SIDM) Assessment of Reasoning Tool (ART) to assess clinical reasoning during oral presentations and (2) the Association of American Medical College's (AAMC) Entrustable Professional Activity 10 (EPA 10) designed to formatively assess a learner's recognition of patients requiring urgent care. A priority of the group was to design a tool applicable to learners in a variety of clinical settings consistent with patient care provided by multiple specialties.

Thammasitboon et al.[20] and SIDM developed and validated the Assessment of Reasoning Tool (ART) to facilitate clinical teaching for oral presentations and clinical reasoning, specifically assessing the learner's proficiency in the domains of data gathering, interpretation, synthesis, and metacognition. This behaviorally anchored tool provides a general framework for assessing and correcting errors in clinical reasoning. Observing a learner in the context of a clinical scenario allows for the assessment of nonverbal and tonal cues, which vary depending on situational stressors.[21] In contrast, the AAMC designed a toolkit for assessing a range of clinical competencies among learners in real-world settings. These core competencies, termed EPAs, include validated proficiencies expected of medical students prior to starting residency.[22] EPA 10 focuses on "recognizing a patient requiring urgent or emergent care and initiating evaluation and management." This toolkit introduces a standard set of behaviors expected in the management of urgent clinical scenarios.

A pilot study was performed at UVA during the 2021 Intern Readiness Course (IRC) for 87 fourth-year medical students who are preparing to transition into internal medicine,

Essential Tasks		Assessment of Behaviors		
Learner Function	**Specified Task**	**Instructions: Circle the box corresponding to the learner's observed behavior**		
Collecting • Data gathering *– recognition of urgent or emergent clinical scenario*	• **Collect/Report** history and exam data in hypothesis-directed manner • **Recognize** patient and disease specific factors as potential etiologies of decompensation • **Recognize** severity of clinical problem and contextual signs of urgency or emergency	○ Non-focused history and exam ○ Included extraneous information ○ Missed key findings ○ Did not recognize contextual clues of urgency	○ History and exam reflect potential diagnoses ○ Limited recognition of urgent contextual clues ○ Included limited pertinent positive and negative findings	○ Logical history and exam for potential diagnoses ○ Questions assessed likelihood of specific diagnoses ○ Full recognition of urgency contextual clues ○ Prioritization on pertinent positive and negative findings
Interpreting • Diagnostic reasoning *– differential diagnosis*	• **Generate** prioritized differential diagnosis of most likely, less likely, unlikely in urgent clinical situations	○ Differential diagnosis missing likely or "can't miss" diagnoses in urgent situations ○ Includes inappropriate diagnoses	○ Differential diagnosis included likely and "can't miss" diagnoses in urgent situations but missed key diagnoses ○ Inappropriate rank-order of diagnoses	○ Accurately ranked differential diagnosis including key, likely, and "can't miss" diagnoses in urgent situations ○ Prioritized urgent diagnoses appropriately
Managing • Management reasoning *- initial management option selection* *- response to dynamic information*	• **Direct** evaluation and treatment towards high priority diagnoses • **Initiate** management in patient with urgent decompensation • **Recognize** need to escalate patient care	○ Directed evaluation and treatment toward unlikely/unimportant diagnoses ○ Did not evaluate or treat most likely urgent diagnoses ○ Did not evaluate for response to initial management plan	○ Major focus of evaluation and treatment on likely and urgent diagnoses ○ Included non-essential testing ○ Evaluated for response to initial management plan	○ Efficiently directed management towards most likely and urgent diagnoses ○ Deferred tests directed towards less likely or less important diagnoses ○ Evaluated for response to initial interventions
Communicating • Patient-centered care *- communicate care plan and goals of care* *- achieve a shared mental model*	• **Communicate** with health care team members according to role and responsibility to improve task efficiency • **Maintain** focus and two-way communication with patient or surrogate decision-maker in urgent clinical situations	○ Did not engage health care team members ○ Did not communicate care plan with patient/family ○ Did not clarify patient's goals of care	○ Limited engagement of health care team members to improve task efficiency ○ Limited effort to communicate with patient/family but used medical jargon or failed to ensure understanding ○ Limited clarification of patient's goals of care	○ Fully engaged health care team members according to role and responsibility to improve task efficiency in urgent situation ○ Effectively communicated care plan with patient/family and ensured understanding ○ Effective clarification of patient's goals of care
Reflecting • Metacognition *- reflection of urgent or emergent clinical scenarios*	• **Demonstrate** the ability to think about one's own thinking (metacognition) in urgent clinical situations • **Mitigate** cognitive tendencies or emotional/situational factors influencing clinical decision making	○ Unaware of cognitive tendencies or emotional/situational factors that may have influenced decision-making ○ Unable to mitigate cognitive tendencies or emotional/situational factors	○ Limited awareness of cognitive tendencies or emotional/situational factors that may have influenced decision-making ○ Limited mitigation of cognitive tendencies or emotional/situational factors	○ Full awareness of cognitive tendency or emotional/situational factors that may influence decision-making ○ Effectively mitigates reactions to urgent clinical situations

Figure 1 Rapid Evaluation Assessment of Clinical Reasoning Tool (REACT).

psychiatry, family medicine, emergency medicine, or anesthesia residencies. A core goal of the IRC is to provide fourth-year medical students with the skills to appropriately respond to and manage common urgent clinical situations such as hypotension, chest pain, hypoxemia, or altered mental status. Much of this education is accomplished through simulated scenarios with manikins. Each student plays the role of an intern in at least two unique case scenarios, and the scenario is curated by a nurse with relevant clinical experience. At the conclusion of the simulation, students are immediately led through debriefing exercises with a clinician who directly observed the simulation. The simulations for the 2021 IRC were recorded. Table 1 provides a description of each case. An independent, multidisciplinary group of clinicians with no additional training in the clinical reasoning process observed these recordings and scored each medical student's performance using REACT. A scoring system was added to the behavioral anchors in order to analyze the tool's performance for internal consistency across raters and observations. REACT scores were generated using a 3-point scale for each behavioral domain, with a maximum total score of 15 and a minimum score of 5.

Determination of optimal sample size for the study utilized estimates based on minimizing measurement error, both in the number of observations and the number of raters used.[23] Cronbach's alpha was used to assess the internal consistency across the group of raters. Inter-rater reliability for the overall rating score among the group of raters was assessed using the weighted kappa statistic. All analyses were performed using SPSS v28.[24]

The UVA Institutional Review Board reviewed this project and determined that it met the criteria for exempt review (ref # 4234).

RESULTS

REACT is comprised of four learner functions essential to the clinical reasoning process during urgent patient care: data collecting, interpreting, managing, and communicating. A fifth learner function, reflecting, highlights the centrality of metacognition in effective clinical reasoning.[25] For each function, specified tasks are described. A range of exemplar behavioral anchors are described for each function to allow for formative feedback.

Seven raters comprising clinicians from internal and hospital medicine, obstetrics and gynecology, pediatric critical care, emergency medicine, and anesthesiology scored 41 recorded case scenarios representing 41 individual students. Each rater scored the same 41 case scenarios in full. Internal consistency as assessed by Cronbach's alpha was measured for the summed overall rating (score) for the 41 video clips and was .86, a value considered sufficient for high-stakes assessment.[26] Due to the ordinal nature of the ratings, a weighted kappa statistic was used to measure inter-rater reliability which for overall ratings was .56, generally interpreted as a moderate degree of agreement.[27] Supplemental Table 1 provides domain-specific weighted kappa data and Supplemental Table 2 provides descriptive statistics of domain-specific scores for individual raters.

Table 1 Description of Clinical Cases Requiring Rapid Evaluation

Case	Patient description	Objectives
Asthma exacerbation	67-year-old female with sudden onset shortness of breath	• Recognize tachycardia, tachypnea, and hypoxia • Identify asthma exacerbation as a most likely diagnosis • Initiate management for asthma exacerbation
Ruptured ectopic pregnancy	28-year-old female with nausea, vomiting, and lower abdominal pain	• Demonstrate an organized approach to a patient with hypotension • Recognize ectopic pregnancy as a possible cause of abdominal pain and hypotension • Call for obstetric consultation and initiate hypovolemic shock management
Myocardial infarction	72-year-old female with slight pressure in her epigastrium	• Recognize acute coronary syndrome may present atypically in female patients • Obtain EKG and call "STEMI" alert • Initiate management of acute coronary syndrome
Transfusion reaction	45-year-old male with dizziness, nausea, abdominal pain, and shortness of breath	• Recognize hypoxia, hypotension, tachycardia, and fever as possible reactions to transfusion • Stop the blood transfusion and initiate management of possible transfusion reaction
Anaphylaxis	70-year-old female with dizziness, shortness of breath, and pruritis	• Recognize hypotension and tachycardia • Identify anaphylaxis as a potential etiology with recent antibiotic administration • Initiate a care plan for the decompensating patient
Septic shock	76-year-old male with altered mental status	• Recognize fever, hypotension, tachycardia, and altered mental status • Prioritize septic shock as the most likely diagnosis • Initiate stabilization management for septic shock
Cardiac arrest	57-year-old male with shortness of breath and chest pain	• Create a differential for acute chest pain and shortness of breath • Recognize PEA arrest and create a differential for the causes • Initiate management for cardiac arrest
COVID-19 pneumonia	65-year-old female with cough and fever	• Recognize symptoms of hypoxia and fever as potential viral pneumonia • Initiate management of worsening hypoxia
Heart failure exacerbation	58-year-old male with hypotension and chest pain	• Demonstrate an organized approach to a patient with hypotension • Collect an organized history to determine potential causes • Initiate appropriate management for hypotension
Hypertensive emergency	52-year-old male with confusion	• Demonstrate an organized approach to a patient with altered mental status • Recognize hypertensive emergency and hypertensive encephalopathy • Initiate management of hypertension and recognize risks of rapid blood pressure reduction
Acute alcohol withdrawal syndrome	52-year-old male with agitation	• Recognize alcohol withdrawal syndrome and ensure patient and staff safety • Demonstrate understanding of behavioral emergency medications • Initiate management of alcohol withdrawal syndrome
Hypoglycemic seizure	46-year-old female with altered mental status	• Demonstrate an organized approach to a patient with altered mental status • Consider hypoglycemia on the differential • Initiate glucose replacement therapy

EKG electrocardiogram
STEMI ST elevation myocardial infarction
PEA pulseless electrical activity

DISCUSSION

To our knowledge, REACT is the first tool specifically designed for formative assessment of a learner's clinical reasoning performance during simulated urgent clinical situations. Built on the strength of validated instruments specific to clinical reasoning and clinical urgency, REACT was thoughtfully designed by a multidisciplinary group of clinician educators with expertise in teaching and assessing clinical reasoning across the spectrum of UME and GME education. This approach provides evidence of content validity and our analysis demonstrates both moderate inter-rater reliability and a high degree of internal consistency of REACT to assess clinical reasoning performance in simulated urgent clinical situations. This was notably achieved with no additional rater training or standard setting and among a population of clinician educators from multiple medical specialties. This finding is a particularly intriguing observation in contrast to guidelines informing best practice for direct observation of clinical skills in medical education, recommending both rater and frame of reference training.[28]

SCT indicates that numerous potential variables may influence clinical reasoning performance in urgent clinical situations. Variables include those intrinsic to the clinician such as years of experience or training, as well as variables intrinsic to

the patient, the rater, the clinical reasoning task, and the environment.[17] For example, urgent clinical encounters can in part be defined by the need for early management and an accelerated response to dynamic information. Although these variables exist uncontrolled in authentic urgent clinical situations, implementation of REACT in a simulated urgent clinical situation affords an environment in which control over many of these variables generates an opportunity to isolate and measure variables of interest.

REACT, in fact, was not designed to directly assess the myriad relationships and interactions between variables that might affect the performance of clinical reasoning in urgent clinical scenarios. Rather, REACT is focused on the empiric tasks essential to the formative assessment of clinical reasoning performance in urgent clinical situations. It is therefore our hypothesis that REACT may not perform as well in authentic, real-life urgent clinical situations where the rapid interactions may at times compete with the educational goals, threatening the ability to achieve educational goals and objectives. In simulated urgent clinical situations, where educational goals maintain primacy, the numerous factors surrounding the clinician, patient, rater, assessment method, task, and environment are all controllable and modifiable. In a limited way, REACT provides a tool to isolate and study variance attributable to individual factors, creating an understanding of how each is associated with clinical reasoning performance. Such analysis may inform other potential relationships and interactions in non-urgent clinical environments.

Future directions include the study of REACT as a tool to predict the need for remediation among different levels of learners and as a summative assessment tool before and after clinical reasoning coaching. This includes the creation of a quantitative system with performance thresholds for scoring in addition to the current behavioral anchors. Friedman et al.[29] found that residents are more often over-confident in their diagnoses and more prone to errors in diagnostic reasoning. In contrast, medical students made more diagnostic errors but had the least confidence; attending physicians made mistakes, albeit with the highest confidence and accuracy. These findings suggest that opportunities for improved clinical reasoning exist throughout the UME and GME continuum and beyond. Future studies may also employ faculty development for raters on the basics of diagnostic and management reasoning.

CONCLUSION

REACT is a novel tool designed for formative assessment and feedback of a learner's clinical reasoning performance during urgent clinical situations. This tool can reliably serve as a guide to clinician educators in their assessment of learners and may assist in the identification of learners who struggle with clinical reasoning skills in high-risk urgent clinical scenarios, with the goal of reducing diagnostic and management errors. A study of validity in other learner populations is necessary to determine the broadest application for this formative assessment instrument.

Corresponding Author: *Andrew S. Parsons, MD, MPH; University of Virginia School of Medicine, Charlottesville, VA, USA (e-mail: asp5c@virginia.edu).*

Supplementary Information *The online version contains supplementary material available at https://doi.org/10.1007/s11606-022-07513-5.*

Declarations:

Conflict of Interest: *The authors have no conflicts of interest to disclose.*

REFERENCES

1. Gruppen LD. Clinical Reasoning: Defining It, Teaching It, Assessing It, Studying It. West J Emerg Med 2017;18(1):4-7.
2. Geoff R, Norman CPMvdV, David I Newble, Diana H.J.M Dolmans, et al. International Handbook of Research in Medical Education. Dordrecht, the Netherlands: Kluwer Academic Publishers, 2002.
3. Norman G. Research in clinical reasoning: past history and current trends. Med Educ 2005;39(4):418-27.
4. Norman GR, Eva KW. Diagnostic error and clinical reasoning. Med Educ 2010;44(1):94-100.
5. Patel JJ, Bergl PA. Diagnostic vs Management Reasoning. JAMA 2018;320(17):1818.
6. Cook DA, Sherbino J, Durning SJ. Management Reasoning: Beyond the Diagnosis. JAMA 2018;319(22):2267-68.
7. Lighthall GK, Vazquez-Guillamet C. Understanding Decision Making in Critical Care. Clin Med Res 2015;13(3-4):156-68.
8. Ozel F. Time pressure and stress as a factor during emergency egress. Safety Science 2001;38(2):95-107.
9. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. Am J Med 2008;121(5 Suppl):S2-23.
10. Yao DC, Wright SM. National survey of internal medicine residency program directors regarding problem residents. JAMA 2000;284(9):1099-104.
11. Dupras DM, Edson RS, Halvorsen AJ, Hopkins RH, Jr., McDonald FS. "Problem residents": prevalence, problems and remediation in the era of core competencies. Am J Med 2012;125(4):421-25.
12. Guerrasio J, Garrity MJ, Aagaard EM. Learner deficits and academic outcomes of medical students, residents, fellows, and attending physicians referred to a remediation program, 2006-2012. Acad Med 2014;89(2):352-8.
13. Warburton KM, Shahane AA. Mental Health Conditions Among Struggling GME Learners: Results From a Single Center Remediation Program. J Grad Med Educ 2020;12(6):773-77.
14. Lessing JN, Rendon P, Durning SJ, Roesch JJ. Approaches to Clinical Reasoning Assessment. Acad Med 2020;95(8):1285.
15. Monteiro SM, Norman G. Diagnostic reasoning: where we've been, where we're going. Teach Learn Med 2013;25 Suppl 1:S26-32.
16. American-Heart-Association. Advanced cardiovascular life support: Provider Manual. Dallas, Texas: 2016.
17. Rencic J, Schuwirth LWT, Gruppen LD, Durning SJ. Clinical reasoning performance assessment: using situated cognition theory as a conceptual framework. Diagnosis (Berl) 2020;7(3):241-49.
18. Parsons A, Warburton K. A novel clinical reasoning coaching program for the medicine learner in need. MedEdPublish, ed. 2019.
19. Parsons AS, Clancy CB, Rencic JJ, Warburton KM. Targeted Strategies to Remediate Diagnostic Reasoning Deficits. Acad Med 2021.
20. Thammasitboon S, Rencic JJ, Trowbridge RL, Olson APJ, Sur M, Dhaliwal G. The Assessment of Reasoning Tool (ART): structuring the conversation between teachers and learners. Diagnosis (Berl) 2018;5(4):197-203.
21. Ende J. Feedback in clinical medical education. JAMA 1983;250(6):777-81.

22. Obeso V BD, Aiyer M, Barron B, Bull J, et al. Core Entrustable Professional Activities for Entering Residency—EPA 10 Schematic: Recognize a Patient Requiring Urgent or Emergent Care and Initiate Evaluation and Management. Washington, DC: Association of American Medical Colleges; 2017.
23. Gwet KL. Handbook of Inter-Rater Reliability. 4th ed. Advanced Analytics LLC: 2014.
24. IBM-Corp. IBM SPSS Statistics for Windows Version 28. 28 ed. Armonk, NY: IBM Corp; 2021.
25. Marcum JA. An integrated model of clinical reasoning: dual-process theory of cognition and metacognition. J Eval Clin Pract 2012;18(5):954-61.
26. Bland JM, Altman DG. Cronbach's alpha. BMJ 1997;314(7080):572.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-74.
28. Kogan JR, Hatala R, Hauer KE, Holmboe E. Guidelines: The do's, don'ts and don't knows of direct observation of clinical skills in medical education. Perspect Med Educ 2017;6(5):286-305.
29. Friedman CP, Gatti GG, Franz TM, et al. Do physicians know when their diagnoses are correct? Implications for decision support and error reduction. J Gen Intern Med 2005;20(4):334-39.