

# Optimizing OSCE for clinical reasoning assessment: correlation of Clinical Reasoning Indicators-History Taking Scale with traditional checklist

Hanaa Elhoshy<sup>1</sup>, Doaa S. Zaky<sup>2</sup>, Nagwa E. Saad<sup>3</sup>, and Omayma Hamed<sup>4</sup>

<sup>1</sup>MD, Lecturer, Department of Medical Education, Faculty of Medicine, Alexandria University; Department of Medical Education, Armed Forces College of Medicine, Cairo, Egypt

<sup>2</sup>MD, Professor, Department of Internal Medicine, Al-Azhar Faculty of Medicine for Girls, Cairo, Egypt

<sup>3</sup>MD, Professor, Department of Internal Medicine, Armed Forces College of Medicine, Cairo, Egypt

<sup>4</sup>MD, Head, Department of Medical Education, Armed Forces College of Medicine, Cairo, Egypt

## Abstract

**Background:** Clinical reasoning is defined as a multidimensional construct which has multiple assessment methods available to measure it. The Clinical Reasoning Indicators—History Taking—Scale (CRI-HT-S) is a novel tool that assesses clinical reasoning during history taking on a 5-point Likert scale. This study aimed to measure clinical reasoning of undergraduate medical students using this novel tool and correlate its score with traditional checklist-based Objective Structured Clinical Examination (OSCE) scores. **Methods:** This study employed correlational design. A non-probability sample of 115 female undergraduate medical students in the final year of the integrated medical program for the bachelor's degree provided by the faculty of medicine for girls, Al-Azhar University, Cairo, participated in the study. After training the assessors, the study incorporated the Clinical Reasoning Indicators—History Taking—Scale (CRI-HT-S) into the abdominal pain history-taking OSCE station as an additional assessment tool. Descriptive statistics were used for

quantitative data. The Pearson correlation test assessed the association between history-taking checklist scores and CRI-HT-S scores, with a significance level set at  $p < 0.05$ . **Results:** The clinical reasoning skills scores were variable. The highest scores belonged to “taking the lead” skill, and the lowest were checking with patients, collecting data and effectiveness of the conversation skills. There was a statistically significant weak positive correlation between history-taking checklist scores and the total CRI-HT-S score among undergraduate medical students, and results also exhibited weak positive correlations with four of the individual clinical reasoning indicators' scores. **Conclusions:** Study participants are relatively skillful during history-taking at taking the lead with the patient and recognizing and responding to relevant information. Incorporating the CRI-HT-S into OSCE provided a **structured approach** to measuring clinical reasoning in history-taking stations.

**Keywords:** Assessment, Tool, Clinical reasoning, History taking, OSCE, CRI-HT-S, Undergraduate

**Date submitted:** 15-April-2025

**Email:** Hanaa Elhoshy (hanaa.elhoushy@alexmed.edu.eg)

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-Share Alike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

## Background:

Clinical reasoning is defined as a multidimensional construct encompassing a skill to be developed, a cognitive process to be optimized, and an outcome reflecting the effectiveness of clinical decision-making.<sup>1</sup> From a skills-based perspective, clinical reasoning is viewed as a fundamental competency that clinicians acquire and enhance over time.<sup>2,3</sup> It

Citation: Elhoshy H, Zaky D, Saad N, and Hamed O. Optimizing OSCE for clinical reasoning assessment: correlation of Clinical Reasoning Indicators-History Taking Scale with traditional checklist. *Educ Health* 2025;38:227-236

**Online access:** [www.educationforhealthjournal.org](http://www.educationforhealthjournal.org)

DOI: 10.62694/efh.2025.327

Published by The Network: Towards Unity for Health

involves a structured set of abilities, including data gathering, hypothesis generation, problem representation, differential diagnosis formulation, diagnostic justification, and management planning.<sup>4</sup> As a process, clinical reasoning involves the cognitive steps and mechanisms clinicians use to assess and manage patient cases.<sup>5</sup> In addition, clinical reasoning can be defined through the

accuracy and quality of clinical decisions, including diagnosis, treatment planning, and patient management.<sup>6</sup>

Despite the recognition of its multifaceted complex nature, the skills-based perspective dominates assessment methods.<sup>7</sup> In undergraduate medical education, Non-Workplace-Based Assessments (NWBAs), such as multiple-choice questions (MCQs), extended matching questions (EMQs), and key feature examinations (KFEs), are commonly used to assess specific abilities, including differential diagnosis formulation, leading diagnosis selection, and management planning.<sup>1</sup> Additionally, assessments in simulated clinical environments, such as Objective Structured Clinical Examinations (OSCEs), provide a more interactive evaluation by incorporating standardized patients and structured observer ratings.<sup>8</sup> Workplace-Based Assessments (WBAs), such as direct observation using the Mini-Clinical Evaluation Exercise (mini-CEX), and global faculty evaluations, enable supervisors to assess core clinical competencies in real-world settings, including medical interviewing, physical examination skills, and clinical judgment.<sup>9</sup> Objective Structured Clinical Examinations (OSCEs) have been widely adopted as standardized assessments of clinical competence; however, they have been criticized for their limitations in evaluating clinical reasoning.<sup>10,11</sup> Their fragmented, time-constrained structure prioritizes reliability but fails to capture the complexity and integration required for reasoning assessment.<sup>11</sup> While modified formats, such as Task-Integrated OSCEs (TIOSEs) and clinically discriminatory checklists (CDCs), have demonstrated some potential for improvement, their effectiveness remains limited and underexplored in broader educational settings.<sup>10</sup> Traditional OSCEs primarily assess isolated clinical skills (e.g., history-taking, basic decision-making) but lack the contextual integration essential for reasoning.<sup>12</sup> Their station-based design fragments the clinical encounter, making it difficult to evaluate higher-order thinking processes. Checklist-driven assessments further reinforce procedural task adherence rather than deeper cognitive and metacognitive engagement, limiting their ability to measure true reasoning ability.<sup>11</sup>

Several modifications have been proposed to enhance clinical reasoning assessment within OSCEs. Task-Integrated OSCEs (TIOSEs) introduce sequential decision-making tasks to improve contextual realism, but their implementation and validation remain limited.<sup>13,14</sup>

Clinically Discriminatory Checklists (CDCs) have shown promise in capturing reasoning skills with greater specificity, though at the expense of lower reliability and increased standard error.<sup>15</sup> Alternative approaches, such as longitudinal, multi-station case progression or narrative-based assessments, may better reflect real-world diagnostic reasoning.<sup>14</sup> Additionally, technology-enhanced assessments, including computer-based simulations and think-aloud protocols, offer potential solutions for capturing cognitive processes while maintaining station efficiency.<sup>16</sup>

Experts aim to translate the process-oriented perspective of clinical reasoning into a skills-based approach by identifying observable behaviors during history-taking.<sup>17</sup> Using a grounded theory approach, expert assessors analyzed student encounters, focusing on indicators such as questioning patterns, symptom recognition, logical sequencing, and patient interaction.<sup>17</sup> For structured assessment of clinical reasoning based on these identified indicators, the Clinical Reasoning Indicators—History Taking—Scale (CRI-HT-S) is a novel tool designed to assess clinical reasoning based on eight observable clinical reasoning behavioral indicators during history taking.<sup>18</sup> These indicators measure a clinician's ability to lead the conversation, organize questions, recognize relevant information, and verify patient responses on a 5-point Likert scale, ranging from 1 (does not meet the criterion at all) to 5 (fully meets the criterion), providing a standardized, quantitative approach to assessing clinical reasoning in history taking.<sup>18</sup>

While the tool has demonstrated promise for assessing clinical reasoning during history taking, previous studies have not integrated this tool into Objective Structured Clinical Examination (OSCE) history taking stations, limiting its applicability in standardized high-stakes assessments. To address this gap, this study aims to assess the clinical reasoning skills of undergraduate medical students in history taking during internal medicine OSCE station using the CRI-HT-S and validate the CRI-HT-S by examining its correlation with OSCE scores. The current study is designed to answer three questions: (1) How do undergraduate medical students perform in clinical reasoning skills during history taking in an internal medicine OSCE station, as measured by the CRI-HT-S? (2) Are there differences in students' performance levels of clinical reasoning skills as measured by the CRI-HT-S? (3) What is the correlation between CRI-HT-

S scores and OSCE scores in assessing clinical reasoning skills during history taking?

### **Intervention**

The end of clerkship OSCE consisted of 5 stations. Each station involved a 15-minute student-simulated patient encounter: abdominal pain history taking station, neurological examination, cardiopulmonary examination station, cardiopulmonary interpretation linked station, and investigations interpretation station. The study incorporated the Clinical Reasoning Indicators-History Taking-Scale (CRI-HT-S) into the abdominal pain history taking OSCE station as an additional assessment tool. In addition to ensuring alignment with the OSCE blueprint, which encompassed various body systems, the abdominal pain history-taking station was deliberately selected due to its clinical relevance and diagnostic complexity. Abdominal pain is a common presenting complaint that necessitates consideration of a broad differential diagnosis involving multiple organ systems. Importantly, it enables the assessment of clinical reasoning from the initial stages of patient interaction—particularly during history taking—which aligns directly with the focus of this study and the strengths of the CRI-HT-S tool. Students' clinical reasoning abilities were evaluated using the CRI-HT S alongside the traditional OSCE checklist, which remained unchanged.

The internal medicine OSCE was conducted across three parallel circuits, each comprising five stations. Four of the stations in each circuit were assessed by a single examiner, while the history-taking station included two independent assessors—one using the standard checklist and the other applying the CRI-HT-S to evaluate clinical reasoning. Each station was allotted 15 minutes, leading to a total examination duration of 3 hours and 45 minutes per circuit (active station time). The examination was delivered in two sessions, one in the morning and one in the afternoon, each session lasted for three hours (active station time, transit time, and instructions time), separated by a one-hour break for assessors. In total, 18 assessors were involved in administering and scoring the stations across all circuits, ensuring structured and consistent evaluation throughout the OSCE process. The students' performance was evaluated simultaneously using the traditional history taking 10 items binary checklist by an expert clinical assessor. The clinical assessors assign scores based on whether an action or task is "performed" (1 point) or "not performed" (zero points) during an

Objective Structured Clinical Examination (OSCE) with a maximum score of 10 marks.

The CRI-HT-S was used by a trained clinical reasoning assessor to measure 8 indicators of clinical reasoning on a 5-point Likert scale, ranging from 1 (does not meet the criterion at all) to 5 (fully meets the criterion). The CRI-HT-S indicators represent fundamental clinical reasoning skills during history taking including: Taking the lead in the conversation, recognizing and responding to relevant information, specifying symptoms, asking specific questions that point to pathophysiological thinking, putting questions in a logical order, checking with the patient, summarizing, collecting data and effectiveness of the conversation.<sup>18</sup>

The study involved three trained assessors in clinical reasoning, who are faculty members at the Internal Medicine Department, as well as the Medical Education Unit at the Al-Azhar University, Cairo. Notably, these assessors were independent of the individual who employed the traditional history taking checklist and voluntarily participated in the study. Before expanding participation, each assessor was provided with a detailed explanation of the study's objectives and methodology. To enhance scoring consistency and minimize interrater variability, they completed three structured training sessions on the CRI-HTS scoring system. While the study adhered to the pre-validated five-point scoring rubric, particular attention was given during assessors' training to elaborating the conceptual distinctions between adjacent score points—especially between scores 2 and 3, and 3 and 4—to enhance inter-rater reliability. Assessors were provided with detailed explanations and examples illustrating the level of completeness, accuracy, and coherence required for each score point. By standardizing their interpretation of the scoring system assessors developed a shared understanding of the rating criteria, which enhanced consistency across judgments and contributed to the reliability and validity of the scoring process.

### **Methods:**

This study employed a correlational design, conducted at Al-Zahraa University hospital (Training hospital for the medical students of Faculty of Medicine for girls Al-Azhar University), Cairo, Egypt. The data were collected during the end of the Internal Medicine clerkship OSCE of fifth year undergraduate students in September 2023.

A non-probability sample of 115 female undergraduate medical students in the final year of the integrated medical program for bachelor’s degree provided by the Al-Azhar University participated in the present study. The sample size was calculated using G power 3.9.1.4 (developed by University Düsseldorf) to achieve a correlation of 0.285 between diagnostic accuracy and clinical reasoning scale based on the results of Park et al.<sup>5</sup> at 80% power and 0.05 significance.

The IBM Statistical Package for the Social Sciences (SPSS) software package, version 22.0 (Armonk, NY: IBM Corp.) was used for quantitative data analysis. Descriptive statistics including percentages, mean and Standard Deviation (SD) were used to present students’ scores. The Pearson correlation test was used to study the possible association between two variables (history taking checklist scores, and CRI-HT-S scores). A significance level of  $p < 0.05$  was set to assess whether correlations are statistically meaningful.

This study was conducted following ethical research principles, ensuring participant anonymity, confidentiality, voluntary participation, and informed consent. The study proposal was reviewed and approved by the Institutional Research Ethics Board of the Al-Azhar University, as well as obtaining administrative approval of Al-Azhar University. Participation in this study was entirely voluntary. All assessors and student participants were informed about the study's objectives, procedures, potential risks, and benefits before

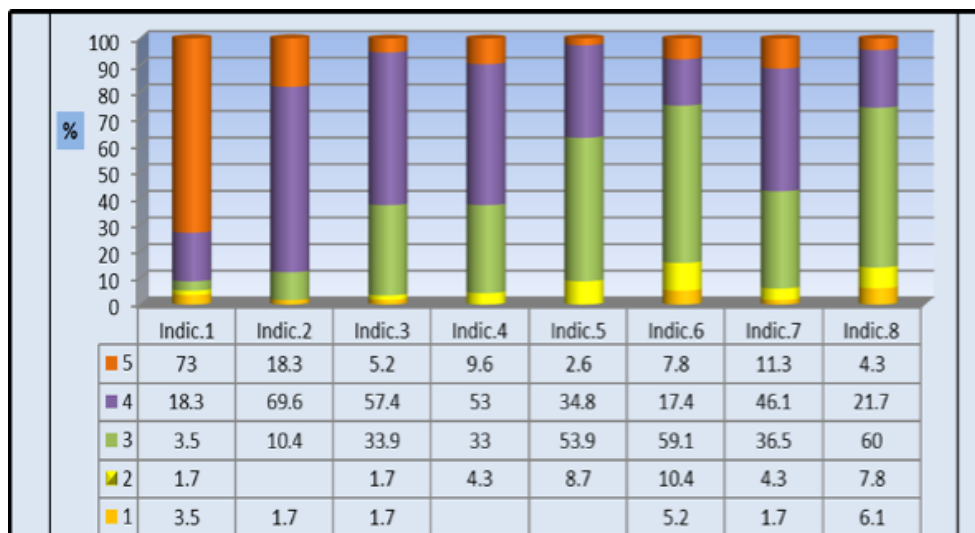
providing their oral consent. They were assured that their participation or withdrawal would not influence their academic standing or professional relationships. All data collected were kept strictly confidential and used solely for research purposes. Personal identifiers were removed or coded to maintain participant anonymity.

**Results:**

***Assessment of Clinical Reasoning Skills***

One hundred and fifteen undergraduate female undergraduate medical students in the final year of the integrated medical program for the bachelor’s degree provided by the faculty of medicine for girls, Al-Azhar University, participated in the study. The CRI-HT-S scores of the undergraduate medical students for each indicator of clinical reasoning were variable, as shown in Figure 1. The students demonstrated competence in taking the lead in the conversation (clinical reasoning indicator 1), with 73% of students performing it with excellence (scoring 5). The students also have high reasoning skills in terms of recognizing and responding to relevant information (clinical reasoning indicator 2), with 69.6% scoring 4, which indicates good performance. However, more than 15 % of the students are unable to check their ideas with the patients in fair quality (10.4% scoring 2 and 5.2 % scoring 1 in clinical reasoning indicator 6). Furthermore, more than 13% of the students lacked collecting data and effectiveness of the conversation skills skill (7.8% scoring 2 and 6.1 % scoring 1 in clinical reasoning indicator 8).

**Figure 1: Comparing the frequency (%) of students’ scores across the CRI-HT-S indicators**



The clinical reasoning indicators are ranked according to their mean scores in Table 1. The highest scores belonged to the “taking the lead” indicator and the lowest were checking with patients, collecting data and effectiveness of the conversation skills indicators, which highlights a potential gap in clinical education. The total CRI-

HT-S score of all students included in the study has a mean of  $29.03 \pm 3.8$  points out of 40 points that graded as 72% with range of 17 points as minimum score and 39 points as maximum score, while the history taking checklist scores ranged from 6-10 (out of 10) with a mean score of  $8.2 \pm 1.1$  out of 10 marks (graded 82%).

**Table 1: Mean clinical reasoning indicators scores using CRI-HT-S**

Clinical reasoning indicators		Number of students = 115		
		Mean $\pm$ SD	Range (min-max)	Rank
Indicator 1	Taking the lead in the conversation	$4.56 \pm 0.92$	1-5	1
Indicator 2	Recognizing and responding to relevant information	$4.03 \pm 0.6$	1-5	2
Indicator 3	Specifying symptoms	$3.63 \pm 0.69$	1-5	4
Indicator 4	Asking specific questions that point to pathophysiologic thinking	$3.68 \pm 0.71$	2-5	3
Indicator 5	Putting questions in a logical order	$3.31 \pm 0.67$	2-5	6
Indicator 6	Checking with patient	$3.1 \pm 0.89$	1-5	7
Indicator 7	Summarizing	$3.61 \pm 0.81$	1-5	5
Indicator 8	Collection of data and effectiveness of the conversation indicators	$3.1 \pm 0.84$	1-5	8
CRI-HT-S	Total	$29.03 \pm 3.8$	17-39	N/A

### ***Correlation of history taking checklists scores and the CR-HT-S scores***

There is a significant weak positive correlation between history taking checklist scores and the total score of CRT-HT-S ( $r=0.175$ ,  $P = 0.031$ ) in the undergraduate medical students as shown in Figure 2. Similarly, there is significant weak positive correlation between history taking checklist scores and indicator 3 (Specifying symptoms)  $r=0.179$  and

$p= 0.028$  (Figure 3), indicator 4 (Asking specific questions that point to pathophysiologic thinking)  $r= 0.247$  and  $P = 0.004$  (Figure 4), indicator 7 (Summarizing)  $r= 0.195$  and  $p= 0.018$  (figure 5), and indicator 8 (collecting data and effectiveness of the conversation skills)  $r =0.211$  and  $p=0.012$  (Figure 6). However, there is no significant correlation between history checklist scores and other indicators of clinical reasoning.

## Discussion

Among the various clinical reasoning skills assessed, the study participants demonstrated relatively strong proficiency in leading the conversation during history taking and in recognizing and appropriately responding to relevant information. These competencies are essential abilities to guide the patient through the consultation and extract fundamental details critical to the clinical decision-making process. Furthermore, the study revealed positive correlations between the history-taking checklist scores and the CRI-HT-S scores, highlighting the potential for incorporating the CRI-HT-S within the Objective Structured Clinical Examination (OSCE) as an effective tool for assessing clinical reasoning. The inclusion of the CRI-HT-S provides an additional layer of assessment, enhancing the robustness of clinical reasoning evaluation during OSCEs. This suggests that incorporating such tools can significantly contribute to a more comprehensive and nuanced assessment of students' clinical reasoning abilities.

### *Students' performance in reasoning skills during history taking as measured by the CRI-HT-S*

Consistent with Fürstenberg et al.,<sup>18</sup> results showed that students performed best in the skill of "Taking the Lead in the Conversation." Other skills, such as asking specific questions related to pathophysiological thinking, specifying symptoms, and checking with patients, received similar rankings in both this study and in the study by Fürstenberg et al.<sup>18</sup> Notably, the skill of recognizing and responding to relevant information achieved a higher mean score in the current study compared to the previous one, ranking as the second-highest skill overall. Similarly, the skill of summarizing was ranked higher and had a higher mean score in this study.

In contrast, the skills of organizing questions logically and collecting data, as well as the effectiveness of the conversation, had relatively lower mean scores in our study, with rankings much lower than in the previous study. Despite these differences in performance across specific skills, both studies resulted in similar total scores.<sup>18</sup> The overall mean score for the CRI-HT-S in the study by Fürstenberg et al.<sup>18</sup> was similar to that in our study, where the total mean score was slightly higher. This suggests that although individual competencies varied across different aspects of clinical reasoning, the overall proficiency of students in both studies remained comparable. A possible explanation is that

students compensated for weaker performance in certain areas by excelling in others, thus maintaining balanced total scores.

### *The relationship between students' performance in OSCEs and CRI-HT-S*

The significant weak correlation between history taking checklist scores, and the total CRI-HT-S score confirms the efficacy to some extent of the traditional clinical assessment via OSCE checklists to measure the clinical reasoning skills of undergraduate medical students. In addition, there are positive correlations between history taking checklist scores and specific clinical indicators scores such as: specifying symptoms, asking specific questions that point to pathophysiologic thinking, summarizing, and collecting data and effectiveness of the conversation skills. This contradicts previous research investigating the relationship between students' performance in OSCEs and clinical reasoning abilities. A prior study assessed clinical reasoning using a structured post-encounter table within the OSCE, in which students documented differential diagnoses and identified symptoms or signs that either supported or contradicted each diagnosis.<sup>11</sup> Clinical reasoning scores were calculated based on the total number of correct findings, as evaluated by two independent physician raters. A key finding of the previous study was that clinical reasoning scores were significantly correlated with GPA, but not with OSCE scores or clinical knowledge.<sup>11</sup> Additionally, total OSCE scores showed no significant correlation with GPA.<sup>11</sup>

### *Comparing CRI-HT-S to other tools of assessing clinical reasoning during OSCEs*

Besides the CRI-HT-S, prior research efforts were aimed at validating structured tools for assessing clinical reasoning during OSCEs. For instance, a previous study involving 136 fourth-year medical students incorporated a modified version of the Assessment of Reasoning Tool (ART) within OSCEs to evaluate reasoning across domains such as problem representation, differential diagnosis, and initial management. Although the ART-based tool did not assess reasoning within the context of history-taking, it demonstrated a positive correlation between clinical reasoning scores and OSCE performance.<sup>19</sup> In contrast, the incorporation of the CRI-HT-S within OSCEs uniquely captures clinical reasoning as it unfolds in real-time during the history-taking encounter. This embedded, performance-based approach allows for a more integrated assessment of diagnostic reasoning in the

early stages of clinical interaction, which is critical yet often underrepresented in traditional reasoning tools like the ART.

In contrast, a study employing the Oral Debrief (OD), a post-history-taking exercise, focused more on the reasoning process than the skill of data gathering itself. Using a rubric adapted from the Manchester Clinical Reasoning Tool (MCRT), the OD targeted cognitive aspects of clinical reasoning such as problem identification, diagnosis ranking, and management planning.<sup>20</sup> While the OD proved effective, it diverged from our study in both timing and scoring. The OD was assessed separately from the history-taking encounter, whereas the current study approach incorporated the CRI-HT-S directly within the history-taking station. This methodological difference may explain variations in observed correlations, as our tool assessed reasoning as it unfolded during the patient encounter, offering a skills-integrated perspective, rather than a post hoc analysis.

Other tools, such as Script Concordance Tests (SCTs) and Clinical Reasoning Problems (CRPs), were evaluated alongside OSCEs to explore their relationship with diagnostic reasoning. However, no significant correlation was found between SCT/CRP scores and OSCE performance, suggesting these tools assess different cognitive dimensions of reasoning.<sup>21</sup> This highlights the advantage of CRI-HT-S, which captures real-time reasoning during clinical interactions.

Another study attempted to embed clinical reasoning assessment into a simulation-based OSCE using remote scoring, but faced logistical barriers such as legal constraints on video recording and time limitations. As a result, the evaluation relied on traditional checklist components rather than dedicated reasoning instruments, limiting its ability to measure reasoning independently.<sup>22</sup> In contrast, our study demonstrates that clinical reasoning can be feasibly and effectively assessed within the standard OSCE structure through the CRI-HT-S, without requiring external resources or major procedural changes.

Group-based history-taking exercises have also been used to evaluate reasoning, with findings showing a strong link between diagnostic accuracy and the ability to extract relevant clinical data and develop a broad differential diagnosis. However, such approaches were not embedded in OSCEs, limiting their scalability for high-stakes

assessments.<sup>23</sup> The current study using CRI-HT-S within an OSCE setting addresses this gap by offering a standardized and objective method for assessing diagnostic reasoning in individual learners.

Finally, observational tools like the Observation Rating Tool (ORT) and the Post-Encounter Rating Tool (PERT) have been implemented to evaluate clinical reasoning in history-taking contexts during internal medicine clerkships. Despite their utility, these instruments were not integrated into OSCE formats, which limits their application in summative, high-stakes evaluations.<sup>24</sup> Embedding the CRI-HT-S directly within an OSCE station supports a more holistic and pragmatic approach to assessing clinical reasoning in a structured and time-efficient manner.

Based on the findings of this study, it is recommended that medical schools implement **routine formative assessments** using CRI-HT-S throughout clerkships, enabling students to **develop and refine their reasoning abilities before high-stakes examinations**. Additionally, integrating **CRI-HT-S into summative OSCEs** would enhance **the reliability and validity of clinical reasoning evaluation**, ensuring that reasoning skills are assessed alongside technical competencies. Faculty development initiatives should also focus on **training examiners in reasoning-based assessment** to provide **structured feedback** that fosters students' diagnostic thinking.

As this study serves as a **correlational validation study of CRI-HT-S**, future research should explore its **longitudinal impact on clinical reasoning development**, its **applicability across diverse medical curricula**, and its **effectiveness in predicting real-world diagnostic performance**. Expanding validation studies across multiple institutions would further establish **CRI-HT-S as a standardized tool** for clinical reasoning assessment, ultimately enhancing the quality of **clinical education and patient care**.

## Conclusions

According to the CRI-HT-S mean total score, the overall clinical reasoning performance of the students was moderate to high. However, the clinical reasoning factors/indicators sub-scores revealed variable levels of performance of specific clinical reasoning skills among the participating students. Undergraduate medical students are relatively skillful during history taking at “taking

the lead with the patient” and “recognizing and responding to relevant information”; while “checking with patients” and “collecting data and effectiveness of the conversation skills” were defective. These findings highlight a **significant gap in clinical reasoning education**, emphasizing the need for **more effective clinical training methods** to ensure all clinical reasoning skills are covered.

A significant weak positive correlation between history taking checklist scores and the total score of CRT-HT-S, as well as a significant weak positive correlation between the history taking checklist scores and four specific clinical reasoning indicators (specifying symptoms indicator, asking specific questions that point to pathophysiologic thinking, summarizing, collecting data and effectiveness of the conversation skills) validates its incorporation into OSCE assessments to provide a **structured approach** to measuring clinical reasoning in history-taking stations. By integrating **CRI-HT-S within OSCE checklists**, examiners can systematically evaluate **students' ability to gather, analyze, and synthesize patient information**, ensuring that reasoning skills are assessed alongside

technical competencies. Future improvements should focus on **refining assessment criteria, providing faculty training on reasoning-based evaluation, and implementing structured feedback mechanisms** to enhance students' diagnostic reasoning skills within the OSCE framework.

#### **Acknowledgements:**

The authors gratefully acknowledge the valuable feedback and support of the OSCE examiners during the study process: Prof Dr. Seham S. Alsaedy, professor of internal medicine, Al-Azhar Faculty of Medicine for girls, Cairo, Egypt, Prof. Dr. Fatema M. Kotb, assistant professor of internal medicine Al-Azhar Faculty of Medicine for girls, Cairo, Egypt, and Dr. Sara M. Elhadad, lecturer of internal medicine, Al-Azhar Faculty of Medicine for girls, Cairo, Egypt. The authors also extend our sincere thanks to all study participants for their time and cooperation, without whom this work would not have been possible.

#### **References**

1. Daniel M, Rencic J, Durning SJ, Holmboe E, Santen SA, Lang V, et al. Clinical reasoning assessment methods: a scoping review and practical guidance. *Academic Medicine*. 2019;94(6):902-912.<https://doi.org/10.1097/acm.0000000000002618>.
2. Cumming A, Ross M. The tuning project for medicine–learning outcomes for undergraduate medical education in Europe. *Medical Teacher*. 2007;29(7):636-641.<https://doi.org/10.1080/01421590701721721>.
3. Modi JN, Anshu, Gupta P, Singh T. Teaching and assessing clinical reasoning skills. *Indian Pediatrics*. 2015;52:787-794.<https://doi.org/10.1007/s13312-015-0718-7>.
4. Plackett R, Kassianos AP, Mylan S, Kambouri M, Raine R, Sheringham J. The effectiveness of using virtual patient educational tools to improve medical students' clinical reasoning skills: a systematic review. *BMC Medical Education*. 2022;22(1):365.<https://doi.org/10.1186/s12909-022-03410-x>.
5. Connor DM, Durning SJ, Rencic JJ. Clinical reasoning as a core competency. *Academic Medicine*. 2020;95(8):1166-1171.<https://doi.org/10.1097/acm.0000000000003027>.
6. Audétat M-C, Laurin S, Dory V, Charlin B, Nendaz MR. Diagnosis and management of clinical reasoning difficulties: Part I. Clinical reasoning supervision and educational diagnosis. *Medical Teacher*. 2017;39(8):792-796.<https://doi.org/10.1080/0142159x.2017.1331033>.
7. Young ME, Thomas A, Lubarsky S, Gordon D, Gruppen LD, Rencic J, et al. Mapping clinical reasoning literature across the health professions: a scoping review. *BMC Medical Education*. 2020;20(1):107.<https://doi.org/10.1186/s12909-020-02012-9>.

8. Thampy H, Willert E, Ramani S. Assessing clinical reasoning: Targeting the higher levels of the pyramid. *Journal of General Internal Medicine*. 2019;34:1631-1636.<https://doi.org/10.1007/s11606-019-04953-4>.
9. Rencic J, Durning SJ, Holmboe E, Gruppen LD. Understanding the assessment of clinical reasoning. In: Wimmers PF, Mentkowski M, editors. *Assessing Competence in Professional Performance across Disciplines and Professions*. Cham: Springer International Publishing; 2016. p. 209-235.
10. Dewan P, Khalil S, Gupta P. Objective structured clinical examination for teaching and assessment: Evidence-based critique. *Clinical Epidemiology and Global Health*. 2024;25:101477.<https://doi.org/10.1016/j.cegh.2023.101477>.
11. Park WB, Kang SH, Myung SJ, Lee Y-S. Does objective structured clinical examinations score reflect the clinical reasoning ability of medical students? *The American Journal of the Medical Sciences*. 2015;350(1):64-67.<https://doi.org/10.1097/maj.0000000000000420>.
12. Talha KA, Mra A, Selina F, Aung T, Rao CV, Maung H. Evaluation of performance in different types of Objective Structured Clinical Examination (OSCE) stations among Undergraduate Medical Students. *Borneo Journal of Medical Sciences* 2017;11(1):24-31.<http://dx.doi.org/10.51200/bjms.v11i1.635>.
13. Hassan S. Task integrated objective structured clinical examination (TIOSCE): A modified OSCE. *Education in Medicine Journal*. 2012;4(1).<http://dx.doi.org/10.5959/eimj.v4i1.15>.
14. Harden RM, Lilley P, Patricio M. An introduction to the OSCE. In: Harden RM, Lilley P, Patricio M, editors. *The definitive guide to the OSCE: The Objective Structured Clinical Examination as a performance assessment*: Elsevier Health Sciences; 2015. p. 1-12.
15. Myung SJ, Kim JW, Kim CW, Kim DH, Eo E, Kim JH, et al. Effect of limiting checklist on the validity of objective structured clinical examination: A comparative validity study. *Medical Teacher*. 2024;1-7.<https://doi.org/10.1080/0142159x.2024.2430364>.
16. Jawaid M, Bakhtiar N, Masood Z, Mehar A-K. Effect of paper-and computer-based simulated instructions on clinical reasoning skills of undergraduate medical students: a randomized control trial. *Cureus*. 2019;11(11).<https://doi.org/10.7759/cureus.6071>.
17. Haring CM, Cools BM, van Gorp PJ, van der Meer JW, Postma CT. Observable phenomena that reveal medical students' clinical reasoning ability during expert assessment of their history taking: a qualitative study. *BMC Medical Education*. 2017;17:1-9.<https://doi.org/10.1186/s12909-017-0983-3>.
18. Fürstenberg S, Helm T, Prediger S, Kadmon M, Berberat PO, Harendza S. Assessing clinical reasoning in undergraduate medical students during history taking with an empirically derived scale for clinical reasoning indicators. *BMC Medical Education*. 2020;20:1-7.<https://doi.org/10.1186/s12909-020-02260-9>.
19. Siegelman J, Bernstein L, Goedken J, Lewin L, Schneider J, Ward M, et al. Assessment of clinical reasoning during a high stakes medical student OSCE. *Perspectives on medical education*. 2024;13(1):629-634.<https://doi.org/10.5334/pme.1513>.
20. Régent A, Thampy H, Singh M. Assessing clinical reasoning in the OSCE: pilot-testing a novel oral debrief exercise. *BMC Medical Education*. 2023;23(1):718.<https://doi.org/10.1186/s12909-023-04668-5>.
21. Dory V, Charlin B, Vanpee D, Gagnon R. Multifaceted assessment in a family medicine clerkship: a pilot study. *Family Medicine*. 2014;46(10):755-760. Available from <https://www.stfm.org/familymedicine/vol46issue10/Dory755>.
22. Bußenius L, Harendza S. A simulation-based OSCE with case presentation and remote rating—development of a prototype. *German Medical Science Journal for Medical Education*. 2023;40(1):Doc12. <https://doi.org/10.3205/zma001594>.

23. Lai J-H, Cheng K-H, Wu Y-J, Lin C-C. Assessing clinical reasoning ability in fourth-year medical students via an integrative group history-taking with an individual reasoning activity. *BMC Medical Education*. 2022;22(1):573.<https://doi.org/10.1186/s12909-022-03649-4>.
24. Haring CM, Klarwater CC, Bouwmans GA, Cools BM, van Gurp PJ, van der Meer JW, et al. Validity, reliability and feasibility of a new observation rating tool and a post encounter rating tool for the assessment of clinical reasoning skills of medical students during their internal medicine clerkship: a pilot study. *BMC Medical Education*. 2020;20:1-7.<https://doi.org/10.1186/s12909-020-02110-8>.