Using Group History-Taking and Individual Reasoning to Identify Shortcomings in Clinical Reasoning for Medical Students

Journal of Medical Education and Curricular Development
Volume 11: 1–7

DOI: 10.1177/23821205241280946

Kuan-Hao Cheng[1,2,#], Chi-Yu Lee[1,2,#], Yih-Jer Wu[1,2] and Ching-Chung Lin[1,2]

[1]Department of Internal Medicine, MacKay Memorial Hospital, Taipei, Taiwan. [2]Department of Medicine, MacKay Medical College, New Taipei City, Taiwan.

ABSTRACT

OBJECTIVE: It is crucial that teaching faculties determine and remain informed of medical school learners' clinical reasoning competence. We created an innovative assessment method for fourth-year medical students to identify deficiencies in various components of their clinical reasoning ability.

METHODS: This was a cross-sectional observational study of fourth-year medical students' reasoning assessments from 2019 to 2022. Teams of four–five trainees questioned standardized patients in clinical scenarios, including fever, abdominal pain, and weight loss. They then individually documented key information to reflect comprehension of patient problems. Trainees were tasked with differentiating diagnoses and associated statuses and reaching the most likely diagnosis along with two tentative diagnoses. The correlations observed between 2020 and 2022 for abdominal pain were analyzed using student *t*-tests.

RESULTS: A total of 177 students participated in this study. Across the scenarios, there was no significant difference in key information representation scores (56%–58%). Reasoning ability scores were 49% for fever, 57% for abdominal pain, and 61% for weight loss. A comparison between 2020 and 2022 revealed a significant improvement in the objective structured clinical examination scores and differential diagnoses ($P < .01$). Shortcomings included brief chief complaint duration, lack of detailed presentation, and insufficient description of negative information. Differential diagnosis and diagnostic justification were inadequate for acute and chronic conditions, and disease location clarity within the organ system was lacking. On average, students presented two correct diagnoses.

CONCLUSIONS: Fourth-year medical students exhibited inadequate reasoning abilities, particularly in fever and abdominal pain scenarios, with deficiencies in hypothesis generation and differential diagnosis. Group history-taking with individual reasoning assessment identified students' shortcomings and provided faculty feedback to improve their teaching strategies.

KEYWORDS: group history-taking, clinical reasoning, medical students, standardized patient, individual reasoning

RECEIVED: May 25, 2024. **ACCEPTED:** August 4, 2024.

TYPE: Original Research Article

FUNDING: The authors received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

[#]These authors contributed equally to this article.

CORRESPONDING AUTHOR: Ching-Chung Lin, Department of Medical Education, MacKay Memorial Hospital, No. 92, Sec. 2, Chung-Shan North Road, Taipei, Taiwan.
Email: sunny.lin56@msa.hinet.net

Introduction

Clinical reasoning, an essential skill and core ability of medical students, presents challenges for clinical teachers in defining and delineating their processes.[1] Clinical reasoning requires both conscious and unconscious cognitive operations to work with the patient's unique circumstances, preferences, and the characteristics of the practice environment. Expertise is correlated with multiple coordinated representations that are adaptable to diverse situations.[2] Eva KW emphasizes the significance of analytical and non-analytic reasoning strategies as guiding frameworks.[3]

The content of clinical reasoning is complex, encompassing a comprehensive process and theoretical basis extensively debated in the literature. Gruppen LD et al identified multiple components of clinical reasoning, including (1) information gathering (IG), (2) hypothesis generation (HG), (3) forming a problem representation (PR), (4) generating a differential diagnosis (DD), (5) selecting a leading or working diagnosis (LD), (6) providing a diagnostic justification (DJ), and (7) developing a management or treatment plan (MT).[4] Therefore, an innovative method to assess this ability is needed when discussing clinical reasoning processes, and it is crucial to precisely define the aspects of clinical reasoning we are teaching, assessing, and studying.[5] These results can serve as references for teaching strategies in clinical diagnosis.

Numerous methods are available to assess trainees' clinical reasoning abilities. This assessment can take place in either a school or hospital, with the reliability and validity of different teaching and assessment strategies varying depending on the reasoning components. Daniel et al demonstrated the

effectiveness of assessing different components in simulated clinical environments, such as the objective structured clinical examination (OSCE). In these assessments, the examinees executed different clinical tasks while interacting with standardized patients (SPs), and their performance was evaluated through observer ratings and written notes. The results revealed that all clinical reasoning components achieved a good to very good status.[6]

Regular assessment of medical students' clinical reasoning abilities is important. Sunju Im et al observed insufficient clinical reasoning skills in third-year medical students, necessitating training in problem identification, reasoning skills, and accurate record keeping.[7] Charlin B et al proposed embedding clinical reasoning training in scripts to refine learners' illness scripts with clinical experience.[8]

Based on our previous literature, we created a group history-taking with individual reasoning activity at Medical College for fourth-year medical students (M4s). The findings included the average history-taking score, average key information number, average diagnosis number, and average correct diagnosis rate, which are clinical reasoning abilities.[9] Although we understand the importance of clinical reasoning problems in learners, most teachers cannot identify the shortcoming of its various components in students. A novel aspect of this study is the use of an easy method to determine these shortcomings in M4 students' clinical reasoning ability. We identified the knowledge gap addressed in this article assessing different components of M4 clinical reasoning ability. We need to design a new tool because traditional methods cannot easily evaluate these abilities. We used group history-taking with an individual reasoning method to help teachers assess more effectively, and the results provided valuable feedback to our school's M4 teaching faculty.

Table 1. The setting of group history-taking with individual reasoning to assess clinical reasoning ability.

Process	Subject steps	Learning outcome focus
1. Create test topics	Create clinical scenarios for M4 students' group test	Scenarios including abdominal pain, fever, and weight loss
2. Group history-taking with individual reasoning	The trainees were taking turns to ask question while gathering clinical information from the standardized patients	Trainees were expected to write down 12 key information for clinical reasoning, one most likely with two tentative diagnosis including reasons
3. Feedback	Post-reasoning test immediate feedback with the answer	Self-learning from our school made clinical reasoning videos

Materials and Methods

Developing an innovative assessment method for different components of clinical reasoning for fourth-year medical students

In Taiwanese medical schools, students enrolled in the Department of Medicine undergo two years of general and basic education. The third-year curriculum integrates content based on the general structure, functions, pathology, and behaviors of healthy individuals. Subsequently, the fourth-year curriculum focuses on studying abnormal and diseased individuals, as well as the acquisition of clinical knowledge and skills. The fifth- and sixth-year curricula involve clinical training in hospitals.

Approximately 45 M4 medical students each year at Medical College receive reasoning training through a systemic module that includes problem-based learning (PBL), clinical diagnosis medicine, and practical exercises. Clinical diagnostic medicine is integrated into the semesters for reasoning training. This observational and cross-sectional study of M4 medical student's reasoning ability was conducted from August 2019 to July 2023 at the Medical College in Taiwan. The purpose of the study was to identify deficits across the classes of students and create an intervention to teach diagnostic reasoning once global deficits were identified, without focusing on the precise assessment of any single student's diagnostic ability in one case. The primary endpoint was to identify the shortcomings of mid-term M4 students' clinical reasoning abilities using an effective and convenient assessment method. The subsequent teaching intervention aimed to share the results with problem-based learning (PBL) educators, clinical diagnosis medicine instructors, and practice teaching faculties. We developed a group history-taking activity combined with individual reasoning for the M4 mid-term group OSCE test. This test included an SP-simulated scenario with 12 min allotted for performance and 15 min for writing, totaling 27 min and focusing on clinical reasoning.

The setting of group history-taking with individual reasoning

The process of establishing the clinical reasoning station involved three steps. First, the clinical stations were established for group testing, encompassing three clinical scenarios: abdominal pain, fever, and weight loss, from 2019 to 2022. We conducted group interviews with four to five students to obtain more information than with one individual student, ensuring every student had a representative history. The purpose was to ensure active participation in gathering clinical information and selecting key details rather than having someone read the patient's full history or observe a master clinician. Second, the trainees rotated (four to five students per team) asking questions while gathering clinical information

from the SP. The third step was administering the reasoning test, followed by immediate feedback containing the correct answers (see Table 1). The clinical reasoning records the format in which information was elicited during interactions with SPs, including the chief complaint, 12 key informational facets (by teaching faculty consensus decision), three differential diagnoses, the most likely diagnosis, and the hypothesis generated (see Table 2). We decided to use 12 key informational elements for all three cases based on the national OSCE exam in Taiwan, which selects 10 to 15 key informational elements in the checklist. This is sufficient to form a differential diagnosis. Our teaching team, therefore, decided to use 12 key informational elements, asking students to write them down. If students recorded half of one key information element, we awarded a score of 1/2.

Participants and long station scenario contents

All fourth-year medical students from Medical College were enrolled in the mid-term OSCE group for clinical diagnosis, medicine, and practice. At the long station, the same teacher evaluated the trainees' performance based on the group OSCE checklist. Additionally, the SPs who simulated the symptoms and signs of the teaching scenario were trained. All methods were performed in accordance with the relevant guidelines and the "Declaration of Helsinki" to ensure ethical considerations and patient safety (IRB was approved by Memorial Hospital 23MMHIS146e).

Table 2. The answer format of M4 medical students' clinical reasoning test.

Group No. _____ Name:_____ Date: _____	
Chief complaint (12%):	
Problem (key information) representation (48%):	History-taking with key information gathering:
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
Differential diagnosis (15%) and Diagnostic justification (10%)	Hypothesis generation (15%)
1. The most likely diagnosis	
2. Tentative diagnosis	
3. Tentative diagnosis	

In 2019, the reasoning test topic was fever, with the chief complaint being fever for three days. The 12 key information representations included fever up to 39°C, chills, headache, and more. The most likely diagnosis was left acute pyelonephritis, with the other two tentative diagnoses being acute viral hepatitis and pneumonia. In 2020 and 2022, the reasoning test topic was abdominal pain, with the chief complaint of right upper quadrant (RUQ) pain for one day. The 12 key information elements included RUQ pain, midnight onset, and post-meal pain. The most likely diagnosis was acute pancreatitis, while the other two tentative diagnoses were acute cholangitis and cholecystitis. In 2021, the reasoning test focused on weight loss, with the chief complaint being a weight loss of 5 kg in one month. The 12 key information elements included weight loss of 5 kg, dry mouth, and mild shortness of breath. The most likely diagnosis was hyperthyroidism, with the other two tentative diagnoses being diabetes mellitus (DM) and malignancy. Concerning the weights of the score categories, three clinical teachers discussed and agreed on the allocation of points for the group history-taking with individual reasoning test. They agreed that the presentation of the chief complaint and the key information was the most important for M4 students, comprising 60% of the score. The most likely diagnosis accounted for 10%, the ability to list three relevant diagnoses correctly for 15%, and hypothesis generation for 15%. What were the students expected to include in the hypothesis generation? For example, they could write "fever, RUQ pain, and tea-colored urine" to support the diagnosis of acute cholangitis. The educational plans and materials for clinical reasoning in each scenario were reviewed by three clinical teachers from the teaching faculty (see Table 3).

Assessment and statistical analysis

In this study, we focused on four major components: (1) the trainee's total score in the long station by individual clinical reasoning; (2) the score of IG, which included points for the chief complaint and key information representation; (3) the scores of HG, DD, and DJ; and (4) identifying the shortcomings in the clinical reasoning abilities of M4. We have followed this guideline (The Do CTRINE Guidelines: Defined Criteria To Report Innovations in Education)[10] and used a recording paper at the end of the OSCE to collect the data, allowing us to determine the number of key information representations documented by each trainee. Additionally, we checked the number of diagnoses made by the trainees and whether they were correct. The data from group OSCE scores were shown as the mean ± standard deviation (SD). The correlations

Table 3. The topics, key information, and diagnoses of clinical reasoning in mid-term test.

School year	2019	2020 and 2022	2021
Topics:	Fever	Abdominal pain	Body weight loss
Chief complain	Fever for 3 days	RUQ pain for 1 day	Weight loss 5 kg in one month
12 Key information	Fever up to 39°C; chillness, headache, dry cough, diarrhea, left flank pain, dysuria, distention, hold the pee, ketamine used, safe sex, malaise.	RUQ pain, midnight onset, post-meal pain, pain score 7, lying down got worst and sitting subside, back pain, no skin lesions, no vomiting or diarrhea, tea color urine, no dysuria, mild fever, hyperlipidemia	Weight loss 5 kg, dry mouth, mild short of breath, feel hot, hand tremor, easy malaise, under stress, sister had hyperthyroidism, normal intake, activity no change, normal urination and defecation, no fever.
Differential diagnosis and Diagnostic justification	1. (Left) acute pyelonephritis: fever, chillness, dysuria, left flank pain and left CV angle knocking pain 2. Acute viral hepatitis: fever, malaise, distention, diarrhea, poor intake 3. Pneumonia: fever, chillness, dry cough	1. Acute pancreatitis: lying down got worst and sitting subside. hyperlipidemia; check amylase, lipase 2. Acute cholangitis: mild fever, RUQ pain, tea color urine, check GOT/GPT, bilirubin D/T, ALKP and WBC 3. Acute cholecystitis: RUQ pain, post-meal pain, midnight onset, back pain, do abdominal echo	1. Hyperthyroidism: weight loss 5 kg, feel hot, hand tremor, under stress, sister had hyperthyroidism; check TSH, free T4 2. DM: weight loss 5 kg, dry mouth; check AC sugar and HbA1c 3. Malignancy: weight loss 5 kg, easy malaise, Mild short of breath, do abdominal echo, CXR, check CBC and Lab data (ex: liver function, renal function)

observed between 2020 and 2022 for abdominal pain, whose scores were normally distributed, were analyzed using Student's *t*-test. Statistical analyses were performed using the SPSS software (version 23.0; SPSS, Chicago, IL, USA), which has a copyright license. All statistical analyses were based on two-sided hypothesis tests with a significance level of $P < .05$.

Results

In total, 177 fourth-year medical students participated in this study between 2019 and 2022. The average group OSCE score for the long station, which involved history-taking for information gathering, was 62% for the fever scenario, 87% for the weight loss scenario, and 90%–95% for the abdominal pain scenario (Supplementary data). Regarding trainees' long-term scores in group history-taking with individual reasoning, the students' score percentage for key information representation ability was 56% for the test topic of fever in 2019, 56% for the test topic of abdominal pain in 2020, and 58% for the topic of weight loss in 2021. There were no significant differences among these three different scenarios (see Table 4).

Regarding the ability to list the three diagnoses, the students' score percentage was poor on fever (27%), and better on abdominal pain and weight loss (67%–68%). Concerning the ability to list the most likely diagnosis, the students' score percentages were also lower in fever (34%) but higher in abdominal pain (82%) and weight loss scenarios (98%). Concerning the ability to generate hypotheses, the percentage scores explaining the clinical findings were generally poor at 26% for fever, 35% for abdominal pain, and 50% for weight loss. The students' overall reasoning total score percentage was the lowest for fever (49%), followed by abdominal pain (57%) and weight loss (63%) (see Table 4). Comparing the clinical reasoning scores of four-year medical students between 2020 and 2022 in the same abdominal pain scenario, we can see that the OSCE and DD scores improved significantly ($P < 0.01$, Table 5).

Several issues were observed regarding shortcomings in clinical reasoning ability among fourth-year medical students. (1) Students tend to forget to describe the duration of symptoms or mention too many irrelevant symptoms in the chief complaint, leading to incomplete and less focused information. (2) Students often ignored important details when writing the key information. For instance, they may describe fever without specifying the temperature, cough without describing whether it is dry or with sputum, or flank pain without a specific location (left or right). (3) Students frequently neglect important negative information, such as the absence of fever, gastrointestinal (GI) symptoms, genitourinary (GU) symptoms, and skin lesions.

Several challenges were identified in specific clinical scenarios regarding the shortcomings of differential diagnoses and diagnostic justifications in M4 students. In the case of fever for three days with flank pain, students had difficulty differentiating between various GU diseases, including urinary tract stones, urinary tract infections (UTI), (acute) pyelonephritis, (acute) nephritis, kidney trauma, pelvic inflammatory disease (PID), or flank sprain. In the case of RUQ pain for one day with worsening pain while lying down and subsiding while sitting, students faced difficulties in distinguishing between different GI diseases, including (acute or chronic) pancreatitis, (acute or chronic) cholangitis, and (acute or chronic) cholecystitis. In the case of weight loss of 5 kg in one month with hand tremors and dry mouth, students were uncertain about how to diagnose hyperthyroidism. Additionally, students struggled to establish clear correlations between weight loss and poor DM control (Table 6).

Table 4. The long station mid-term scores of the fourth-year medical students' individual clinical reasoning.

Items/school year (students no.)	2019 ($N=41$)	2020 ($N=45$)	2021 ($N=45$)
Score component/test topics	Fever	Abdominal pain	Body weight loss
Problem representation 60 points Chief complain 12 points Key information 48 points	11.1 ± 2.5	6.6 ± 4.9	7.6 ± 4.1
	26.7 ± 6.2	26.8 ± 5.4	27.8 ± 4.3
Hypothesis generation 15 points Explain the clinical findings	3.9 ± 2.5	5.2 ± 3.6	7.5 ± 2.9
Differential diagnosis 15 points List three diagnoses	4.1 ± 2.4	10.2 ± 4.2	10.0 ± 2.5
Diagnostic justification 10 points The most likely diagnosis	3.4 ± 4.8	8.2 ± 3.2	9.8 ± 1.5
Total score 100 points	49.0 ± 11.0	57.3 ± 10.7	62.7 ± 8.4

Table 5. The differences in reasoning scores between 2020 and 2022 in abdominal pain scenario.

Items/school year (students no.)	2020 ($N=45$)	2022 ($n=46$)	*P*
Group OSCE score	90.3 ± 5.9	95.1 ± 4.4	.000
Problem representation 60 points Chief complain 12 points Key information 48 points	6.6 ± 4.9	7.0 ± 3.2	.629
	26.8 ± 5.4	27.5 ± 4.4	.543
Hypothesis generation 15 points Explain the clinical findings	5.2 ± 3.6	6.3 ± 3.1	.115
Differential diagnosis 15 points List three diagnoses	10.2 ± 4.2	12.8 ± 3.7	.002
Diagnostic justification 10 points The most likely diagnosis	8.2 ± 3.2	7.5 ± 4.2	.358
Total score 100 points	57.3 ± 10.7	61.1 ± 11.2	.096

Table 6. The shortcomings of reasoning ability in long station by group history-taking with individual reasoning.

School year	2019	2020 and 2022	2021
Topics:	Fever	Abdominal pain	Body weight loss
Shortcomings of chief complaint	The time is not described. Describe too many other symptoms.	The time is not described. Describe too many other symptoms.	The time and reduced weight are not described. Describe too many other symptoms.
Shortcomings of problem representation	Lack of detail such as: fever (up to 39°C), (dry) cough, (safe) sex, (left) flank pain	Easy to ignore negative information such as: no skin herpes zoster, no diarrhea	Easy to ignore negative information such as: no fever, normal urination
Shortcomings of differential diagnosis and Diagnostic justification	1. The students cannot differentiate these GU diseases including urinary tract stone, UTI, (acute) pyelonephritis, (acute) nephritis, kidney trauma, PID, flank sprain. 2. The students cannot differentiate these GI diseases including IBS, acute appendicitis, AGE, acute cholecystitis, infectious diarrhea, intra-abdominal infection. 3. The students cannot differentiate lung diseases including URI, influenza, TB infection.	1. The students cannot differentiate GI diseases including (acute or chronic) pancreatitis, (acute or chronic) cholangitis and (acute or chronic) cholecystitis. 2. The students were unclear when to order the exams including plain abdomen, abdominal echo, abdominal CT, ERCP, MRCP, and PTCD. 3. The students were unclear how to order the correct liver function tests.	1. The students were unclear about how to diagnose hyperthyroidism, and the correlations between weight loss and DM with poor control. 2. The students knew weight loss is correlated with lung or thyroid CA, and wanted to prove by CT, MRI, or PET. 3. There were other diagnoses for SOB and palpitation such as arrhythmia, ACS, heart failure, TB, stress, depression, or anxiety

IBS: irritable bowel syndrome; ACS: acute coronary syndrome.

Discussion

According to previous literature, individual clinical reasoning training activities for M4 medical students using group history-taking with individual reasoning showed that the students' reasoning skills were insufficient, especially in achieving a high diagnosis rate in scenarios including abdominal pain, anemia, and fever. One potential reason for this may be the system-based teaching modules used by the fourth-year medical students.[9] In response to these findings, we created a cross-system problem-based learning scenario. During the assessment of the

different components of clinical reasoning, we found that the students' PR score percentage was below 60%, indicating that they struggled to effectively present and organize clinical information. The key information scores were similar across the three cases, whereas the other categories improved over time. This change may be due to student preparation or previous training in diagnostic reasoning. Why were the key information scores stable over time? This may be due to the students applying the basic interviewing skills they were taught in the M3 or M4 years. Additionally, students' ability to perform DD was poor in the fever scenario (score percentage of 27%) and slightly improved but was still suboptimal in the abdominal pain (score percentage of 68%) and weight loss scenarios (score percentage of 67%). It is difficult to judge the comparison across years using different cases. In the most difficult cases, fever had the lowest scores for differential diagnosis, diagnostic justification, and hypothesis generation. This may be because fever was the most difficult case for M4 to encounter, it was the first time this assessment was used in the curriculum, or there was a difference in trainee ability over time. Students preparing for such an assessment during the second year of curricular change often perform better because they are better prepared based on class information. To address these challenges, we recommend incorporating SP in the PBL class to facilitate students' information-gathering ability and practice in listing the three diagnoses. This approach aimed to improve students' skills in writing information and subsequently formulating a list of three potential diagnoses, thus enhancing their overall clinical reasoning abilities.

Clinical reasoning ability is fundamental to clinical practice in the workplace. It requires the teacher to accompany the learner to the bedside and perform activities such as shared decision-making and collaborative practice or entrustable professional activities to assess moment-in-time clinical reasoning.[11] PBL at school, which aims to integrate basic scientific knowledge into clinical presentations, has been widely used in medical education. However, previous studies have disclosed the impact of this teaching method, which predominantly employs a backward-directed hypothetical deductive model, which was accompanied by a tendency to generate diagnostic errors.[12] Conversely, Groves reported that clinical reasoning problems are a reliable method to assess clinical reasoning at different levels of individual students, both in theory and practice.[13] In our study, we disclosed the shortcomings of clinical reasoning among M4 medical students, which include missing the duration of symptoms in the chief complaint, neglecting detailed presentations, and tending to overlook negative important data during IG. Furthermore, the students' DD and DJ performance was suboptimal, particularly in differentiating between acute and chronic disease conditions. They also encountered challenges in accurately localizing the disease within the correct organ system. Overall, our results indicate that the M4 mid-term medical students were able to present an average of two correct diagnoses.

Clinical reasoning consists of numerous components, as described in Daniel et al's published practical guidance. The authors provided definitions for each of these clinical reasoning units, including IG, HG, problem presentation, DD, diagnosis, DJ, and management.[6] Identifying the different processes of clinical reasoning is crucial to assign proper training and remediation strategies, such as case information, presented teaching methods, or serial-cue approach teaching methods.[14] Regarding IG and HG, physicians first encounter patients. Gruppen et al emphasized that IG and integration are both important processes in diagnostic decision-making.[15] Four clinical reasoning problems have been identified in previous literature, including inadequate knowledge, faulty data gathering, faulty data processing, or faulty metacognition. Educators are encouraged to consider the point at which a learner's reasoning is breaking down and to employ various strategies to improve their clinical reasoning.[16] In our study, we found that students tended to recognize numerous symptoms beyond the chief complaint and often overlooked describing the duration and onset of primary symptoms presented by the patients. Additionally, they neglected to detail clinical clues such as the degree of fever, severity of pain, and other pertinent information. Furthermore, our participants tended to easily overlook patients' negative symptoms, which could have provided valuable diagnostic insights.

After acquiring adequate clinical information and identifying the problem, the process of presenting a list of hypotheses for confirmation is called a DD.[6] Our findings revealed that medical students have difficulty differentiating between diagnoses across several organ systems. The process of providing reasons and evidence to support a diagnosis was termed "diagnostic justification" (DJ). Williams RG et al reported that more than 20% of scheduled graduate students exhibited borderline or poor diagnostic justification performance in more than 50% of the cases.[17] Graber et al suggested that novel web-based decision support systems could enhance the ability to handle complicated cases.[18] Subsequently, we initiated a video-assisted teaching program in 2021 and found significant improvements in the OSCE and DD scores in the abdominal pain scenario from 2020 to 2022 years, respectively (Table 5).

Moreover, retaining the ability to make clinical decisions and arrange further treatment based on the diagnosis plays an important role in clinical reasoning. Stojan et al, who investigated the notion of diagnostic thresholds in medical students, highlighted that they revised their management of uncertainty.[19] Additionally, Goldszmidt et al highlighted the importance of identifying the factors influencing physicians' reasoning tasks, offering valuable teaching strategies to enhance clinical reasoning.[20] Although our study had a limited focus on medical students' diagnostic testing and treatment plans, we emphasized the importance of an immediate feedback system. This system facilitates knowledge absorption

and encourages students to conduct self-learning using clinical reasoning videos from schools.

Our study has several limitations. First, a sample size analysis was not performed. The relatively small sample size of 177 participating pre-clerkship students and the fact that the study was conducted at a single center and institute may limit its generalizability to other centers. Secondly, the testing topics varied and included fever, abdominal pain, and weight loss. Regarding the variation in complexity between different case settings, a standardized process is warranted for further evaluation. Additionally, our study focused on teaching clinical reasoning in practical environments by utilizing interactions with SP, which presented relatively simplified scripts that may not fully represent the complexity and diversity encountered in real-life patient scenarios. Overall, learning clinical reasoning is essential for physicians, and its importance persists throughout their careers. It is widely believed that the early training phase of clinical reasoning, which involves training medical students, illuminates the common presentation of diseases and reinforces the typical patterns in their memories.[11] Thus, further research is required to bridge this gap in clinical reasoning learning between medical students and novice physicians.

In conclusion, applying this group history-taking method to an individual reasoning assessment can identify students' shortcomings in the reasoning process. This approach enabled us to effectively enhance reasoning teaching strategies.

Abbreviations

IG, information gathering; HG, hypothesis generation; PR, problem representation; DD, differential diagnosis; LD, leading diagnosis; DJ, diagnostic justification; MT, management and treatment; OSCE, objective structured clinical examination; SP, standardized patient; PBL, problem-based learning; RUQ, right upper quadrant; DM, diabetes mellitus; GI, gastrointestinal; GU, genitourinary; UTI, urinary tract infection; PID, pelvic inflammatory disease.

Acknowledgments

Not applicable.

Author contributions

Lin substantially contributed to the conception and design of the study; Cheng and Lee: acquisition of data, analysis and interpretation of data; Lin, Cheng, and Lee drafted the article; Lin and Wu critically revised it for important intellectual content and final approval of the version to be published; all authors agree to be accountable for all aspects of the work.

Availability of data and materials

The datasets generated and/or analyzed during this study are not publicly available because this is a pilot study in M4 clinical diagnosis medicine practice for reasoning but are available from the corresponding author upon reasonable request.

Consent for publication

Not applicable.

Ethical considerations

This study was approved and informed consent was waived by the Memorial Hospital Institutional Review Board (23MMHIS146e), Taipei, Taiwan.

ORCID iD

Ching-Chung Lin https://orcid.org/0000-0001-8917-5293

Supplemental material

Supplemental material for this article is available online.

REFERENCES

1. Young M, Thomas A, Lubarsky S, et al. Drawing boundaries: the difficulty in defining clinical reasoning. *Acad Med*. 2018;93(7):990-995.
2. Norman G. Research in clinical reasoning: past history and current trends. *Med Educ*. 2005 Apr;39(4):418-427. doi:10.1111/j.1365-2929.2005.02127.x
3. Eva KW. What every teacher needs to know about clinical reasoning. *Med Educ*. 2004;39(1):98-106. doi:10.1111/j.1365-2929.2004.01972.x
4. Gruppen LD, Frohna AZ, et al. Clinical reasoning. In: Norman GR, van der Vleuten CPM, Newble DI, eds. *International handbook of research in medical education*. Vol 1. Kluwer Academic; 2002:205-230.
5. Gruppen LD. Clinical reasoning: defining it, teaching it, assessing it, studying it. *West J Emerg Med*. 2017;18(1):4-7.
6. Daniel M, Rencic J, Durning SJ, et al. Clinical reasoning assessment methods: a scoping review and practical guidance. *Acad Med*. 2019 Jun;94(6):902-912. doi:10.1097/ACM.0000000000002618
7. Im S, Kim DK, Kong HH, Roh HR, Oh YR, Seo JH. Assessing clinical reasoning abilities of medical students using clinical performance examination. *Korean J Med Educ*. 2016 Mar;28(1):35-47. doi:10.3946/kjme.2016.8
8. Charlin B, Boshuizen HP, Custers EJ, Feltovich PJ. Scripts and clinical reasoning. *Med Educ*. 2007 Dec;41(12):1178-1184. doi:10.1111/j.1365-2923.2007.02924.x
9. Lai JH, Cheng KH, Wu YJ, Lin CC. Assessing clinical reasoning ability in fourth-year medical students via an integrative group history-taking with an individual reasoning activity. *BMC Med Educ*. 2022 Jul 26;22(1):573. doi:10.1186/s12909-022-03649-4
10. Blanco M, Prunuske J, DiCorcia M, Learman LA, Mutcheson B, Huang GC. The DoCTRINE guidelines: defined criteria to report innovations in education. *Acad Med*. 2022;97(5):689-695.
11. Bowen JL. Educational strategies to promote clinical diagnostic reasoning. *N Engl J Med*. 2006 Nov 23;355(21):2217-2225. doi:10.1056/NEJMra054782
12. Patel VL, Groen GJ, Norman GR. Effects of conventional and problem based medical curricula on problem solving. *Acad Med*. 1991;66:380-389.
13. Groves M, Scott I, Alexander H. Assessing clinical reasoning: a method to monitor. Its development in a PBL curriculum. *Med Teach*. 2002 Sep;24(5):507-515. doi:10.1080/01421590220145743
14. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. *Med Educ*. 2015 Oct;49(10):961-973. doi:10.1111/medu.12775
15. Gruppen LD, Wolf FM, Billi JE. Information gathering and integration as sources of error in diagnostic decision making. *Med Decis Making*. 1991 Oct–29 Dec;11(4):233-239. doi:10.1177/0272989X9101100401
16. Cutrer WB, Sullivan WM, Fleming AE. Educational strategies for improving clinical reasoning. *Curr Probl Pediatr Adolesc Health Care*. 2013 Oct;43(9):248-257.
17. Williams RG, Klamen DL, Markwell SJ, Cianciolo AT, Colliver JA, Verhulst SJ. Variations in senior medical student diagnostic justification ability. *Acad Med*. 2014 May;89(5):790-798. doi:10.1097/ACM.0000000000000215
18. Graber ML, Tompkins D, Holland JJ. Resources medical students use to derive a differential diagnosis. *Med Teach*. 2009 Jun;31(6):522-527. doi:10.1080/01421590802167436
19. Stojan JN, Daniel M, Morgan HK, Whitman L, Gruppen LD. A randomized cohort study of diagnostic and therapeutic thresholds in medical student clinical reasoning. *Acad Med*. 2017 Nov;92:S43-S47. (11S Association of American Medical Colleges Learn Serve Lead: Proceedings of the 56th Annual Research in Medical Education Sessions). doi:10.1097/ACM.0000000000001909
20. Goldszmidt M, Minda JP, Bordage G. Developing a unified list of physicians' reasoning tasks during clinical encounters. *Acad Med*. 2013;88:390-397.