



OPEN ACCESS

EDITED BY

Goran Hauser,
University of Rijeka, Croatia

REVIEWED BY

Ziyi Yan,
Sichuan University, China
Maria Do Carmo Barros De Melo,
Universidade Federal de Minas Gerais, Brazil

*CORRESPONDENCE

Xingxing He
✉ hexingxing@whu.edu.cn
Mingxia Yu
✉ dewrosy520@whu.edu.cn
Zhibing Lu
✉ luzhibing222@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 16 March 2025

ACCEPTED 18 June 2025

PUBLISHED 09 July 2025

CITATION

Yang M, Liu Y, Dong Y, Ding Y, Lu Z, Yu M and He X (2025) The application of objective structured clinical examinations for evaluating clinical internship in long-term clinical medical students in China.
Front. Educ. 10:1594705.
doi: 10.3389/feduc.2025.1594705

COPYRIGHT

© 2025 Yang, Liu, Dong, Ding, Lu, Yu and He. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The application of objective structured clinical examinations for evaluating clinical internship in long-term clinical medical students in China

Miao Yang^{1†}, Yao Liu^{2†}, Yuhan Dong^{1†}, Yang Ding², Zhibing Lu^{3*}, Mingxia Yu^{4*} and Xingxing He^{2*}

¹Teaching Office, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China, ²Department of Gastroenterology, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China, ³Department of Cardiology, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China, ⁴Department of Laboratory Medicine, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China

Background: Clinical internship plays a vital role in enhancing medical students' theoretical knowledge, clinical operational abilities, and clinical thinking development. This study aims to establish an efficient evaluation system to assess interns' proficiency in clinical skills, and analyze the teaching effectiveness of clinical internships.

Methods: This is a retrospective descriptive study. A total of 75 long-term clinical medical students, including 26 eight-year students and 49 "5 + 3" students, who had already completed a 3-month clinical internship in internal medicine, participated in an Objective Structured Clinical Examination (OSCE). Four stations during the OSCE included history taking, abdominal physical examination, thoracentesis, and interpretation of auxiliary examination results. Both candidates and examiners voluntarily completed anonymous Likert scale questionnaire immediately after the OSCE. Results were analyzed using the Chi-square test and the *t*-test as appropriate.

Results: In history taking station, new examiners awarded significantly higher scores compared to examiners with experience in conducting examinations ($t=6.21$, $p<0.0001$). "5 + 3" candidates scored significantly higher than 8-year clinical medicine doctoral program candidates in physical examination and thoracentesis station ($t=5.316$, $p<0.0001$; and $t=2.145$, $p=0.0353$, respectively). According to the questionnaire survey, the majority of candidates and all examiners were quite satisfied with its design, organization, quality and effectiveness. More than half of the candidates and examiners believed that insufficient practice and inadequate preparation for revision were the factors most likely to affect performance. 75% of the examiners felt that candidates needed the most training in operative steps and content, but only 56 and 40% of the candidates agreed with this view. Furthermore, 56% of candidates and half of the examiners identified operational methods and humanistic care as areas that still required intensive improvement.

Conclusion: This study is an effective attempt to construct an OSCE evaluation system for the skills of long-term clinical medical students in the internship

stage. This system can objectively and comprehensively reflect the students' clinical skills, promote the realization of evaluation feedback and quality management of clinical internship teaching, which can be promoted in the evaluation of clinical internship.

KEYWORDS

medical education, internal medicine, OSCE, clinical internship, clinical skills

Introduction

Clinical skills, such as history taking and physical examination, are fundamental abilities that every medical student should acquire. The accurate and error-free execution of these clinical skills holds immense significance in the diagnosis, identification, and treatment of diseases (Reilly, 2003). The judicious utilization of clinical skills by doctors facilitates a comprehensive assessment and understanding of the patient's physical condition and disease status (Verghese and Horwitz, 2009). Moreover, the display of humanistic care and aseptic concept during clinical skill operations help to establish a harmonious doctor-patient relationship, reduce the occurrence of complications, and alleviate patient discomfort.

Clinical internship is a vital link between theory and practice during medical education. Through methods, such as observation, hands-on experience, and demonstrations, theoretical knowledge is effectively translated into clinical applications, and gaps in theoretical teaching are filled (Dev et al., 2020). Thus, clinical internship serves as a significant stage for medical students to acquire medical knowledge, clinical skills, clinical reasoning, medical ethics, and professional demeanor. This experience enhances the understanding and facilitates the transformation of medical students into competent doctors.

Competency-based medical education necessitates the implementation of appropriate learning and competency assessment tools (Lee and Chiu, 2022). In this context, the establishment of a scientific and rational quality monitoring system is critical for ensuring the quality of internships and medical education. Since the introduction of the concept in (Harden et al., 1975), the Objective Structured Clinical Examination (OSCE) has become a standardized method for assessing clinical competence and has been adopted in medical education worldwide. China began implementing the OSCE, using a multi-station assessment format, in the practitioner skills examination in 2015. This format not only replicates real clinical environments for diagnosis and treatment but also avoids the uncertainties in certain clinical scenarios (Ouldali et al., 2023). Through well-designed examination stations, examiners and/or standardized patients (SPs) evaluate candidates' abilities to gather clinical information, perform clinical procedures, and analyze clinical data. This approach effectively assesses whether students can seamlessly integrate the knowledge they have acquired at the undergraduate level with their clinical practice effectiveness (Sarfranz et al., 2021).

In China's clinical physician training system, the majority of them are based on the 5-year undergraduate program. Long-term

clinical physician training programs, including "5 + 3" integrated clinical medicine training program and 8-year clinical medicine doctoral program, are elite training programs that only a few medical students can access in China. The "5 + 3" integrated training program is a 5-year undergraduate program, plus 3 years of national standardized residency training, and at the same time, training for professional clinical medicine master's degree; the 8-year training program is a 5-year undergraduate study in clinical medicine, followed by 3 years of research training and clinical internship, and graduates are directly awarded a Doctor of Medicine degree, generally, after graduation, an additional 3 years of national standardized residency training are required.

In order to gain a comprehensive understanding of the interns' clinical skills and the quality of teaching during a 3-month clinical internship in internal medicine, an assessment was carried out through the OSCE. The aim of this assessment was to enhance students' learning motivation and guide clinical instructors in refining their teaching methods to ultimately improve the clinical skills gained during internship for long-term clinical medical students.

Materials and methods

Pre-examination preparation

We recruited examiners from our clinical teaching staff in the internal medicine department, including experienced and new examiners. Experienced examiners refer to those who hold the title of attending physician or above, have more than 3 years of continuous examination experience, and have participated in at least 1 large-scale national OSCE exam (such as medical examination, national proficiency test, residential training graduation exam, etc.). New examiners refer to those who have 0~2 years of experience as an examiner, but have not participated in large-scale national level exams. Examiners develop questions and establish marking criteria based on the standards of the Practitioner Proficiency Test (PPT) Skills Examination. The questions and criteria were then entered into the system by the Examination Centre. Prior to the examination, the examiners and standardized patients (SPs) received training to meet the requirements of (1) having a clear understanding of the test station settings and scoring standards, (2) refraining from providing any form of prompts or feedback to candidates during the examination, and (3) strictly following the established standards for scoring. The SPs were required to understand the scoring standards, encourage candidates to proactively inquire about their medical history, and provide consistent answers to each candidate. This study was

Abbreviations: OSCE, Objective Structured Clinical Examination; SP, standardized patients.

approved by the Medical Ethics Committee of Zhongnan Hospital of Wuhan University (approval number: 2023035K).

Subject of assessment

This study adopted a retrospective descriptive design. A total of 75 medical students, who had just completed both theoretical training and a 3-month clinical internship in internal medicine at Zhongnan Hospital of Wuhan University, participated in this study. Among them, there were 26 candidates from the 2017 eight-year program, with an average age of 23.04 ± 1.0 and a male-to-female ratio of 1:1. Forty-nine candidates were from the 2018 5 + 3 integration (“5 + 3”) program, with an average age of 22.24 ± 0.72 and a male-to-female ratio of 23:26. The difference in cultivation programs resulted in the 8-year program for the 2017 class starting the internship in the 6th year of enrollment, while the “5 + 3” program for the 2018 class started the internship in the 5th year. Therefore, there was a statistically significant difference in the average age of candidates between the two academic programs ($t = 3.989$, $P = 0.0002$).

Examination Station Setting

The examination took place at the Clinical Skills Centre of the Second Clinical College of Wuhan University, utilizing a multi-site OSCE model consisting of three parallel examination groups, each with four stations. The question bank is set according to the content of the National Medical Practitioner Qualification Examination, covering more than 20 symptomatic consultations such as fever, abdominal pain, cough, and so on, and standard operations for overall physical examination. In order to ensure the comparability of the scores of the single examination, this OSCE assessment used “Cough and sputum” history taking in Station 1, “Abdominal physical examination” physical examination in Station 2, “Thoracentesis” puncture operations in Station 3, “ECG (3 questions), X-ray (1 question) and CT (1 question)” interpretation of auxiliary examination results in Station 4 (Table 1, For detailed test questions and scoring tables, please refer to [Supplementary Material](#)). Each station had a time limit of 8 min, and a maximum score of 100 points. In response to the COVID-19, the first station was conducted using an online questioning system, while the third station involved volunteers acting as assistants.

TABLE 1 Examination station setting.

Station	Examination items	Examination contents
1	History taking (Standardized patient)	Cough and sputum consultation
2	Physical examination (Human patient simulator)	Abdominal physical examination
3	Four major puncture operations (Human patient simulator)	Thoracentesis
4	Interpretation of auxiliary examination results (Computer)	ECG, X-ray and CT

Examination process

Candidates sign in to the system and are randomly assigned to a station. The examiner at each station initiates the examination and starts a timer. During this time, the electronic marking scheme is used to evaluate the candidate's performance and assign marks. Once the examination at a station is completed, the candidate proceeds to a waiting station before being assigned to the next station by the system.

Post-exam feedback

The Examination Centre utilizes the Questionnaire Star system to generate and distribute anonymous feedback questionnaires. These questionnaires consist of 12 items across four categories, including the general evaluation of the OSCE, self-performance in the examination, support for the OSCE, and reflection on the examination results. Candidates and examiners voluntarily complete these questionnaires after the examination. The first three categories are scored on a 5-point Likert scale, ranging from 1 (strongly disagree) to five (strongly agree), where both “strongly agree” and “agree” are considered favorable. The last category includes two multiple-choice questions regarding factors that may impact performance and areas that need improvement, with five and six choices, respectively. A total of 29 valid feedback questionnaires (25 from candidates and 4 from examiners) were collected after the examination.

Statistical analysis

GraphPad Prism 8 was employed for data analysis. Data were presented as mean $\pm$ standard deviation (SD), and the independent sample *t*-test was utilized for group comparisons. The count data were expressed as rates, and the χ^2 test was employed for group comparisons. A statistically significant difference was considered if the *P*-value was less than 0.05.

Results

Overall performances

Although all examiners received uniform training prior to the examination, new examiners awarded significantly higher scores at the history taking station than experienced examiners (Table 2,

TABLE 2 Score discrepancy between two examiner groups.

Station	New examiners (<i>n</i> = 3)		Experienced examiners (<i>n</i> = 6)		Significance
	Mean	SD	Mean	SD	
1	90.02	4.16	77.13	10.30	$p < 0.0001$
2	78.48	6.25	77.78	5.33	$p = 0.6187$
3	78.67	5.35	80.48	9.11	$p = 0.3697$

$t = 6.21, p < 0.0001$). Candidates classified as “5 + 3” achieved significantly higher scores in physical examination ($t=5.316, p<0.0001$) and thoracentesis ($t=2.145, p=0.0353$) compared to 8-year candidates (Table 3). Figure 1 illustrates the distribution of candidates from different academic programs across each mark band at each station.

Loss of points by station

The assessment items that had high failure rates in the first station were primarily focused on gathering information about the history of the present illness (Figure 2A). The failure rates was similar between the “5 + 3” candidates and the 8-year candidates (Figure 2B). Additionally, 65.3% of the “5 + 3” candidates had an incomplete systematic review, while 57.7% of the 8-year candidates tended to overlook the patient’s concomitant symptoms, such as abdominal pain, diarrhea, and urinary tract irritation (Figures 2C,D).

The top 5 items that were missed in the second station were all related to abdominal palpation examinations (Figure 3A), with a similar pattern between the “5 + 3” candidates and the 8-year candidates (Figure 3B). Furthermore, the failure rates for performing the examination of vibratory sounds and hepatic jugular venous reflux sign were 80.8 and 69.2%, respectively, among the 8-year candidates (Figures 3C,D).

The items that resulted in a higher loss of marks at the third station are illustrated in Figure 4A. The loss of marks for “5 + 3” candidates in the preoperative measurement of vital signs with lung percussion [(77.6% vs 100%), $\chi^2 = 6.84, P = 0.0089$] and in the retesting of vital signs after puncture [(57.1% vs 80.8%), $\chi^2 = 4.187, P = 0.0407$] were significantly lower compared to the 8-year candidates (Figure 4B). Additionally, 69.4% of the “5 + 3” candidates demonstrated a poor understanding of the different puncture draw volumes, while 69.2% of the 8-year candidates needed to strengthen their concept of asepsis (Figures 4C,D).

Nearly half of the candidates answered questions 1 and 2 incorrectly at Station 4. Question 1 stated “Young male, sudden panic attack for half an hour,” was only answered correctly by 50.7% of candidates, with 28% and 18.7% choosing ventricular tachycardia and sinus tachycardia, respectively. Similarly, question 2 described “Young female with intermittent panic attacks for more than 1 month,” was only answered correctly by 45.3% of candidates, with 34.7% choosing second-degree type I AV block. In question 5, the interpretation of abdominal CT findings was correct in 66.7% of

cases, with 30.7% of candidates mistakenly choosing hepatocellular carcinoma, as shown in Figure 5 (the black bar represents the correct answer).

Questionnaire analysis

Analysis of the questionnaire feedback revealed that, apart from the difficulty of the examination, the majority of candidates and all examiners expressed satisfaction with the overall assessment of the OSCE, their own performance in the examination, and their support for the OSCE (Table 4).

An analysis of the results revealed that the two most influential factors on performance were insufficient general practice and inadequate preparation before the assessment (Figure 6A). 75% of the examiners believed that the candidates needed the most improvement in terms of operational steps and content, but only 56 and 40% of the candidates agreed, respectively. Furthermore, other areas that still require significant improvement were humanistic care (56% of the candidates) and operational methods (50% of the examiners) (Figure 6B).

Discussion

The setting of the OSCE examination station depends on the examination’s objective, with some emphasizing the application of knowledge while others focusing on the demonstration of clinical skills (Casey et al., 2009). In the past, this assessment method was used to evaluate the advantages of new teaching systems compared to traditional ones (Yao et al., 2020), as well as to assess whether interns have met discharge criteria (Zhang et al., 2010), among other purposes. In comparison to other institutions, the OSCE examination station in this work was designed to closely resemble a practitioner-level test, with an emphasis on assessing common clinical skills of long-term clinical medical students. The examination content aligns with the internship syllabus, allowing for the evaluation of students’ internship effectiveness and the identification of any gaps between their clinical skill level and the practitioner-level test standard.

The ability of examiners to organize, administer, adjudicate, provide feedback, and take responsibility for their practice in the OSCE, as well as their correct utilization of the marking criteria for assessment, significantly impacts the normality, fairness, reliability, and validity of the examination. Therefore, effective training of examiners is crucial for ensuring a high-quality OSCE. Lauren Chong’s study found that examiners’ seniority and experience influence the marking of communication skills but not operational skills (Chong et al., 2018), which aligns with the findings of this study. We observed that, despite receiving uniform training prior to the examination, examiners varied in their understanding of the purpose and rules of the examination. In the history-taking station, which primarily assessed candidates’ ability to communicate with patients, experienced examiners demonstrated more advanced questioning skills compared to new examiners. They also had higher expectations for candidates in areas such as humanistic care and the use of non-medical terminology, resulting in significantly lower scores awarded than new examiners. However, in the

TABLE 3 Score discrepancy in candidates from different academic systems.

Station	5 + 3 integration (n = 49)		Eight-year program (n = 26)		Significance
	Mean	SD	Mean	SD	
1	82.09	7.06	82.4	8.88	$p = 0.8684$
2	81.05	6.80	70.98	9.44	$p < 0.0001$
3	81.33	7.91	77.21	7.94	$p = 0.0353$
4	69.39	15.87	70.00	19.80	$p = 0.8846$

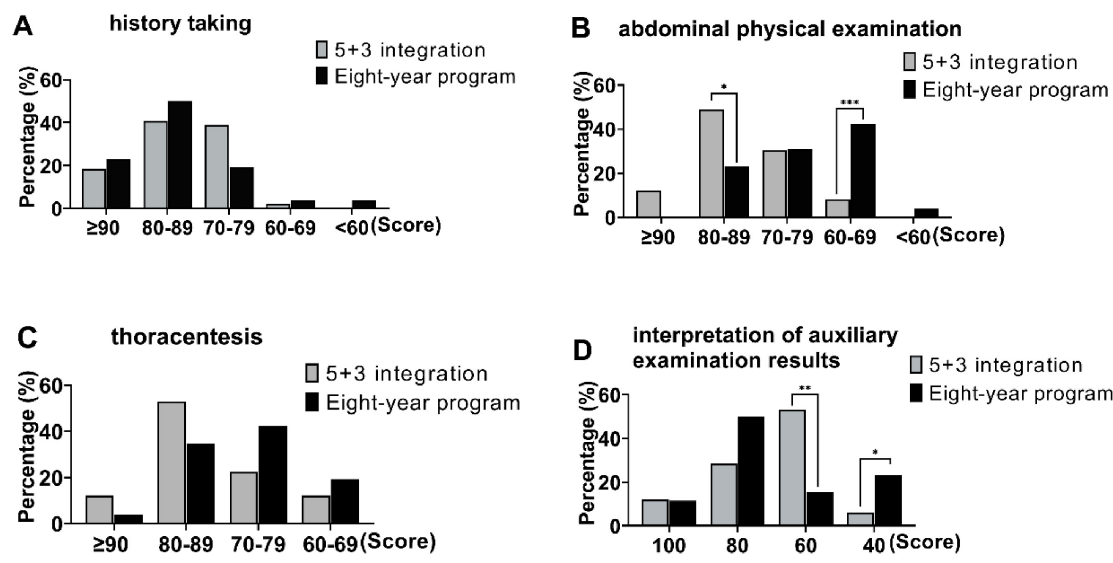


FIGURE 1 Proportion discrepancy in score segment among candidates from different academic systems (5 + 3 integration and 8-year program). (A) Station 1, history taking, (B) Station 2, abdominal physical examination, (C) Station 3, thoracentesis, (D) Station 4, interpretation of auxiliary examination results. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

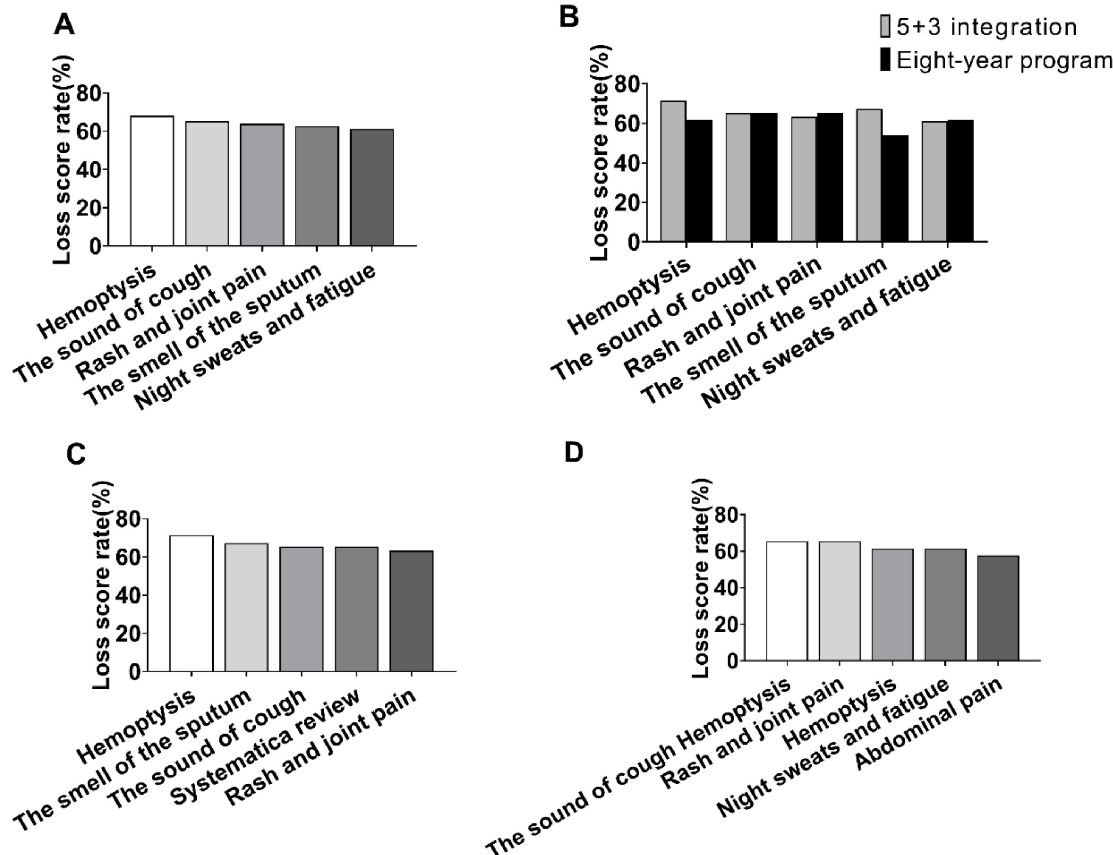


FIGURE 2 Loss score rate analysis in station 1. (A) Top five loss score rates in station 1. (B) Top five loss score rates discrepancy between candidates from 5 + 3 integration and 8-year program. (C) Top 5 loss score rates in 5 + 3 integration candidates. (D) Top five loss score rates in 8-year program candidates.

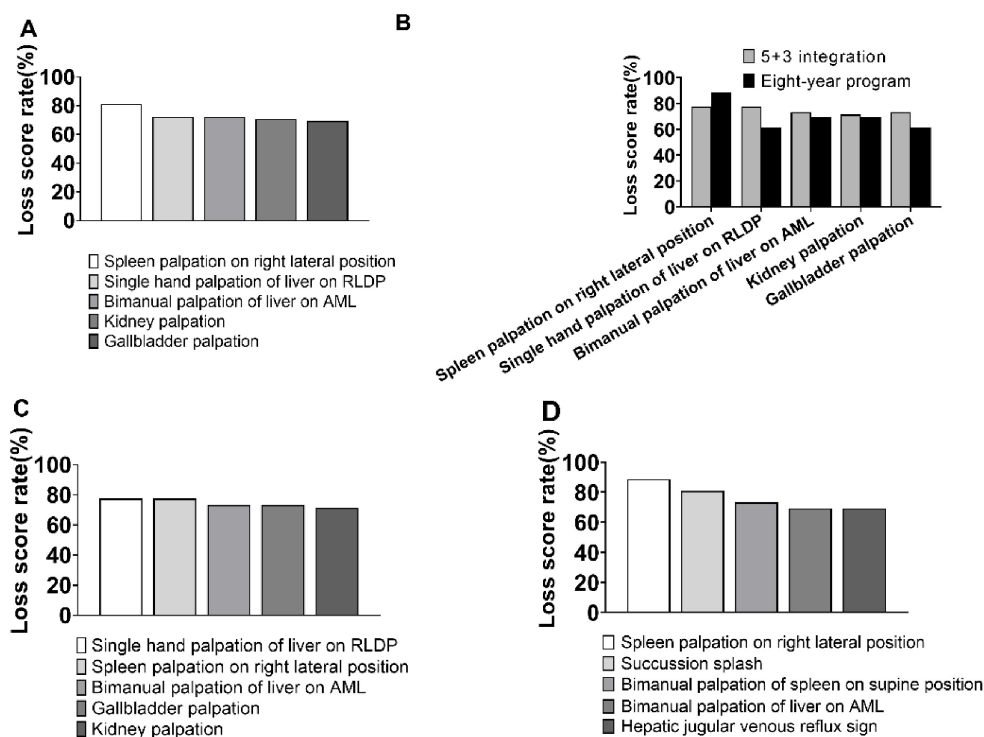


FIGURE 3

Loss score rate analysis in station 2. (A) Top five loss score rates in station 2. (B) Top 5 loss score rates discrepancy between candidates from 5 + 3 integration and 8-year program. (C) Top 5 loss score rates in 5 + 3 integration candidates. (D) Top five loss score rates in 8-year program candidates. AML, Anterior Median Line; RLDP, Right Lateral Decubitus Position.

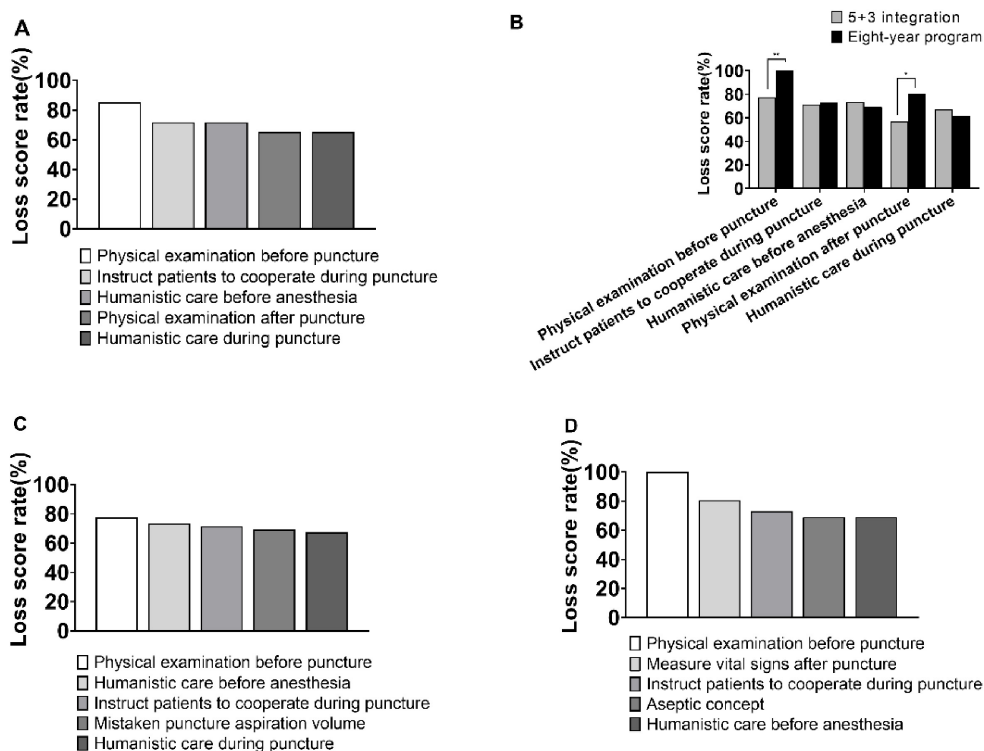


FIGURE 4

Loss score rate analysis in station 3. (A) Top five loss score rates in station 3. (B) Top 5 loss score rates discrepancy between candidates from 5 + 3 integration and 8-year program. (C) Top 5 loss score rates in 5 + 3 integration candidates. (D) Top five loss score rates in 8-year program candidates. * $p < 0.05$, ** $p < 0.01$.

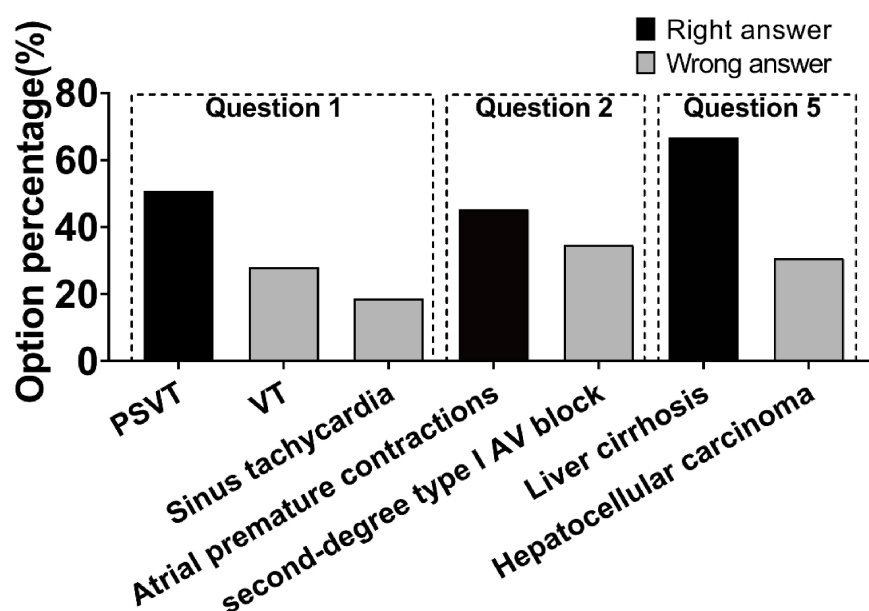


FIGURE 5

Loss score rate analysis in station 4, question1, 2 and 5. PSVT, Paroxysmal supraventricular tachycardia; VT, Ventricular tachycardia; AV, Atrioventricular.

TABLE 4 Analysis of the candidates' and examiners' responses about OSCE.

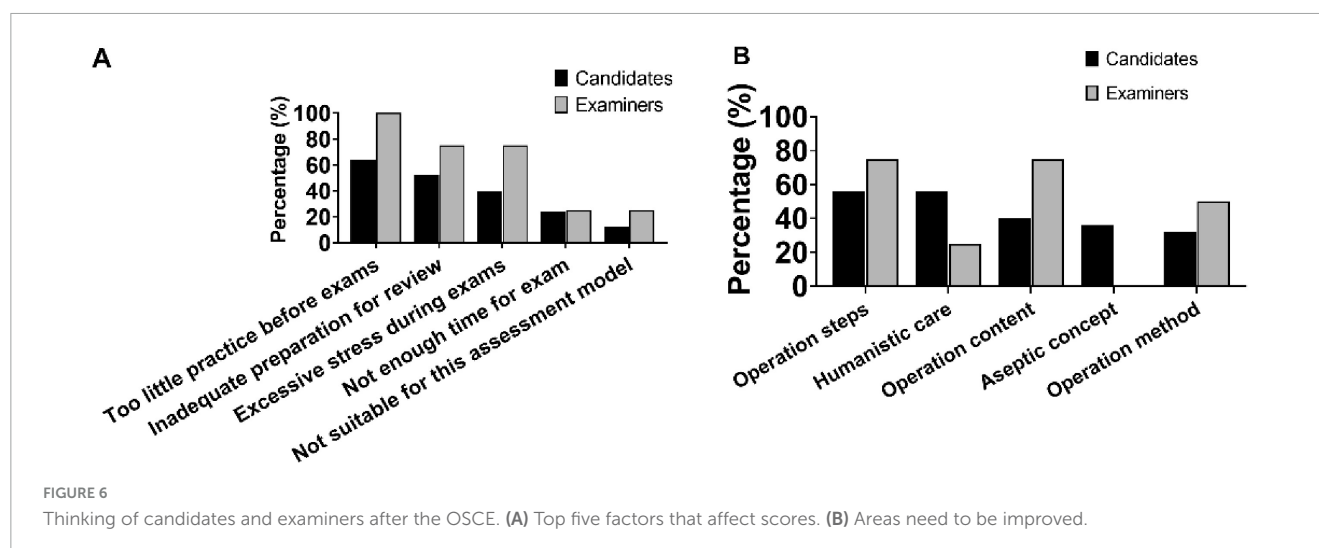
Question	Positive response (%)	Mean response
	Candidates/examiners	Candidates/examiners
1. Overview of OSCE		
Examination time setting is reasonable	76/100	3.96/4.75
Difficulty of examination is appropriate	48/75	3.48/3.75
SP's cooperation is satisfactory	76/100	4.04/5.00
2. Performance during OSCE		
Understand the content of this exam	84/100	4.20/4.25
Familiar with the scoring criteria	84/100	4.12/4.25
Nervous during the exam	44/0	3.28/0
Strictly follow the scoring criteria	0/100	0/4.00
3. Effectiveness of the OSCE		
Improve study motivation	60/100	3.72/4.00
Improve clinical skills	60/100	3.64/4.00
Continue to promote skill test	64/100	3.44/4.00

physical examination and thoracentesis station, which focused on candidates' operational skills, there were no significant differences in the scores awarded. This can be attributed to the presence of strict marking criteria and the objectivity of different examiners in assessing the correctness or incorrectness of the examination. These findings suggest that it is essential to establish a well-trained examiner team for conducting future OSCEs. This includes establishing a standardized selection mechanism for examiners and regularly assessing and updating the examiner team. Utilizing

online platforms, recording screens, catechism, and other targeted and regular training methods for different types of clinical skills tests can enhance the homogeneity of examiners and improve the fairness and impartiality of the examination (Chen and Shen, 2018). In future examinations, we will introduce quality control measures, either by establishing a calibration mechanism for examiners and conducting pre-marking visits, or by double marking and bringing in a third-party quality control team to increase the confidence of the marks.

According to the current academic degree training and education system in China, 3 years of residency training requires "5 + 3" students to be proficient in clinical skills. Upon passing the end-of-training examination, these students receive a standardized training certificate. On the other hand, during the Doctoral phase of 8-year clinical medicine program, these students do not undergo residency training, but instead focus on scientific research as a major part of their studies. They will undergo an additional 3 years of standardized clinical training at their workplace after graduation (Liu et al., 2023). The training programs at our university for these two different academic systems differ in that "5 + 3" students are required to possess strong clinical competence and meet the requirements for standardized residency training by the time of graduation, while the 8-year students are expected to have a strong aptitude for medical science research.

Results from this study show that "5 + 3" candidates scored significantly higher in abdominal physical examination and thoracentesis, which are two stations that reflect the level of clinical operation. Although the average scores of the candidates from the two academic tracks at the fourth station were similar, 61.5% of the 8-year candidates scored in the range of 80 points or higher, while 59.2% of the "5 + 3" candidates scored in the range of 60 points or lower. This discrepancy may be attributed to the fact that teachers expect "5 + 3" students to be more competent



in standardized residency training after 1 year and may provide more opportunities for hands-on experience during internships, which enhances the effectiveness of clinical internships. Despite having a stronger theoretical foundation, 8-year students face higher graduation requirements, greater research pressure, limited time and energy to dedicate to the internship course, and no further opportunities for clinical practice after the completion of the internship (Luo et al., 2017). Therefore, this study suggests that practical training needs to be further strengthened for 8-year students.

In terms of the areas where candidates are likely to lose points, there are several factors to consider. In the history-taking station, candidates often struggle due to a lack of clinical thinking ability and time constraints. As a result, they may neglect to ask for important details regarding the main symptom, such as the sound of coughing and the nature and odor of the sputum. Additionally, candidates may focus too heavily on respiratory diseases and overlook accompanying symptoms, such as skin rashes and osteoarthritis, which are commonly associated with rheumatological and immune system diseases. To address these issues, teachers should provide students with more opportunities to actively ask diagnostic questions during in-patient admission and outpatient consultation. Additionally, utilizing online simulation systems or encouraging students to act as standardized patients for each other can help facilitate practical training. Saunders's study found that the use of virtual patients improved students' ability to ask questions, which is consistent with our view (Raafat et al., 2024). In the abdominal physical examination station, beginners often struggle with spleen palpation compared to liver palpation. This can lead to missed diagnoses, irregularities, or errors. Furthermore, locating normal kidneys in the bilateral retroperitoneum can be challenging, and students may lack an intuitive feeling during clinical practice, resulting in a failure rate of 70.7%. Additionally, 69.3% of candidates confuse gallbladder palpation with Murphy's sign. To address these challenges, future teaching methods should incorporate problem-based learning (PBL) and case-based learning (CBL) approaches, as well as

multimedia teaching software and simulation mannequins. These tools can help simulate real pathological features and typical signs of the human body, providing students with audio and video demonstrations to compensate for the lack of clinical practice. At the thoracocentesis station, candidates often struggle with preoperative preparation and postoperative precautions. They may also omit the direct percussion method and confuse diagnostic and decompression thoracocentesis aspiration with abdominal puncture. To address these issues, it is recommended that students closely observe and learn from their teachers during internship. To emphasize the importance and necessity of doctor-patient communication, preoperative conversations between student and patient should be encouraged. During operations, teachers should adopt a bedside teaching mode that involves showing, leading, assisting, and guiding to cultivate students' observation, thinking, and problem-solving abilities. In the auxiliary examination results interpretation site, candidates tend to lose more points in ECG interpretation. This is often due to a lack of understanding of the mechanisms behind various types of tachycardia production and an inability to quickly identify typical ECG features commonly used in clinical practice. Additionally, candidates struggle to establish a connection between clinical symptoms and ECG presentations and rely too heavily on imaging reports. To address these challenges, students should be encouraged to actively learn from actual cases and develop the ability to summarize their findings during clinical practice.

We also found that 84% of candidates and all examiners had a clear understanding of the exam content and were familiar with the marking criteria. This indicates that the instructions provided prior to the exam were effective. The majority of candidates and examiners identified three factors that were likely to impact their performance: insufficient practice in general, inadequate preparation before the assessment, and excessive nervousness during the examination. These issues may stem from the lack of student participation in clinical settings during their placements. Similar to other medical schools in China, our hospital's lead

teachers were burdened with heavy medical workloads, leaving them with limited energy to dedicate to standardized bedside teaching. Consequently, students were not provided with sufficient opportunities for practical observation and hands-on experience in a structured manner (Gu et al., 2010; Song et al., 2019)). Furthermore, some patients were unwilling for interns to perform their clinical examinations (Aljoudi et al., 2016; Santen et al., 2005), resulting in even fewer chances for students to gain practical experience during their internships. Moving forward, it is advisable to enhance the requirements for internship instructors and regularly open the Clinical Skills Centre. This will provide students with access to simulators and virtual operating systems for practice purposes.

There was a discrepancy between the conclusions of the examiners and the candidates regarding the section that required the most improvement. This may be attributed to the low recovery rate and the absence of feedback data from the examiners at the thoracentesis station, which is not surprising since completing the questionnaire was voluntary. In the future, we will carefully design the questionnaire, incorporating open-ended questions to allow free expression of ideas and opinions by participants. This will enhance the motivation to complete the questionnaire and improve the accuracy of the feedback. Additionally, we will intensify our efforts in publicizing and distributing the questionnaires through various channels to increase the recovery rate. The results also shed light on the disparity between the feedback provided by examiners and the candidates' awareness of the issues revealed during the examination. Candidates often lack a comprehensive understanding of their clinical competence and are unable to fully evaluate themselves. This indicates that there are shortcomings in our current feedback approach to students. Candidates typically only focus on their total score and are unaware of the specific deductions made. Due to the absence of a comprehensive evaluation system in our regular skills teaching, insufficient teaching time, and limited observation of each student's performance by the teacher, there is a lack of professional feedback for students. Furthermore, there is a clear need for the development of a better digital question bank for clinical skills assessment in our hospital. Overall results from this study highlight the need for a practical and innovative plan to integrate teaching and assessment in medical education.

Main limitations of this study include: (1) the score discrepancy in history taking portion clearly existed between new and experienced examiners, (2) the question bank is currently limited in size and requires improvement in terms of quality, and (3) the questions in the questionnaire were not validated. These will be improved in future teaching research and practice.

Conclusion

In this work, we successfully established an OSCE system for evaluating the clinical skills of long-term clinical medical students in the internship stage. Additionally, the implementation of the OSCE system provides a reasonable assessment of the

students' clinical skill operation learning status and the quality of clinical internship teaching. Given the positive outcome, it is highly recommended to further promote and refine the OSCE system in future clinical internships to achieve more effective training of medical students with excellent clinical skills.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the author's, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of Zhongnan Hospital of Wuhan University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

MY: Writing – original draft, Conceptualization. YL: Writing – original draft, Data curation, Software, Visualization. YDo: Writing – original draft, Data curation, Software, Visualization. YDi: Writing – original draft, Methodology. ZL: Supervision, Writing – review and editing. MY: Supervision, Writing – review and editing. XH: Supervision, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The study was supported by Ministry of Education Industry-University Cooperation Collaborative Education Program (No. 231104794210013), Hubei Province Medical Youth Reserve Talents (Young Top Talents) (No. HB20200409), the Undergraduate Education Quality Construction Comprehensive Reform Project of Wuhan University (Nos. 2022ZG190, 202310, 202311, 2020048), Wuhan University Student Innovation Project (Nos. S202410486378, W202410486251), Wuhan University's "Promoting the Spirit of Educators" Demonstration Project (No. 1606-412600003).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2025.1594705/full#supplementary-material>

References

- Aljoudi, S., Alsolami, S., Farahat, F., Alsawyid, B., and Abuznadah, W. (2016). Patients' attitudes towards the participation of medical students in clinical examination and care in Western Saudi Arabia. *J. Family Community Med.* 23, 172–178. doi: 10.4103/2230-8229.189133
- Casey, P., Goepfert, A., Espey, E., Hammoud, M., Kaczmarczyk, J., Katz, N., et al. (2009). To the point: Reviews in medical education—The objective structured clinical examination. *Am. J. Obstet. Gynecol.* 200, 25–34. doi: 10.1016/j.ajog.2008.09.878
- Chen, J., and Shen, J. (2018). Exploration of training and management mode for examiners of clinical skills multi-level examination platform. *China High. Med. Educ.* 10, 38–91. doi: 10.3969/j.issn.1002-1701.2018.10.019
- Chong, L., Taylor, S., Haywood, M., Adelstein, B., and Shulruf, B. (2018). Examiner seniority and experience are associated with bias when scoring communication, but not examination, skills in objective structured clinical examinations in Australia. *J. Educ. Eval. Health Prof.* 15:17. doi: 10.3352/jeehp.2018.15.17
- Dev, M., Rusli, K., McKenna, L., Lau, S., and Liaw, S. (2020). Academic-practice collaboration in clinical education: A qualitative study of academic educator and clinical preceptor views. *Nurs. Health Sci.* 22, 1131–1138. doi: 10.1111/nhs.12782
- Gu, X., Sun, G., Wang, S., Li, L., Yang, Y., and Sun, Q. (2010). Investigation on the status of medical students, clinical practice. *Res. Med. Educ.* 4, 459–460. doi: 10.3760/cma.j.issn.2095-1485.2010.04.010
- Harden, R., Stevenson, M., Downie, W., and Wilson, G. (1975). Assessment of clinical competence using objective structured examination. *Br. Med. J.* 1, 447–451. doi: 10.1136/bmj.1.5955.447
- Lee, G., and Chiu, A. (2022). Assessment and feedback methods in competency-based medical education. *Ann. Allergy Asthma Immunol.* 128, 256–262. doi: 10.1016/j.anai.2021.12.010
- Liu, X., Feng, J., Liu, C., Chu, R., Lv, M., Zhong, N., et al. (2023). Medical education systems in China: Development, status, and evaluation. *Acad. Med.* 98, 43–49. doi: 10.1097/ACM.0000000000004919
- Luo, J., Wang, H., Guo, Y., Huang, W., and Zhu, C. (2017). Study of the impact of differences in academic structure on the learning of undergraduate students in clinical medicine. *China High. Med. Educ.* 1, 28–29. doi: 10.3969/j.issn.1002-1701.2017.01.014
- Ouldali, N., Le Roux, E., Faye, A., Leblanc, C., Angoulvant, F., Korb, D., et al. (2023). Early formative objective structured clinical examinations for students in the pre-clinical years of medical education: A non-randomized controlled prospective pilot study. *PLoS One* 18:e0294022. doi: 10.1371/journal.pone.0294022
- Raafat, N., Harbourne, A., Radia, K., Woodman, M., Swales, C., and Saunders, K. (2024). Virtual patients improve history-taking competence and confidence in medical students. *Med. Teach.* 46, 682–688. doi: 10.1080/0142159X.2023.2273782
- Reilly, B. (2003). Physical examination in the care of medical inpatients: An observational study. *Lancet* 362, 1100–1105. doi: 10.1016/S0140-6736(03)14464-9
- Santen, S., Hemphill, R., Spanier, C., and Fletcher, N. (2005). 'Sorry, it's my first time!' Will patients consent to medical students learning procedures? *Med. Educ.* 39, 365–369. doi: 10.1111/j.1365-2929.2005.02113.x
- Sarfraz, F., Sarfraz, F., Imran, J., Zia-Ul-Miraj, M., Ahmad, R., and Saleem, J. (2021). OSCE: An effective tool of assessment for medical students. *Pak. J. Med. Health Sci.* 15, 2235–2239. doi: 10.53350/pjmhs211582235
- Song, D., Guo, J., Zhang, W., Wang, Y., Zhang, Q., and Zhai, J. (2019). Investigation and analysis of clinical practice status and satisfaction among medical students. *Chin J Med Edu.* 39, 197–202. doi: 10.3760/cma.j.issn.1673-677X.2019.03.010
- Verghese, A., and Horwitz, R. (2009). In praise of the physical examination. *BMJ* 339:b5448. doi: 10.1136/bmj.b5448
- Yao, L., Li, H., Zeng, M., Yuan, J., Qiu, F., and Cao, J. (2020). Application of OSCE based on organ system in clinical practice teaching. *Chin. J. Med. Edu. Res.* 4, 458–462. doi: 10.3760/cma.j.cn116021-20190303-00105
- Zhang, Y., Zhou, J., Zhang, D., Liu, X., and Li, W. (2010). The application of OSCE in the internal medicine phase exam for the overseas students. *Res. Med. Educ.* 7, 975–977. doi: 10.3760/cma.j.issn.2095-1485.2010.07.041