Assessing clinical reasoning abilities of medical students using clinical performance examination

Sunju Im[1], Do-Kyong Kim[2], Hyun-Hee Kong[3], Hye-Rin Roh[4], Young-Rim Oh[5] and Ji-Hyun Seo[6]

[1]Medical Education Unit, Pusan National University School of Medicine, [2]Department of Medical Humanities, Dong-A University College of Medicine, [3]Department of Parasitology, Dong-A University College of Medicine, [4]Department of Medical Education, Inje University College of Medicine, [5]Department of Obstetrics & Gynecology, Kosin University College of Medicine, Busan, and [6]Department of Pediatrics, Gyeongsang National University School of Medicine, Jinju, Korea

Purpose: The purpose of this study is to investigate the reliability and validity of new clinical performance examination (CPX) for assessing clinical reasoning skills and evaluating clinical reasoning ability of the students.

Methods: Third-year medical school students (n=313) in Busan-Gyeongnam consortium in 2014 were included in the study. One of 12 stations was developed to assess clinical reasoning abilities. The scenario and checklists of the station were revised by six experts. Chief complaint of the case was rhinorrhea, accompanied by fever, headache, and vomiting. Checklists focused on identifying of the main problem and systematic approach to the problem. Students interviewed the patient and recorded subjective and objective findings, assessments, plans (SOAP) note for 15 minutes. Two professors assessed students simultaneously. We performed statistical analysis on their scores and survey.

Results: The Cronbach α of subject station was 0.878 and Cohen κ coefficient between graders was 0.785. Students agreed on CPX as an adequate tool to evaluate students' performance, but some graders argued that the CPX failed to secure its validity due to their lack of understanding the case. One hundred eight students (34.5%) identified essential problem early and only 58 (18.5%) performed systematic history taking and physical examination. One hundred seventy-three of them (55.3%) communicated correct diagnosis with the patient. Most of them had trouble in writing SOAP notes.

Conclusion: To gain reliability and validity, interrater agreement should be secured. Students' clinical reasoning skills were not enough. Students need to be trained on problem identification, reasoning skills and accurate record-keeping.

Key Words: Educational measurement, Reproducibility of results, Clinical competence

Introduction

In order to solve the problem of a patient, it is not only crucial to memorize the concept of the disease but also to reason it. Reasoning is a process of derive a conclusion from the evidence provided by the patients.

Clinical reasoning is a cognitive process, constantly occurring in clinical environment, in which the physician assess, diagnose, and treat the patients based on the information provided by them [1].

The very first step in clinical reasoning or problem solving is problem identification, through which, the physician precisely acknowledges the problem [2]. The

patients may visit the clinic from 1 to 6 concerns [3,4], and their first complaint may not be the most significant one [5]. The patients generally speak of their problems within 60 seconds [6]. Thus, the physician should encourage the patients to unload all the problems the patient is suffering and pay full attention to the details to discover the most compelling problem [3]. Identifying the main problem must precede in order to perform effective and efficient diagnosis.

Once the main problem has been established, it is important to obtain relevant information systematically. An expert should not exclude potential diseases of the problem one by one, but should execute forward reasoning in categorized fashion [7]. For example, if the patient has sudden rise in creatine, the physician should not rule out whether it is post-streptococcal glomerulonephritis or rhabdomyolysis, but rather assess if the underlying cause is prerenal, renal, or postrenal to be efficient [7]. Successful clinical reasoning requires the user to organize and store medical knowledge in long term memory and implement it whenever necessary. An expert distinguishes oneself from a novice in organizing and extracting necessary information [7,8].

Evaluating the student's clinical reasoning capacity, which is identifying the main problem, systematically gathering information, and finding out appropriate diseases or categories, could suggest implication for our future educational system. Up to date, the research primarily evaluated the clinical reasoning ability of the students through written exam such as multiple choice questions, script concordance test [2]. Clinical performance examination (CPX) or objective structured clinical examination (OSCE) has been rarely utilized to evaluate clinical reasoning.

In Korea, CPX and OSCE first introduced in 2009 has displayed positive effects, yet at the same time some negative effects on education field [9,10]. Contrary to our expectation for the students to get a fair chance to see the patient, the students wasted their time by lingering in clinical skills center or spent their time memorizing references. Instead of critically thinking and analyze the chief complaint, the student prepares a list of questions, irrelevant to the patient's symptoms. It is important to identify the problem via conversation with the patient and yet in current CPX, patient provides single clinical symptom, which makes it difficult to assess reasoning process.

CPX allows examinee to write postencounter notes as an adjuvant tool to assess clinical reasoning. In Korea, CPX adopted interstation, where students keep a note after the patient encounter. But there is limitation to assess clinical reasoning skills using current postencounter note [9,10]. The note is comprised of patient assessment and diagnosis, treatment, education plan. This record omits history taking & physical examination and merely states conclusion, which makes it hard grade the whole process behind building the diagnosis from the interview. Although medical records take many different shapes, subjective and objective findings, assessments, plans (SOAP) note attempts to clearly deliver history & physical examination obtained from the patient and impression following it [11].

From this study, CPX case was developed and applied to inspect students' clinical reasoning ability. The primary purpose of this research is to determine how well the case measures student's ability. This can be connected with reliability and validity of the test.

The secondary purpose refers to estimate students' clinical reasoning capacity through developed case. This study aims at describing what percentage of students precisely identified the patient problem according to principles of problem solving process, that of students systematically approaching to chief complaint, that of students accurately recording postencounter note. The

data could suggest what to improve and direction in our education system.

The purpose of this study could be summarized to the following: (1) Is CPX developed to assess clinical reasoning capacity reliable and valid? (2) How well did the students, evaluated by CPX, perform in clinical reasoning?

Subjects and methods

1. Subjects

This research was cross-sectional observational study. Subjects were 313 third-year students of five universities in Busan-Gyeongnam consortium for clinical performance examination. They have participated in joint clinical test from December 1st to 3rd of 2014. The curriculum of five universities consists of either European model of 2+4, total 6-year curriculum or American education system of 4+4, in the last four of which, the students take 2 years of classroom study then another 2 years in clinical clerkship. The third-year students has finished essential clerkship in both curriculums.

2. Development of case

The cases were invented by six researchers, belonging to Research Committee of Busan-Gyeongnam consortium for clinical performance examination. The researchers exhibit highest level of expertise in their respective field and has distinguished experience in task such as these.

The patient of the case was a 22-year-old female with rhinorrhea. Rhinorrhea is common symptom to primary physicians and is usually associated with other clinical symptoms and thus, was viewed appropriate to evaluated the students. The first line of the patient is "I have a headache" and when the students encourage the patients to explain further, the simulated patients adds "I worry whether it is meningitis because I vomited once." And finally on the third line, the simulated patient talks about rhinorrhea. Via 15 minutes of interview with the patient, the examiner must (1) identify chief complaint, (2) enquire about patient's history and perform physical examination related to chief complaint, (3) discuss future diagnostic and therapeutic plans, (4) fill out SOAP note (Appendixes 1, 2).

The assessment of student's performance was divided into five categories: agenda setting, history taking, patient education, clinical reasoning and with a total of 30 items, including global rating and patient-physician interaction (PPI) (Appendix 3). Clinical reasoning process focused on identifying the chief complaint early and whether history taking and physical examination were logical and systematic. The criteria of the assessment were determined after five conferences and discussions.

As for standardized patients (SPs), four candidates were selected who have met the prerequisites: 20 hours of basic training and minimum of 3 years of service to Busan-Gyeongnam clinical performance examination. They have standardized and adjusted the case after 10 hours of case rehearsals. Via mock exam, the SP has had eight training sessions to evaluate PPI checklists.

Two professors simultaneously graded the students in the checklists except PPI, after 1 hour of preceding education and once conformity between their scoring has been checked.

SOAP note, drafted by the students, were checked by two researchers and judged uncertain recording by the agreement of the two.

3. Examination schedule

The exam was performed at four separate clinical skills center within universities for 3 days. The environment,

training of simulated patients, broadcast system, overall management, online grading system were standardized. Six OSCE cases and six CPX cases were prepared in daily exam. Time limit for original CPX was 10 minutes while research CPX was limited to 15 minutes.

4. Survey

Questionnaires for professors and students were developed to extrapolate content validity and face validity. The questionnaire enquired problem familiarity, level of difficulty, authenticity, validity of the case with four questionnaire in 5-Likert scale.

5. Methods of analysis

The reliability of the test was measured with Cronbach α value and Cohen κ coefficient was used to find degree of agreement between the examiners. Content validity and face validity utilized 5-point scoring system of survey of assessors and students and clinical reasoning capacity was based on response rate of checklists and SOAP note. The data were analyzed with SPSS version 21.0 for Windows (IBM Corp., Armonk, USA).

Results

1. Average score for subject case

The average score for rhinorrhea case was 48.14 and was significantly lower than average score of 58.01 of six CPX cases. The mark increased from 42.22 (first day), 46.80 (second day), 55.44 (third day). The highest score was 87.69, and the lowest was 14.31.

2. Reliability analysis

The cases were composed of 30 checklists, of which Cronbach α value was 0.878. The average value for other CPX cases were 0.785. Cronbach α of 12 cases in first, second, third day was 0.743, 0.672, and 0.652, respectively. Once rhinorrhea case was eliminated, overall reliability fell to 0.728, 0.634, and 0.618, respectively. The correlation coefficient of six CPX and rhinorrhea cases was 0.677. Cohen's κ coefficient between two graders was 0.785.

3. Validity analysis

Students responded with problem familiarity by 3.09, level of difficulty by 3.93, and authenticity by 3.71 and validity of the case by 3.62. On the other hand, professors had negative view on validity of the cases by 2.79. Seven out of 14 professors responded negatively, three of them found it improper that rhinorrhea was not mentioned at the very first line, while other three commented it is also important to rule out meningitis (Table 1).

4. Clinical reasoning ability in relation to response rate

Most of the students listened to the end of the first

Table 1. Students' and Assessors' Surveys

No.	Item	Students (n = 313)	Assessors (n = 14)
1	The case was familiar to perform (problem familiarity)	3.09	2.50
2	The case was difficult than traditional CPX cases (level of difficulty)	3.93	3.50
3	The case was authentic than traditional CPX cases (authenticity)	3.71	3.36
4	The case was appropriate to assess student's performance (validity of the case)	3.62	2.79

Data are presented as averages of 5-point Likert scale.
CPX: Clinical performance examination.

line of the patient (94.2%). Students realized patient's concern about meningitis in early stage of the conversation (97.4%, 305 students), while 232 students (74.1%) checked neck stiffness, and 248 students (79.2%) explained unlikeliness of meningitis to the patient (Table 2).

Two hundred seventy-nine students (89.1%) knew patient's rhinorrhea during the interview. One hundred eight students (34.5%) clarified rhinorrhea at the early stage, and 58 students (18.5%) performed sound history taking and physical examination to find out cause of rhinorrhea. One hundred seventy-eight students (55.3%) explained to the patient that common cold was the cause

of rhinorrhea (Table 2).

5. Clinical reasoning capacity based on SOAP note

Based on SOAP note, 108 out of 313 students (34.5%) recorded concern about meningitis on subjective (S) section. Two hundred thirty-six students (75.4%) wrote down rhinorrhea and 61 students among them (19.5%) classified rhinorrhea as main symptom. In objective (O) section, 157 students (50.2%), 60 students (19.2%), 43 students (13.8%) performed neck stiffness test, nasal speculum examination, and throat examination. One

Table 2. Students' Performance according to Checklists (n=313)

	Yes	No
Finish first line	295 (94.2)	18 (5.8)
Concern about meningitis		
Know it in the early phase of interview	305 (97.4)	8 (2.6)
Perform neck stiffness	232 (74.1)	81 (25.9)
Explain low possibility	248 (79.2)	65 (20.8)
Rhinorrhea		
Know it	279 (89.1)	34 (10.9)
Identify it in the early phase of interview	108 (34.5)[a]	205 (65.5)[b]
Perform systematic interview	58 (18.5)[a]	255 (81.5)
Explain possibility of common cold	173 (55.3)	140 (44.7)

Data are presented as number (%).
[a]They meant "doing properly", [b]It was included "doing properly" (117, 37.4%) and "not doing" (88, 28.1%).

Table 3. Comparison of Performance between Checklists and SOAP Notes (n=313)

	Performance	Recording	Recording[a]
Agenda setting & history taking			
Concern about meningitis	305 (97.4)	108 (34.5)	105 (34.4)
Know rhinorrhea	279 (89.1)	236 (75.4)	221 (79.2)
Identify rhinorrhea as chief complaint	108 (34.5)	61 (19.5)	35 (32.4)
Physical examination[b]			
Neck stiffness	232 (74.1)	157 (50.2)	151 (65.1)
Nasal speculum examination	101 (29.3)	60 (19.1)	59 (58.4)
Throat examination	263 (84.0)	43 (13.7)	41 (15.6)
Patient education			
Possibility of common cold	173 (55.3)	158 (50.5)	122 (70.5)

Data are presented as number (%).
SOAP: Subjective and objective findings, assessments, plans.
[a]Students actually kept a record of what they performed, [b]They were included "doing properly" and "doing improperly."

Table 4. Students' Clinical Reasoning Abilities according to Checklists and Recordings

	Recording: wrote common cold as first impression		Total
	Yes	No	
Checklist: "rhinorrhea is due to common cold"			
Yes	122	51	173
No	36	104	140
Total	158	155	313

hundred fifty-eight students (50.5%) wrote common cold as first impression (Table 3).

We also checked whether students actually kept a record of what they performed. Of those who realized the patient was worried about meningitis 34.4% took note on patient's concern in S section. Among the 108 students who acknowledged rhinorrhea as the chief complaint, 35 (32.4%) write it down as the chief complaint. Among the students who performed physical examination, 50.2%, 19.2%, and 13.8% wrote neck stiffness test, nasal speculum test, and throat examination down in O section, respectively. One hundred fifty-eight students (50.5%) wrote common cold as first impression. Fifty-one students explained common cold as impression did not record it. Thirty-six students who did not explain common cold wrote down common cold as first impression (Table 4).

Discussion

The primary objective in this research was investigate reliability and validity of CPX cases developed to evaluate clinical reasoning capacity. Cronbach α across the items within the CPX was 0.878, and overall reliability was diminished to when rhinorrhea case was excluded. Cohen κ coefficient between two graders was 0.785.

Our result on reliability is higher than the one reported by Brannick et al. (0.878 vs. 0.78) [12] but our results on intergrader agreement is lower (0.785 vs. 0.89). In the same research, the number of graders and communication checklists increased reliability [12]. In this research, reliability was reinforced by two graders and training SPs to assess accurately students' communication ability (PPI). Relatively lower agreement between graders could be attributed to lack of orientation about the case, because it was performed an hour earlier for security purposes.

Pell et al. [13] reported that change in overall reliability should be observed when a certain case was removed. It needed to modify the case when removal of the case increases the reliability. When Cronbach value is low and elimination of a certain case increases the reliability, the case might be measuring different content, or be error in case design, or the students might be taught in different fashion, or the evaluators might not assess to designed criteria. This case, although distinguished its content and format from other cases, contributes significantly to overall reliability. This indicates that the case was measuring essentially similar contents with other cases, and that no major error was present, the students learned about the contents, and the evaluators were judging based on the criteria. The score of this case and the total score of CPX has high correlation coefficient of 0.677, which can be translated to low case specificity and conformity in assessment objectives despite the new format.

From the survey, students thought that the rhinorrhea case, when compared to traditional CPX cases, was

unfamiliar (3.09) more difficult (3.93) more authentic (3.71) and appropriate to assess clinical performance level (3.62). On the other hand, professors found it inappropriate to evaluate clinical performance level (2.79). One of the reasons behind unfavorable response was rhinorrhea symptom was not presented at the very first line. The script was designed to presents rhinorrhea on the third line so as to assess whether students can pay attention to patient's history and figure out chief complaint. The argument that this case presentation is unsuitable because the first line does not manifest the main problem seems to overlook this arrangement. It appears to be necessary to provide sufficient time to educate the graders and spell out the objective of this exam.

Another controversy rose from not having enough meningitis checklists. The patient's most concern was meningitis and meningitis could be serious condition than rhinorrhea, it is necessary for students to put emphasis on ruling out meningitis. It is true that check-lists to make a differential diagnosis on meningitis was relatively short. It seems necessary to make symptoms or diseases not differ so much in terms of severity.

The secondary purpose of this study is to evaluate clinical reasoning capacity. In this case, the capacity is determined by whether student can identify rhinorrhea as a chief complaint in meningitis-fearing patient, make a systemic approach to rhinorrhea, disprove meningitis and record their performance. Consequently, it came to our understanding that the students cannot pinpoint patient's chief complaint (34.5%), is unable to make systemic approach (18.5%) and has difficulty filling out SOAP note.

Thirty-four students (10.9%) were unable to find out the patient had rhinorrhea until the end of the interview. These students assumed the first complaint patient commented to be chief complaint and failed to encour-

age patient to unravel all of patient's symptoms. Previous CPX students feel comfortable at, allowed student to identify chief complaint without further inquiries. This probably has contributed to failure of some students.

Even after students succeeded in listening to all of patient's problem, a number of students failed to identify the chief complaint and ended up history taking in all of the symptoms. In other words, 117 students (37.4%) threw questions about fever, headache, vomiting and every possible symptoms until they came across rhinor-rhea then identified it as their chief complaint. Only 34.5% of the students identified rhinorrhea as chief complaint in early interview and concentrate it. The predicament was evident on the survey. Students found rhinorrhea case harder than previous cases (3.93) since the students were likely to get lost when there are more than one symptoms.

This phenomenon was also reported in other docu-mentation where students focused only on first com-plaint, or deviated to other symptoms than the chief complaint [14,15].

Identifying rhinorrhea as chief complaint is one thing but systematically approach to rhinorrhea was another (58 students, 18.5%). Systematical approach to the disease is not about ruling out underlying causes one by one. It is distinguishing the cause into two categories: allergic rhinitis to infectious rhinitis, then if infection is suspicious, student can narrow it down to possible underlying cause. This type of approach involves clinical reasoning process based on schema, structured thinking algorithm. It has been reported that students are better at storing medical knowledge and making diagnosis with it, if the knowledge is learnt with schema [16,17]. Instead, students examine according to their question lists based on impressions associated with each symptom [18].

Although only 29.3% of the students performed nasal

speculum test, neck stiffness test and throat examination was performed 74.1% and 84.0%, respectively. Forty-one students of 263 students who performed throat examination wrote it down in SOAP note. This data implies that student performed this test not because they tried to screen pharyngeal injection or post nasal drip but as a routine examination and this might be an explanation to such a high performance rate. Yudkowsky et al. [19] claims that head to toe physical examination irrelevant to the context of the case, not to point out the cause of the disease deteriorates the efficiency of physical examination, give negative impact to learning, especially in terms of reliability and validity. Wilkerson & Lee [20] states that physical examination in CPX has no correlation to physical examination in OSCE and reports the score of CPX examination similar to that of OSCE was lower than that of OSCE. This results indicate students does not lack in their ability to perform simply but ignores necessary physical examination or cannot squeeze it in time limit. Therefore, physical examination must incorporate clinical reasoning ability, and students must be taught the reason and purpose behind serving certain physical examination [19,20].

A considerable number of students (15.6%–79.2%) who performed physical examination recorded it, illustrating that students had difficulty in filling SOAP notes. Most of the students was aware of patient's concern about meningitis, yet only one-third recorded it. Students remained negligent on patient's feeling and expectation by showing tendency to not discuss worries of the patients and overlook it [21]. Although students could recognized rhinorrhea as chief complaint, fewer students wrote it down (35 out of 108 students). Some students did not concentrate on rhinorrhea during the interview but realized it after the interview, then wrote it down.

The students who made their final diagnosis as common cold were two-thirds (209 out of 313 students).

Although 122 students both mentioned and recorded it as common cold, 51 students mentioned yet not recorded it and 36 students did not mention it but recorded it. Disagreement between performance and recording was shown. It is imperative to emphasize on recording what has been explained to the patient and vice versa. Students had hard time when to fill out postencounter note. During or after the interview, most of them was not comfortable filling note out in front of the patient. Recently, the need for postencounter note is being stressed, and the education on it does not seem to be sufficient. It seems necessary to understand intellectual level of the student and to provide feedback by students filling it out. Students should be disciplined on how to keep a precise record on significant information, perform clinical reasoning based on it and share the information with the patient.

From this study, we can safely conclude the following. First of all, students need to be disciplined on identifying patient's problem accurately and promptly. Both doctor's impression and patient's concern must be considered to extrapolate chief complaint. Secondly, systematical approach toward patient's main symptom is necessary and schema could come in handy. Thirdly, physical examination should not be performed in head to toe fashion but must be linked with clinical reasoning process. Fourthly, students must appreciate the importance of postencounter note and should devote oneself to fill it out precisely and thoroughly based on the interview.

This research attempted to assess clinical reasoning capacity of the students using CPX and can be concluded that CPX is possible methods, once reliability and validity could be reinforced. On top of it, we tried to propose a new direction in our education system. We hope to refine this program to evaluate clinical reasoning skills of the students most accurately and have positive

influence in educational system.

References

1. Hrynchak P, Takahashi SG, Nayer M. Key-feature questions for assessment of clinical reasoning: a literature review. Med Educ 2014; 48: 870-883.

2. Groves M, O'Rourke P, Alexander H. Clinical reasoning: the relative contribution of identification, interpretation and hypothesis errors to misdiagnosis. Med Teach 2003; 25: 621-625.

3. Mauksch LB, Dugdale DC, Dodson S, Epstein R. Relationship, communication, and efficiency in the medical encounter: creating a clinical model from a literature review. Arch Intern Med 2008; 168: 1387-1395.

4. Beasley JW, Hankey TH, Erickson R, Stange KC, Mundt M, Elliott M, Wiesen P, Bobula J. How many problems do family physicians manage at each encounter? A WReN study. Ann Fam Med 2004; 2: 405-410.

5. Baker LH, O'connell D, Platt FW. "What else?" Setting the agenda for the clinical interview. Ann Intern Med 2005; 143: 766-770.

6. Beckman HB, Frankel RM. The effect of physician behavior on the collection of data. Ann Intern Med 1984; 101: 692-696.

7. Mandin H, Jones A, Woloschuk W, Harasym P. Helping students learn to think like experts when solving clinical problems. Acad Med 1997; 72: 173-179.

8. Harasym PH, Tsai TC, Hemmati P. Current trends in developing medical students' critical thinking abilities.

Kaohsiung J Med Sci 2008; 24: 341-355.

9. Park HK. The impact of introducing the Korean Medical Licensing Examination clinical skills assessment on medical education. J Korean Med Assoc 2012; 55: 116-123.

10. Kim JH. The effects and challenges of clinical skills assessment in the Korean Medical License Examination. Korean Med Educ Rev 2013; 15: 136-143.

11. Coulehan JL, Block MR. The medical interview: mastering skills for clinical practice. 5th ed. Philadelphia, USA: F.A. Davis; 2006. p 45-139.

12. Brannick MT, Erol-Korkmaz HT, Prewett M. A systematic review of the reliability of objective structured clinical examination scores. Med Educ 2011; 45: 1181-1189.

13. Pell G, Fuller R, Homer M, Roberts T; International Association for Medical Education. How to measure the quality of the OSCE: a review of metrics – AMEE guide no. 49. Med Teach 2010; 32: 802-811.

14. Roh H, Park KH, Jeon YJ, Park SG, Lee J. Medical students' agenda-setting abilities during medical interviews. Korean J Med Educ 2015; 27: 77-86.

15. Mavis BE, Wagner DP, Henry RC, Carravallah L, Gold J, Maurer J, Mohmand A, Osuch J, Roskos S, Saxe A, Sousa A, Prins VW. Documenting clinical performance problems among medical students: feedback for learner remediation and curriculum enhancement. Med Educ Online 2013; 18: 20598

16. Woloschuk W, Harasym P, Mandin H, Jones A. Use of scheme-based problem solving: an evaluation of the implementation and utilization of schemes in a clinical presentation curriculum. Med Educ 2000; 34: 437-442.

17. Blissett S, Cavalcanti RB, Sibbald M. Should we teach using schemas? Evidence from a randomised trial. Med Educ 2012; 46: 815-822.

18. Hauer KE, Teherani A, Kerr KM, O'Sullivan PS, Irby DM. Student performance problems in medical school clinical skills assessments. Acad Med 2007; 82(10

Suppl):S69-S72.

19. Yudkowsky R, Otaki J, Lowenstein T, Riddle J, Nishigori H, Bordage G. A hypothesis-driven physical examination learning and assessment procedure for medical students: initial validity evidence. Med Educ 2009; 43: 729-740.

20. Wilkerson L, Lee M. Assessing physical examination skills of senior medical students: knowing how versus knowing when. Acad Med 2003; 78: S30-S32.

21. Maguire P, Fairbairn S, Fletcher C. Consultation skills of young doctors: II--Most young doctors are bad at giving information. Br Med J (Clin Res Ed) 1986; 292: 1576-1578.

Appendix 1. Door Instruction

A 22-year-old female Young-Mi Kim (patient number: 14678)

<Vital sign>

BP: 110/80 mm Hg HR: 80/min
RR: 16/min BT: 36.8°C

Applicant has 15 minutes to
- identify chief complaint
- perform history taking and physical examination related to chief complaint
- discuss diagnosis and treatment with the patient
- complete SOAP note during the interview

Appendix 2. Doctor's Questions and Patient's Answers in the Early Phase of Interview: Agenda Setting

Doctor	Patient
What brought you here?	(first line) I had fever two days ago, It came with a headache.. Hmm... (hesitates and slowly speaks) I took some tyrenols and felt better. Headache was gone and fever came down.
Alright, please continue? (And then···)	(second line) But here's the thing. I vomited once yesterday. I was terrified if it was meningitis. One of my friend was admitted with it recently, and what I am going through look exactly like it.
Is there anything else that is bothering you?	(third line) I have runny nose and stuffy nose.
Oh I see. Anything else you would like to add?	No, that is it.
Please tell me more about your runny nose.	It came along with the headache and fever. All other symptoms faded yet this still persists.

Appendix 3. Checklists

Agenda setting		Yes	No
1	Finished first line		
2	Showed concern about meningitis early in the interview		
3	Has runny nose		

History taking		Yes	No
Chief complaint			
4	No head truama		
Allergic/non-allergic			
5	Rhinorrhea independent from location or season		
6	Sneezing nor itching		
7	No concurrent allergic disease (rhinitis/atopic dermatitis/asthma)		
8	No family member suffering allergic disease (rhinitis/atopic dermatitis/asthma)		
Infection/non-infection			
9	Color change from transparent to white		
10	Recently suffered sore throat, common cold		
11	Runny nose from both sides		

Physical examination		Doing properly	Doing improperly	Not doing
Allergic/non-allergic				
12	Checked under the eyes and nasal ridge			
Infection/non-infection				
13	Nasal speculum test			
	Doing properly: ① Insert nasal speculum closed ② Used penlight to observe (card dispensed)			
14	Bilateral compression of cheek bone			
	Doing properly: ① Compressed cheek bones below the eyes			
15	Throat examination			
	Doing properly: ① Open mouth ② Compress tongue ③ Say 'Ah' ④ Shine light			
Meningitis				
16	Neck stiffness			
	Doing properly: ① Supine position ② One hand under patient's head ③ Flexes neck			
17	Hemiplegia			
	Doing properly: ① Checked ankle, knee, elbow, wrist ② Flex and extend with opposing force ③ Compare both sides			
18	Proficiency of physical examination			
	Doing properly: ① Displayed expertise in diagnosis and physical examination			

Appendix 3. Checklists (Cont.)

Patient education		Yes	No
19	Low risk of meningitis was explained		
20	(fever, headache, vomitting) rhinorrhea (is) are due to common cold		

Clinical reasoning process		Doing properly	Doing improperly	Not doing
21	Identified rhinorrhea as chief complaint early stage			
	Doing properly: ① Found out about rhinorrhea early ② Recognized rhinorrhea as the chief problem Doing improperly: ① Enquired about fever, headache, vomiting then asked about rhinorrhea Not doing: ① Could not find out about rhinorrhea ③ Found out about rhinorrhea but asked no questions			
22	History taking and physical examination were systematic			
	Doing properly: ① Rhinorrhea → Non-allergic → Infection			

Overall performance (professor)		Very good	Good	Average	Poor	Very poor
23	Satisfaction at performance level					

Patient-physician interaction		Very good	Good	Average	Poor	Very poor
1	Formed decent bond and started the interview - introduction, interest, respect, confidence & reliability					
2	Listened closely - open question, waiting, responding, listening attitude					
3	Asked questions efficiently - easy to understand, clarify, summarize, set up agenda - avoided: misleading, multiple questions					
4	Showed empathy to my concerns and emotions - attitude, eye contact, empathy, understanding (patient concerns and thoughts)					
5	Explained easily to understand - knowledge and explanations, concise, easy vocabulary, questions - avoided: impatient, unnecessary information					
6	Physical examination - hand washing, discrete, prior explanation (process/purpose), ask permission					

Overall performance (SP)		Very good	Good	Average	Poor	Very poor
7	I would like this physician to take care of me in future					