



Development and psychometrics of script concordance test (SCT) in midwifery



Somayeh Delavari^{1,2}, Mitra Amini³, Zohreh Sohrabi^{2,1*}, Hamidreza Koohestani⁴, Sajad Delavari⁵, Rita Rezaee³, Elaheh Mohammadi⁶, Anne Demeester⁷, Bernard Charlin⁸, Masoomeh Kheirkhah^{9,2}

Received: 20 Nov 2017

Published: 23 Aug 2018

Abstract

Background: Clinical reasoning plays an important role in the accurate diagnosis and treatment of diseases. Script Concordance test (SCT) is one of the tools that assess clinical reasoning skill. This study was conducted to determine the reliability and concurrent and predictive validity of SCT in assessing final lessons and gynecology exams of undergraduate midwifery students.

Methods: At first, 20 clinical scenarios followed by 3 questions were designed by 2 experienced midwives. Then, after examining the content validity, 15 scenarios were selected. The test was used for 55 midwifery students. The correlation of SCT results with grade point average (GPA) was measured. To evaluate the concurrent validity of SCT, the correlation between SCT scores and the final exam of the gynecology course was measured. To measure predictive validity, the correlation of SCT scores with comprehensive exams of midwifery was calculated. Data were analyzed using SPSS software. Descriptive statistics, Pearson correlation, and coefficient Cronbach's alpha were used for analysis. The test's item difficulty level (IDL) and item discriminative index (IDI) were determined using Whitney and Sabers' method.

Results: The internal reliability of the test (calculated using Cronbach's alpha coefficient) was 0.74. All questions were positively correlated with the total score. The highest correlation coefficient was related to GPA and comprehensive test with the score of 0.91. The correlation coefficient between SCT and the final test (concurrent validity) was 0.654, and the correlation coefficient between SCT and comprehensive test (predictive validity) was 0.721. The range of item discriminative index and item difficulty level in this exam was 0.39-0.59 and 0.32-0.66, respectively.

Conclusion: SCT shows a relatively high internal validity and can predict the success rate of students in the comprehensive exams of midwifery. Also, it showed a high concurrent validity in the final test of gynecology course. This test could be a good alternative for formative and summative tests of clinical courses.

Keywords: Clinical decision-making, Problem-solving, Assessment, Educational, Clinical competence, Clinical skill, Midwifery, Psychometrics

Copyright© Iran University of Medical Sciences

Cite this article as: Delavari S, Amini M, Sohrabi Z, Koohestani H, Kheirkhah M, Delavari S, Rezaee R, Mohammadi E, Demeester A, Charlin B. Development and psychometrics of script concordance test (SCT) in midwifery. *Med J Islam Repub Iran.* 2018 (23 Aug);32:75. <https://doi.org/10.14196/mjiri.32.75>.

Introduction

To train future paramedical professionals, there is a need for a wide range of medical knowledge and clinical skills (1). On the other hand, medical and paramedical professions are challenging and complex because of the

fore, to provide effective clinical care to patients with multiple and complex problems, the use of clinical reasoning is inevitable (4), which is a critical enabler for physicians (5, 6). A physician or paramedic applies different kinds of

Corresponding author: Dr Zohreh Sohrabi, zo_sohrabi@yahoo.com

¹ Department of Medical Education, School of Medicine, Iran University of Medical Sciences, Tehran, Iran.

² Center for Educational Research in Medical Sciences, School of Medicine, Iran University of Medical Sciences, Tehran, Iran.

³ Clinical Education Research Center, Shiraz University of Medical Sciences, Shiraz, Iran.

⁴ Department of Medical Education, Saveh University of Medical Sciences, Saveh, Iran

⁵ Health Human Resources Research Center, School of Management & Information Sciences, Shiraz University of Medical Sciences, Shiraz, Iran.

⁶ Department of Medical Education, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran.

⁷ Ecole Supérieure du Professorat et de l'éducation (ESPE), Aix-Marseille Université, EA4671 ADEF, 13 284 Marseille cedex07.

⁸ Centre de pédagogie appliquée aux sciences de la santé (CPASS), Université de Montréal, Montréal, QC, Canada.

⁹ Department of Midwifery and Reproductive Health, Nursing Care Research Center School of Nursing and Midwifery, Iran University of Medical Sciences, Tehran, Iran.

↑What is "already known" in this topic:

SCT is a suitable tool to evaluate the depth and breadth of students' knowledge because of its focus on the structure and organization of the knowledge.

→What this article adds:

SCT had a high correlation with GPA, final exam score, and the comprehensive exam of midwifery; thus, this test could be a good alternative for traditional exams of clinical courses.

knowledge and critically evaluates the evidence and acquisitions to reach a diagnosis (3). For this reason, professors of different fields of medicine and care specialists believe that clinical reasoning is a core competency for doctors or paramedics (3,6,7).

Since clinical reasoning is an important skill for all doctors, regardless of their specialty and primary purpose of medical education (8), in the past decades, clinical reasoning has been considered as one of the important aspects of medical competence (3,6). Students must earn it in universities and professors should ensure the proficiency of the learners (2,3,6,9). Nowadays, various methods are designed to assess clinical reasoning, most of which are based on a clinical scenario (1, 10, 11); and one of these methods is SCT (1,12).

Scripts are actually knowledge networks used to make clinical decisions and solve clinical problems (1, 13, 14). These knowledge structures are acquired during training and clinical experience (15), which is exclusively adapted with doctors' tasks (15). According to the description, scripts are made of the relationship among contextual factors, disease mechanisms, and symptoms (14,15) that are narrative (7), classified (8) and mental advanced organizers (1, 2).

SCT examines the degree of concordance between judgments of the expert panels and responses of the students (5, 16). SCT is used to assess clinical reasoning in different aspects of clinical skills when there is no certainty (1,17). This test requires a degree of ambiguity, for example, when the disease history is unknown or decision-making is done on little information (13,17). This tool allows assessment based on real situations that are not measured by other methods even in questionable and ambiguous situations (13). This method of assessment is a suitable method to evaluate the depth and breadth of students' knowledge because of its focus on the structure and organization of the knowledge (12). Moreover, in this test the agreement and harmony of the intellectual process between students and professors is examined, as the intellectual process of the faculty is different from that of the students (1,12).

SCT is used in various fields of medical education (13, 18-23), nevertheless, to our knowledge, its application in the field of midwifery training has not still been investigated or published in English or Persian languages. However, an article was published in French language about application of SCT in midwifery (24,25). Because most obstetrics and gynecology services are provided by graduates of midwifery, their ability of clinical reasoning and decision-making can play a major role in reducing the incidence of adverse events in the treatment of obstetrics or gynecology related diseases. Thus, it seems that evaluating the clinical reasoning capability, identifying the strengths and weaknesses, and eliminating the weaknesses of midwifery students are of paramount importance. However, currently, students' formative and summative assessments are often a combination of multiple choice and descriptive questions that do not have much potency in evaluating the clinical reasoning skill of midwifery students. It seems like using clinical reasoning tests, including script concordance test,

could have a major role in enhancing knowledge and experiences of the students. This study was conducted to evaluate the SCT reliability and concurrent and predictive validity to test final lessons and gynecology exams of undergraduate midwifery students.

Methods

SCT design

We designed SCT according to the guidelines of Dory et al. (11) and Fournier et al. (15). The first step in evaluation using SCT is to determine the aim of evaluation (12, 15, 26), and the next step is to determine the target group (12, 15). When designing the test, the domain of the knowledge that is supposed to be evaluated should be determined (12). To develop SCT questions, one can write the clinical scenario, but a group of 2 can cause innovations and creativity in developing the questions (15). In this study, the target group was midwifery students in the sixth semester. The test, aiming at determining the concurrent and predictive validity and SCT reliability, was provided using the test specifications table by 2 midwifery faculty based on the subjects taught. Three principles (challenging clinical situations, answering questions based on the Likert scale, and scoring based on the compliance options chosen by the student) were considered in preparing the questions.

Initially, 20 basic scenarios (each followed by 3 questions) were prepared by 2 faculty members of the Midwifery Department. To ensure face and content validity, questions were reviewed independently by 5 other experts. A total of 5 unnecessary and irrelevant scenarios were eliminated and 15 scenarios and 45 questions remained. Therefore, the SCT test included 15 scenarios related to gynecology; each clinical scenario was followed by 3 questions in 3 areas of diagnosis, clinical course, and treatment, which were independent from one another. Questions had 3 columns: (1) "if you think (hypothesis)", (2) "If you recognize the findings (new information including laboratory investigation, clinical examination, etc.)", and (3) the 5-point Likert scale between -2 to +2; the students were required to evaluate the degree of relation of findings to scenarios and hypotheses (12). The score of 2+ is very much in favor of the hypothesis, +1 is much in favor of the hypothesis, the score of 0 represents no communication, -1 is in favor of rejecting the hypothesis, and -2 is strongly in favor of rejecting the hypothesis (12,13,15). **Table 1** demonstrates a typical SCT scenario that is followed by the 3 questions.

A 30-year-old woman, who had a normal delivery 2 years before, was admitted with secondary amenorrhea for 3 months. Her menstrual cycle was normal 3 months after delivery.

SCT was distributed among 15 midwifery faculty members who had 5 to 27 years of work experience in midwifery education. Scoring the questions was based on the panelists' consensus and comments (12, 17). The options that were selected by most of the panelists got the score of 1 and those that were not selected by them got the score of 0. In terms of number of responses, other options that were selected by panelists received the scores between 0 and 1 (12). Panelists' score for every question was regarded as the ref-

Table 1. SCT sample scenario

If you think.....		In the history of examination or paraclinical results you encounter the findings	These findings prove or disprove the diagnostic hypothesis
q1	Hypothyroidism	TSH higher than normal	-2 -1 0 +1 +2
q2	Galactorrhea	Prolactin higher than normal	-2 -1 0 +1 +2
q3	Menopause	FSH higher than normal	-2 -1 0 +1 +2

Table 2. The method of scoring in SCT (12, 27)

Choices	+2	+1	0	-1	-2
Panelists who chose this answer	4	9	2	0	0
The number of panelists who chose this choice divided by the number of responses that were selected by most of the experts	4.9	9.9	2.9	0.9	0.9
The score for this question	.44	1	.22	0	0

erence score, and students received a score of 0 or 1 according to the reference score. For example, in Table 2 panelists did not select choices -1 and -2. Therefore, if a student selected one of these choices, then, the score of 0 was considered. The choice of 0 was selected by 2 panelists; therefore, a student who chose this choice got the score of 0.22. The choice of +1 was selected by 9 experts, and the score of students who selected this choice was 1. The choice of +2 was selected by 4 experts, and the score of students who chose this choice was 0.44. The total score was obtained from the sum of the scores.

Participants and test time

In Iran, students should study for 4 years to attain a bachelor degree in midwifery. The duration of this study was from May 2015 to February 2017. The test was held for the target group of 55 midwifery students in January 2016. To evaluate the concurrent validity, SCT was held a day after the final exam of the gynecology course at the end of the sixth semester, and the scores correlation was examined with the scores of written gynecology exam. With respect to predictive validity, test scores were used to predict students' success in the comprehensive test, which is a standard test required for graduation in midwifery.

Participation in this test was optional. Furthermore, to maintain the anonymity of the students and ensure that participation did not impact the results of the mandatory final tests of the students, codes were given to students. Furthermore, faculty members were assured that the results of the study and test scores remained confidential. Before the test, students were provided with an adequate explanation as how to answer the questions. Despite the lack of knowledge and experience in dealing with these types of tests, the format of questions was attractive to students. They were surprised of this type of questions and looked forward to receiving their score.

Statistical analysis

Data analysis was performed using SPSS software. The evaluation of the mean score of SCT, the GPA, the scores

and that of the final exam was performed by descriptive statistics. The test's reliability was calculated using Cronbach's alpha. For concurrent and predictive validity, Pearson correlation coefficient was used for the SCT, the final exam, and comprehensive score, respectively. Pearson correlation coefficient was used to evaluate the correlation between SCT and GPA score as well as the GPA and the final score. The test's item difficulty level (IDL) and item discriminative index (IDI) were determined using Whitney and Sabers' method.

Results

Internal reliability was calculated using Cronbach's alpha test, which was equal to 0.74. Table 3 demonstrates the correlation coefficient, item difficulty level, item discriminative index of each scenario, and correlation of each scenario with the total scores of the SCT. All scenarios were positively correlated with the test's total score. Thus, the total score and item discriminative index (IDI) for scenario 2 were 0.77 & .66, respectively; the lowest correlation coefficient was related to scenario 11, with the score of .43; and the lowest IDI was related to scenario 11, with the score of .32. An acceptable item difficulty level (IDL) was in the range of .3 to .8 (28). Table 3 demonstrates the item difficulty level of each scenario. The highest IDL belonged to scenario 3, with the score of .59, and the lowest IDL to scenario 9, with the score of .39.

The mean±SD GPA of all students during the sixth semester, from the score of 20, was 16.41±0.94, with the highest total average of 18.67 and the lowest of 14.65. The final score of students in gynecology was calculated from the score of 20, and the mean±SD score on this course was 16.74±1.14, with the highest score of 19.65 and the lowest of 14.8. The comprehensive test score was calculated from the score of 100 and the mean±SD score of the test was 67.27±7.83, with the maximum score of 91 and the minimum of 52.

Score 1 was given to each question of SCT test, and the score of this test was calculated from 45. The mean±SD score of SCT for midwifery experts was 42.73±1.86 and it

Table 3. ITC, IDL and IDI of each question

	SC1	SC2	SC3	SC4	SC5	SC6	SC7	SC8	SC9	SC10	SC11	SC12	SC13	SC14	SC15
ITC*	0.70	0.77	0.63	0.74	0.62	0.76	0.51	0.64	0.60	0.44	0.43	0.45	.063	0.51	0.53
IDL**	0.41	0.51	0.59	0.41	0.43	0.40	0.42	0.45	0.39	0.45	0.46	0.48	0.45	0.48	0.41
IDI***	0.46	0.66	0.43	0.48	0.52	0.62	0.35	0.52	0.50	0.35	0.32	0.36	0.52	0.39	0.38

Correlation was significant at 0.01 level (2-tailed).

*Item total correlation (ITC)

**Item difficulty level (IDL)

***Item discriminative index (IDI)

Table 4. Descriptive statistics of the scores based on the level of expertise

	Sample size	Minimum	Maximum	Mean	Std. Deviation
Faculty member	15	39.00	45.00	42.7333	1.86956
Student	55	18.45	38.30	26.41	3.29

Table 5. Correlation among the final scores of gynecology course, comprehensive exam, SCT test, and GPA

	SCT score	Comprehensive exam score	Grade point average	Gynecology final exam score
SCT score	1	0.721**	0.709**	0.654**
comprehensive exam score	0.721**	1	0.915**	0.799**
Grade point average	0.709**	0.915**	1	0.812**
Gynecology final exam score	0.654**	0.799**	0.812**	1

** Correlation was significant at the 0.01 level (one-tailed)

was 26.41 ± 3.29 for midwifery students. The mean scores of the participants are summarized in Table 4.

Table 5 displays the correlation among the final scores of gynecology course, comprehensive exam, SCT test, and GPA using Pearson correlation coefficient. The correlation coefficient between these items was positive and significant ($p < 0.001$). The lowest correlation belonged to gynecology final score and the SCT test score, which was 0.654, representing the concurrent validity of SCT test with the final score. Moreover, the highest correlation belonged to GPA and the score of comprehensive exam, which was equal to 0.915. Also, the correlation coefficient between the score of SCT and that of comprehensive exam, which represented predictive validity, was 0.721.

Discussion

In SCT test, students research and interpret the exclusive combination of clinical reasoning, which is the critical steps of this process (15). This test measures the degree of consistency of learners' performance in a series of clinical scenarios using a reference panel. SCT measures those aspects of reasoning and knowledge that are different from other measurement tools (19).

To our knowledge, this was the first SCT study on undergraduate midwifery students in gynecology. The ability of clinical reasoning and decision-making can play a major role in reducing the incidence of adverse events in the treatment of obstetrics and gynecology diseases. Thus, this study was conducted to evaluate the reliability and concurrent and predictive validity of SCT tests with the final exam of gynecology course and comprehensive exams of BA in gynecology course, respectively. In this study, we measured the clinical reasoning skills of students using SCT test. The test was able to reflect the level of expertise, and clinical experience was clearly correlated with SCT score (19). SCT test showed a significant difference in the mean scores of experts and students. Similar to other studies, the results of this study significantly confirmed the difference between the clinical reasoning skills of experts and students (29, 30). Thus, it can be concluded that this test can clearly distinguish the levels of expertise; this finding is similar to that of Demeester et al. in 2004 (24).

Because reliability is the prerequisite of validity, if a test is not repeatable, it will not be valid. This test, as shown in previous studies (19, 30-32), has been recognized as a valid and reliable instrument to assess clinical reasoning of midwifery students. This study showed relatively high internal reliability coefficient for this test (.74). A good test reliability for SCT is 0.7 to 0.8 (18, 26, 29), and different studies

have reported different internal consistency coefficients as Cronbach's alpha of 0.78 (1), 0.745 (23), 0.73 (32), .85 (30), 0.79 (21), 0.73 (31), .8 (28), 0.9 (19). Andrew et al. stated that based on SCT 3 or 5 test options, Cronbach's alpha coefficient varies from 0.78 to 0.68 (33). Also, Charlin et al. reported that SCT test with 60 to 50 questions is sufficient to achieve reliability of 0.8 (34).

Innovation of this study was to examine the concurrent and predictive validity of SCT test. The results of this study showed that the final scores of gynecology course and SCT test were highly correlated. Therefore, based on these results, the SCT test has concurrent validity with students' final exam, and if the faculty are trained to design, implement, and score, this test can be a good alternative or a complementary test for specific clinical courses that usually evaluate the students via a combination of MCQ and descriptive questions. However, this test was highly correlated with the comprehensive exam score (based on the syllabus of clinical expertise) held at the end of the academic period as a combination of MCQ questions and clinical skills assessment. Based on these results, the score of SCT test can be considered as a predictor of success of midwifery students in the comprehensive test. According to Goos et al. in 2016, SCT cannot measure the actual knowledge, so it does not correlate with MCQ tests that are used to measure the knowledge level (29). The high correlation of SCT with final and comprehensive tests may be due to the combination of the 2 tests, as the total score is the result of combining MCQ score and descriptive or clinical tests.

Based on these results, the validity of the test can be increased if the SCT test scenarios are taught based on test specifications table (This table contains the objectives and content of the test.). The highest correlation coefficient was between the GPA and comprehensive exam score, so it can be concluded that students with high GPA can be much more successful in comprehensive exams. The high correlation coefficient between GPA and SCT score could indicate that those who are generally more hardworking and have a higher GPA are more successful in clinical reasoning skills than others. Moreover, the clinical reasoning abilities of those who had higher scores than those who received a lower score were better and they were more successful in using their data more efficiently for clinical tasks (19).

Similar to the study of Lambert et al. (19), in this study, from the students and experts' perspective, the test was very interesting because the scenarios were consistent with what is done in clinical practice. This test could be applied

as a complement to other methods of assessing students' clinical reasoning in specialized clinical courses; moreover, its format allows the faculty to assess the ability of students in solving ambiguous clinical issues that they encounter daily (15).

The limitations of this type of test can be its hard design and unfamiliarity of teachers and students with answering such questions and correcting them (17, 29). Those problems can be easily solved by training teachers and involving them in SCT building. Thus, despite the difficulty of the design, implementation, and scoring, this test introduces a new perspective on assessment and evaluation of students' cognitive skills (29). Moreover, it can play an important role in the evaluation of students' clinical reasoning potentials and can be used as a comprehensive assessment tool to predict students' success in specialized clinical tests.

Conclusion

We hope that the results of this study can provide sufficient evidence to support further studies on increasing the reliability and predictive and concurrent validity of this test for midwifery undergraduate students. Because traditional tests, such as MCQ and descriptive tests, are used to assess midwifery students, we hope that conducting these studies using such tests as SCT would continue, so that appropriate tests could be designed to evaluate clinical midwifery students in specialized courses. Moreover, SCT could apply to predict the success rate of students in postgraduate entrance exam; so, through more evidence, predictive validity of SCT could be judged. Moreover, SCT can be used in our programs to train students. For future studies, it is recommended that 4 levels of undergraduate, graduate, doctoral, and academic faculty members of midwifery gather more precise evidence about SCT's power in different levels.

Acknowledgment

This study was the result of a research project approved by the Center of Medical Education Research, Iran University of Medical Sciences (project number: 95-01-133-28278). We thank Iran University of Medical Sciences for financial support of the project, and we extend our gratitude to the midwifery faculty and students who have helped us in the design and implementation of this project. Also, we thank those who have helped in translating and editing the manuscript.

Conflict of Interests

The authors declare that they have no competing interest.

References

1. Amini M, Moghadami M, Kojuri J, Abbasi H, Abadi AAD, Molaee NA, et al. An innovative method to assess clinical reasoning skills: Clinical reasoning tests in the second national medical science Olympiad in Iran. *BMC Res Notes*. 2011;4(1):1.
2. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. *Med Educ*. 2015;49(10):961-73.
3. Modi JN, Gupta P, Singh T. Teaching and assessing clinical reasoning skills. *Indian Pediatr*. 2015;52(9):787-94.
4. Forsberg E, Ziegert K, Hult H, Fors U, editors. Assessing progression of clinical reasoning through Virtual Patients. Transforming Healthcare through Excellence in Assessment and Evaluation, 16th Ottawa Conference, Ottawa, Ontario, Canada, April 25-29, 2014; 2014.
5. Zamani S, Amini M, Masoumi SZ, Delavari S, Namaki MJ, Kojuri J. The comparison of the key feature of clinical reasoning and multiple choice examinations in clinical decision making ability. *Biomed Res*. 2017;28(3).
6. Norman G. Research in clinical reasoning: past history and current trends. *Med Educ*. 2005;39(4):418-27.
7. Kassirer JP. Teaching clinical reasoning: case-based and coached. *Acad Med*. 2010;85(7):1118-24.
8. Pennaforte T, Moussa A, Loye N, Charlin B, Audétat M-C. Exploring a New Simulation Approach to Improve Clinical Reasoning Teaching and Assessment: Randomized Trial Protocol. *JMIR Res Protoc*. 2016;5(1).
9. Ashoorian V, Liaghatdar MJ, Adibi P. What variables can influence clinical reasoning? *J Res Med Sci*. 2012;17(12).
10. Elstein AS. Beyond multiple-choice questions and essays: the need for a new way to assess clinical competence. *Acad Med*. 1993;68(4):244-9.
11. Dory V, Gagnon R, Vanpee D, Charlin B. How to construct and implement script concordance tests: insights from a systematic review. *Med Educ*. 2012;46(6):552-63.
12. Lubarsky S, Dory V, Duggan P, Gagnon R, Charlin B. Script concordance testing: From theory to practice: AMEE Guide No. 75. *Med Teach*. 2013;35(3):184-93.
13. Boulouffe C, Charlin B, Vanpee D. Evaluation of clinical reasoning in basic emergencies using a script concordance test. *Am J Pharm Educ*. 2010;74(10):194.
14. Charlin B, Boshuizen H, Custers EJ, Feltovich PJ. Scripts and clinical reasoning. *Med Educ*. 2007;41(12):1178-84.
15. Fournier JP, Demeester A, Charlin B. Script concordance tests: guidelines for construction. *BMC Med Inform Decis Mak*. 2008;8(1):18.
16. Charlin B, van der Vleuten C. Standardized assessment of reasoning in contexts of uncertainty the script concordance approach. *Eval Health Prof*. 2004;27(3):304-19.
17. Lineberry M, Kreiter CD, Bordage G. Threats to validity in the use and interpretation of script concordance test scores. *Med Educ*. 2013;47(12):1175-83.
18. Duggan P, Charlin B. Summative assessment of 5 th year medical students' clinical reasoning by script concordance test: requirements and challenges. *BMC Med Educ*. 2012;12(1):29.
19. Lambert C, Gagnon R, Nguyen D, Charlin B. The script concordance test in radiation oncology: validation study of a new tool to assess clinical reasoning. *Radiat Oncol*. 2009;4(1):7.
20. Monajemi A. Clinical Reasoning: Concepts, Education and Assessment. Isfahan: Isfahan University of Medical Sciences; 2010. 125 p.
21. Sibert L, Charlin B, Corcos J, Gagnon R, Lechevallier J, Grise P. Assessment of clinical reasoning competence in urology with the script concordance test: an exploratory study across two sites from different countries. *Eur Urol*. 2002;41(3):227-33.
22. Amini M, Kojuri J, Karimian Z, Lotfi F, Moghadami M, Dehghani M, et al. Talents for future: Report of the second national medical science Olympiad in Islamic republic of Iran. *IRAN RED CRESCENT MED J*. 2011(6, Jun):377-81.
23. Nseir S, Elkalioubie A, Deruelle P, Lacroix D, Gosset D. Accuracy of script concordance tests in fourth-year medical students. *Int J Med Educ*. 2017;8:63.
24. Demeester A. Évaluation du raisonnement clinique des étudiants sages-femmes par filetest de concordance de script. Mémoire pour la maîtrise universitaire de pédagogie des sciences de la santé, Paris. 2004;13.
25. Gantelet M, Demeester A, Pauly V, Gagnon R, Charlin B. Impact of the reference panel on the psychometric quality of a script concordance test developed for midwifery training. *Pédagogie Médicale*. 2013;14(3):157-68.
26. Charlin B, Gagnon R, Pelletier J, Coletti M, Abi-Rizk G, Nasr C, et al. Assessment of clinical reasoning in the context of uncertainty: the effect of variability within the reference panel. *Med Educ*. 2006;40(9):848-54.
27. Norcini JJ, Shea JA, Day SC. The use of aggregate scoring for a recertifying examination. *Eval Health Prof*. 1990;13(2):241-51.

28. Iravani K, Amini M, Doostkam A, Dehbozorgian M. The validity and reliability of script concordance test in otolaryngology residency training. *Journal of advances in medical education & professionalism*. 2016;4(2):93.
29. Goos M, Schubach F, Seifert G, Boeker M. Validation of undergraduate medical student script concordance test (SCT) scores on the clinical assessment of the acute abdomen. *BMC Surg*. 2016;16.
30. Nough T, Boutros M, Gagnon R, Reid S, Leslie K, Pace D, et al. The script concordance test as a measure of clinical reasoning: a national validation study. *Am J Surg*. 2012;203(4):530-4.
31. Sibert L, Darmoni SJ, Dahamna B, Hellot M-F, Weber J, Charlin B. On line clinical reasoning assessment with Script Concordance test in urology: results of a French pilot study. *BMC Med Educ*. 2006;6(1):45.
32. Park AJ, Barber MD, Bent AE, Dooley YT, Dancz C, Sutkin G, et al. Assessment of intraoperative judgment during gynecologic surgery using the Script Concordance Test. *Am J Obstet Gynecol*. 2010;203(3):240. e1-. e6.
33. Bland AC, Kreiter CD, Gordon JA. The psychometric properties of five scoring methods applied to the script concordance test. *Acad Med*. 2005;80(4):395-9.
34. Charlin B, Roy L, Brailovsky C, Goulet F, van der Vleuten C. The Script Concordance test: a tool to assess the reflective clinician. *Teach Learn Med*. 2000;12(4):189-95.