Le test de concordance de script comme outil d'évaluation sanctionnante en cardiologie

Script concordance test as a sanctionnal evaluation in cardiology

Khadija Mzoughi[1], Ihsen Zairi[1], Mohamed Ali Kedous[2], Sana El Mhamdi [4], Mohamed Ben Dhiab[5], Fethia Mghaieth[3], Sondos Kraiem[1]

1- Service de cardiologie- Hôpital Habib Thameur- Université Tunis El Manar, Faculté de Médecine de Tunis
2- Service d'orthopédie- CTGB- Université Tunis El Manar, Faculté de Médecine de Tunis
3- Service des Explorations Fonctionnelles et de Réanimation Cardiologiques. Hôpital la Rabta -Université Tunis El Manar, Faculté de Médecine de Tunis
4- Service de médecine préventive et d'épidémiologie, CHU de Monastir -Faculté de médecine de Monastir
5-. Service de Médecine Légale - Hôpital Farhat Hached Sousse- Faculté de Médecine de Sousse

RÉSUMÉ

Introduction : L'évaluation sanctionnante du module de cardiologie est réalisée à l'aide de questions à choix multiples, de questions à réponses ouvertes et courtes, de cas cliniques et de questions rédactionnelles. Or, ces méthodes ne permettent pas d'évaluer la capacité de raisonnement clinique en situation d'incertitude, contrairement au test de concordance de script (TCS).

Objectif : Le but de notre travail était de comparer les notes obtenues par les étudiants en 3ème année médecine au TCS versus l'épreuve sanctionnante de cardiologie et d'étudier la corrélation entre ces deux méthodes d'évaluation.

Méthode : Il s'agit d'une étude prospective incluant 31 étudiants de 3ème année ayant réalisé leur stage de cardiologie au sein du service de cardiologie de l'hôpital Habib Thameur au cours du premier semestre en 2016.

Nous avons comparé les notes obtenues à l'épreuve de 13 TCS (39 items) à celles de l'épreuve sanctionnante de cardiologie.

Résultats : Le score moyen des externes au TCS était significativement inférieur à celui des experts (66,6 ± 10,2 vs 86±6,7%, p<0,0001). La note moyenne obtenue par les étudiants au TCS était significativement supérieure à celle de l'épreuve sanctionnante de cardiologie (p<0,001). Le coefficient alpha de Cronbach était de 0,71. Nous n'avons pas trouvé de corrélation entre les deux épreuves (r=0,329 ; p=0,07).

Conclusion: L'évaluation de nos étudiants par le TCS a montré des notes moyennes statistiquement plus élevées que les questions d'une épreuve classique, sans corrélation entre les épreuves. Ceci devrait nous inciter à intégrer le TCS dans nos méthodes d'évaluation afin de favoriser le raisonnement clinique.

Mots-clés

Test de concordance de script, Evaluation, Compétence, Raisonnement, cardiologie.

SUMMARY

Introduction: Sanctioning evaluation in cardiology is carried out using multiple choice questions, short-answer questions, clinical cases and editorial questions. However, these methods do not assess clinical reasoning in a context of uncertainty in contrast with script concordance tests (SCT).

Objective: The aim of our study was to compare the scores obtained by the students in the 3rd year of medicine with the SCT versus the sanctioning test of cardiology and to study the correlation between these two evaluation methods.

Methods: This is a prospective study including 31 3rd year students who completed their cardiology clerckship in the Cardiology Department of the Habib Thameur Hospital during the first half of 2016. We compared the scores obtained in the 13 SCT test (39 items) with those of the cardiology normative test.

Results: Students 'mean score at SCT was significantly lower than that of experts (66.6 ± 10.2 vs 86 ± 6.7%, p <0.0001).

The mean score obtained by students at the SCT was significantly higher than that of the cardiology sanctioning test (p <0.001). Cronbach alpha coefficient was 0.71. There was no correlation between the two tests (r = 0.329; p= 0.07).

Conclusion: The evaluation of our students by the SCT showed mean score statistically higher than the questions of a classic test, without correlation between them. This should encourage us to incorporate SCT into our assessment methods to promote clinical reasoning.

Key-words

Script concordance test, Evaluation, Competence, Reasoning, cardiology.

L'évaluation sanctionnante du module de cardiologie des étudiants en première année du deuxième cycle des études médicales (DCEM1) à la faculté de médecine de Tunis (FMT) se base sur une épreuve écrite semestrielle constituée de questions à choix multiples (QCM), de questions à réponses courtes et ouvertes (QROC), de cas cliniques et de questions rédactionnelles.
L'un des inconvénients de ce système d'évaluation est qu'il se limite à l'évaluation des connaissances dans un contexte clinique bien défini. Or, dans la pratique quotidienne, les professionnels de santé rencontrent des problèmes variés et complexes. Pour les résoudre, ils font appel à leurs connaissances, à leur expérience et à leur raisonnement clinique [1,2].
Le test de concordance de script (TCS) est une méthode d'évaluation basée sur la théorie hypothético-déductive et conçue pour mesurer la capacité de raisonner en contexte d'incertitude [3-5] .
Développé par Charlin et al. et apparu au début des années 2000, le TCS est devenu progressivement un outil d'usage courant utilisé dans de nombreuses spécialités pour l'évaluation formative [6-12]. Bien que sa validité soit clairement établie et que les modalités techniques de construction d'un test sont bien connues, le TCS reste peu utilisé dans l'évaluation sanctionnante [11,13-15].
Le but de notre travail était de comparer les notes obtenues par les étudiants en 3ème année médecine au TCS versus l'épreuve sanctionnante de cardiologie et d'étudier la corrélation entre ces deux méthodes d'évaluation.

METHODES

Population de l'étude

Nous avons réalisé une étude prospective analytique incluant les étudiants en DCEM1 de la FMT, ayant réalisé leur stage de cardiologie au courant du premier semestre de l'année universitaire 2016-2017 au sein du service de cardiologie de l'hôpital Habib Thameur.
L'ensemble des étudiants ont participé lors de leur stage de quatre semaines à :
-une séance d'allo-confrontation à l'aide d'une vidéo-case sur l'interrogatoire d'un patient présentant une douleur thoracique,
- une séance d'allo-confrontation à l'aide d'une vidéo-case sur l'interrogatoire d'un patient présentant une dyspnée,
-une séance d'apprentissage du raisonnement clinique sur le thème de la douleur thoracique,
-et une séance d'apprentissage du raisonnement clinique sur le thème de la dyspnée.

TCS

La construction du TCS a été basée sur les recommandations de l'université de Montréal. Le TCS était composé de 13 vignettes avec 3 items chacune, rédigées par un sénior du service (maitre de conférences agrégé) et relues par un deuxième sénior (assistant hospitalo-universitaire).
Chaque vignette est constituée d'un énoncé bref décrivant une situation clinique contextualisée, non exceptionnelle, susceptible d'être rencontrée par l'étudiant dans son exercice actuel ou futur.
Après avoir exposé le contexte clinique en quelques lignes, trois items sont proposés. Chaque item est composé de trois parties. La première partie est une attitude, une décision clinique, diagnostique ou thérapeutique, et est titrée « *si vous pensiez à…* ». La deuxième partie correspond à une nouvelle donnée dans cette même situation clinique et s'intitule « *et qu'alors vous apprenez que …* ».
Le 3ème item débute par « *alors cette nouvelle information rend votre attitude initiale…* » .
L'étudiant ou l'expert doit alors évaluer, sur une échelle de Likert, si l'attitude est plus ou moins adaptée en fonction cette nouvelle donnée.L'échelle de Likert a cinq entrées (ou ancrage) de -2,-1, 0, 1, et 2. Si l'information donnée rendait l'hypothèse, l'examen initialement prévu, la prescription ou le geste envisagé(e) plus probable ou plus justifié (e), selon le domaine de connaissance à évaluer, le score était de 1 ou 2 ; si elle était sans effet, il était de 0 ; si elle le (la)rendait moins probable ou moins justifiée, le score était de -1 ou -2.
La table de spécification du TCS est décrite dans le tableau I.

Tableau 1: table de spécification du TCS en cardiologie

Situation clinique problématique	Domaines d'évaluation des connaissances		
	Diagnostic	Investigation	Traitement
Douleur thoracique	X	X	
Dyspnée aigue	X		
Syndrome coronarien aigu avec sus-décalage du segment ST			X
Syndrome coronarien aigu sans sus-décalage du segment ST			X

Afin d'établir la grille de correction du TCS, un panel de référence ou « groupe d'experts » a été constitué. Les réponses de ces derniers vont servir de référence, chacune d'elles étant prise en compte. Notre panel était composé de 10 experts (2 professeurs, 3 maitres de conférences agrégés, 2 assistants hospitalo-universitaires et 2 cardiologues de libre pratique). Les consignes sur l'épreuve et la durée étaient précisées sur la première page du TCS.
Chaque groupe d'étudiants a été informé en fin de stage qu'ils seraient invités après la session d'examens de janvier 2018 à passer un TCS réalisé dans le cadre du projet d'un mémoire de pédagogie.

Les étudiants ont bénéficié d'une séance de préparation au TCS d'une heure pendant leur stage pendant laquelle les principes et les buts de cette méthode d'évaluation ont été expliqués.
Les étudiants ont été recontactés par mail afin de venir passer le TCS le mercredi 25 janvier 2017 à la salle de staff du service de cardiologie de l'hôpital Habib Thameur. La durée de l'épreuve était de 30 minutes. Chaque étudiant a reçu une copie du TCS. Les consignes ont été données avant le démarrage de l'épreuve.

L'épreuve écrite sanctionnante de cardiologie
L'épreuve de cardiologie était composée de deux QCM, six cas cliniques avec des QROC, et un cas clinque avec une question rédactionnelle.
Chaque question a été proposée par l'enseignant ayant assuré le cours à la faculté de médecine. La validation des questions a été ensuite réalisée lors d'une réunion de la section de cardiologie avec établissement des grilles de réponses.

Analyse statistique
L'analyse statistique a été réalisée à l'aide du logiciel Epi Info.
Les résultats ont été exprimés en moyennes ± écart-type.
Les comparaisons de moyennes entre les notes obtenues au TCS entre les expertes et les étudiants et entre le TCS et l'épreuve sanctionnante de cardiologie, ont été réalisées à l'aide du test de Wilcoxon.
Une corrélation entre la note obtenue aux TCS et celle de l'épreuve sanctionnante de cardiologie a été recherchée en utilisant le test de corrélation de Spearman.
Une valeur de p inférieure à 0,05 a été considérée comme significative.
La cohérence interne (fidélité) du TCS a été évaluée par le coefficient α de Cronbach à l'aide d'une feuille de calcul Excel téléchargeable sur le site de l'université de Montréal. S'il est supérieur à 0,70, cela signifie que l'on peut comparer les étudiants entre eux (un étudiant qui aura une meilleure note sera réellement meilleur), alors que si ce coefficient est faible, le manque de cohérence interne ne permet pas de comparer des étudiants entre eux.

RESULTATS

Trois groupes de 13 externes chacun ont réalisé leur stage au sein du service de cardiologie de l'hôpital Habib Thameur lors du premier semestre de l'année universitaire 2016-2017.
Trente-et-un externes ont participé à l'épreuve du TCS soit un taux de participation de 79%. Le sexe-ratio était de 2,9 (23 sexe féminin, 8 sexe masculin).
Les experts ont eu un score moyen, rapporté sur 100, de 86±6,7avec des extrêmes allant de 72,1 à 94,2%. Le score moyen des externes était de 66,6 ± 10,2 avec des extrêmes allant de 40,2 à 82,9%. Les notes moyenne des externes étaient significativement inférieures aux notes des experts (p<0,0001) comme indiqué dans la figure 1.

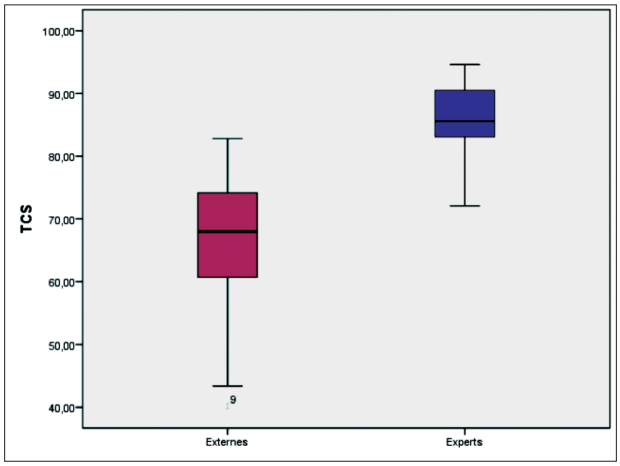

Figure 1: Comparaison des notes moyennes au TCS obtenues par les externes et les experts

La note moyenne obtenue par les étudiants au TCS était significativement supérieure à la note obtenue à l'épreuve sanctionnante de cardiologie (66,6 ± 10,2 versus 57±9%, p<0,001) comme indiqué dans la figure 2.

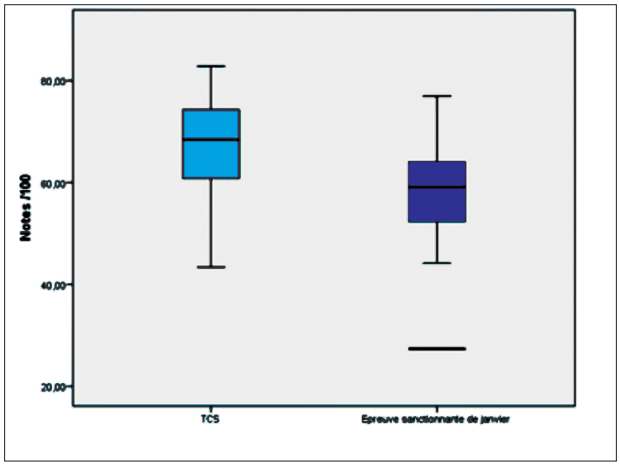

Figure 2 : Comparaison des notes moyennes des étudiants entre TCS et épreuve sanctionnante

Le coefficient alpha de Cronbach du TCS était de 0,71 pour nos externes.
Nous n'avons pas trouvé de corrélation entre les deux épreuves (r=0,329 ; p=0,07) comme indiqué dans la figure 3.

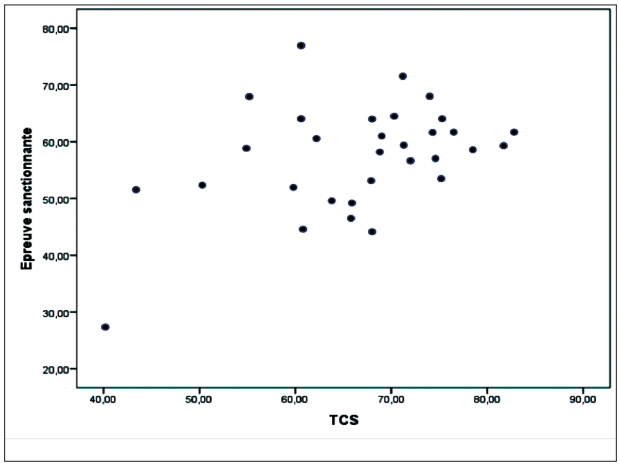

Figure 3 : Corrélation entre les notes obtenues au TCS et à l'épreuve sanctionnante

DISCUSSION

Nos résultats ont montré que la note moyenne obtenue par les étudiants au TCS était significativement supérieure à celle de l'épreuve sanctionnante de cardiologie, sans corrélation entre les deux épreuves.

Le coefficient alpha de Cronbach du TCS de notre étude était de 0,71 ; dénotant d'une bonne validité et fiabilité [12]. Les scores moyens des experts étaient significativement supérieurs à celui des apprenants de l'étude rendant valide le TCS de notre travail [4].

En 2001, Brailovsky et al. ont noté une corrélation significative entre le TCS et un cas clinique avec QROC (0,451, p=0,013) [16]. En revanche, l'examen clinique objectif structuré (ECOS) n'était pas corrélé au TCS (0,340, p=0,052). Leur étude a inclus 24 étudiants en médecine au Québec suivis pendant leur internat et leur résidanat. Les auteurs expliquent l'absence de corrélation entre TCS et ECOS par le fait que le TCS évalue le raisonnement clinique en condition d'incertitude alors que l'ECOS évalue les compétences cliniques [16]. La limite de ce travail reste son faible effectif.

En 2006, Fournier et al. n'ont pas trouvé de corrélation entre le TCS et les QCM dans l'évaluation du raisonnement clinique (r2 = 0,016, p = 0,59) dans un travail incluant 43 étudiants [17].

En 2008, Gibot et al. ont étudié la corrélation entre un TCS et un examen de type examen national classant avec dossier clinique en réanimation médicale [9]. L'étude a inclus 18 étudiants. Les auteurs ont trouvé une corrélation intermédiaire entre les deux épreuves (R=0,55, intervalle de confiance à 95% 0,10-0,81) [9].

En 2011, Jouneau et al. ont introduit le TCS dans l'examen écrit de pneumologie de 85 étudiants de DCEM2 à la faculté de médecine de Rennes [11]. L'épreuve comportait un cas clinique (2/3 du score) et un TCS (1/3 du score). Les auteurs ont noté une corrélation significative entre les cas cliniques et le TCS (R=0,304, p=0,005). Toutefois, ce coefficient de corrélation reste faible. De plus, le coefficient α de Cronbach était de 0,35 pour le TCS rendant sa cohérence interne beaucoup moins importante. Il est à noter que les étudiants de cette étude avaient bénéficié d'une séance d'une heure de préparation en groupe pour le TCS [11]. Cette séance de préparation est recommandée avant l'introduction du TCS dans l'évaluation afin de préparer les étudiants [12].

Plus récemment, Kelly et al. ont comparé la fiabilité et la validité des QCM aux TCS dans une étude ayant inclus 163 étudiants. Les auteurs ont conclu à une meilleure fiabilité du TCS en raison de sa meilleure corrélation avec la performance clinique. Toutefois, le questionnaire de satisfaction a montré une préférence des étudiants pour les QCM [18].

En 2017, Nseir et al. ont publié une large étude incluant 519 étudiants comparant le TCS à une épreuve de cas clinique QROC en cardiologie et en gynécologie [19].

Les scores en cardiologie étaient significativement plus élevés dans le TCS que les cas cliniques QROC (13,5±2,4 versus 11±2,2 ; p<0,001). A l'inverse, en gynécologie, les scores au TCS étaient significativement plus bas (10,8±2,6 versus 11,4±2,7 ; p=0,001). Cette différence entre les deux spécialités a été expliquée par la possible différence dans l'expérience clinique des étudiants. Les deux épreuves étaient significativement corrélées (p <0,05) [19].

Notre étude est l'une des premières à évaluer le TCS dans l'évaluation sanctionnante à la FMT. L'équipe de néonatologie de l'hôpital Charles Nicolle a introduit le TCS dans l'évaluation formative en 2015 [20].

Certaines limites de notre étude peuvent expliquer l'absence de corrélation entre les deux épreuves :

-le faible effectif de notre série,

-les thèmes abordés dans l'épreuve écrite sanctionnante étaient plus larges et le type de questions plus variées (QROC, QCM, cas cliniques) rendant la comparaison des deux épreuves plus difficiles. Le choix des thèmes de notre TCS a été basé sur les thèmes de la douleur thoracique et de la dyspnée aigue qui ont fait l'objet de séances d'apprentissage du raisonnement clinique et d'allo-confrontation par la méthode des vidéos lors du stage de cardiologie.

-le nombre réduit des items (39) ne permettait pas d'élargir les thèmes des vignettes. Ce choix est dû au fait que nos étudiants ont été convoqués après la session principale des examens de janvier, sur leur temps libre (un mercredi après-midi) ; une durée d'examen de 30 minutes a été jugé préférable.

-le TCS a été réalisé par un sénior du service et relu par un second. Les deux séniors n'avaient pas d'expérience antérieure dans l'élaboration du TCS. Ainsi, Gagnon et al recommandent la participation de trois experts à l'élaboration d'un TCS afin de multiplier les expériences et les jugements [21].

-le panel d'experts était composé de dix experts du secteur hospitalo-universitaire et libéral avec une expérience de cinq ans d'exercice au moins. Aucun des experts n'avaient d'expérience antérieure avec le TCS. Si le nombre de 10 experts est suffisant pour l'évaluation formative, le nombre de 15 au moins est nécessaire dans l'évaluation sanctionnante [12,22].

CONCLUSION

L'évaluation de nos étudiants par le TCS a montré des notes moyennes statistiquement plus élevées que les questions d'une épreuve classique. Ceci devrait nous inciter à intégrer le TCS dans nos méthodes d'évaluation afin de favoriser le raisonnement clinique tout en assurant une formation des enseignants des différentes spécialités à la construction, validation et notation du TCS et en initiant les étudiants à ce nouvel outil d'évaluation.

REFERENCES

1. Présentation - Centre de pédagogie appliquée aux sciences de la santé (CPASS) – Faculté de Médecine - Université de Montréal [Internet]. [cité 4 sept 2017]. Disponible sur: https://www.cpass. umontreal.ca/ recherche/ concordance/tcs/presentation_tcs/ http://www.cpass.umontreal.ca/tcs.html
2. Bordage G. Elaborated knowledge: a key to successful diagnostic thinking. Acad Med J Assoc Am Med Coll. 1994;69(11):883-5.
3. Charlin B, Desaulniers M, Gagnon R, Blouin D, van der Vleuten C. Comparison of an aggregate scoring method with a consensus scoring method in a measure of clinical reasoning capacity. Teach Learn Med. 2002;14(3):150-6.
4. Charlin B. Standardized Assessment of Reasoning in Contexts of Uncertainty: The Script Concordance Approach. Eval Health Prof.2004;27(3):304-19.
5. Charlin B, Kazi-Tani D, Gagnon R, Thivierge R. Le test de concordance comme outil d'évaluation en ligne du raisonnement des professionnels en situation d'incertitude. Rev Int Technol. 2005;2(2):22-7.
6. 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.
7. Sibert L, Charlin B, Gagnon R, Corcos J, Khalaf A, Grise P. Évaluation du raisonnement clinique en urologie: l'apport du test de concordance de script. Prog Urol. 2001;11:1213-9.
8. Marie I, Sibert L, Roussel F, Hellot M-F, Lechevallier J, Weber J. Le test de concordance de script : un nouvel outil d'évaluation du raisonnement et de la compétence clinique en médecine interne ? Rev Med Interne. 2005;26(6):501-7.
9. Gibot S, Bollaert P-E. Le test de concordance de script comme outil d'évaluation formative en réanimation médicale. Pédagogie Médicale. 2008;9(1):7-18.
10. Caire F, Marin B, Cuny E. Utilisation du test de concordance de script au cours du deuxième cycle des études médicales : expérience dans l'enseignement de la neurochirurgie. Pédagogie Médicale. 2011;12(1):29-35.
11. Jouneau S, Luraine R, Desrues B. Intérêt des tests de concordance de script pour évaluer le raisonnement et l'organisation des connaissances des étudiants de quatrième année des études médicales en France. Pédagogie Médicale. 2012;13(4):225-32.
12. Giet D, Massart V, Gagnon R, Charlin B. Le test de concordance de script en 20 questions. Pédagogie Médicale. 2013;14(1):39-48.
13. Fournier JP, Demeester A, Charlin B. Script concordance tests: guidelines for construction. BMC Med Inform Decis Mak. 2008; 6;8-18.
14. Charlin B, Gagnon R, Lubarski S, Lambert C, Meterissian S, Chalk C, et al. Assessment in the Context of Uncertainty Using the Script Concordance Test: More Meaning for Scores. Teach Learn Med. 2010;22(3):180-6.
15. Jouquan J. L'évaluation des apprentissages des étudiants en formation médicale initiale.Pédagogie Médicale. 2002;3(1):38-52.
16. Brailovsky C, Charlin B, Beausoleil S, Coté S, Van Der Vleuten C. Measurement of clinical reflective capacity early in training as a predictor of clinical reasoning performance at the end of residency: an experimental study on the script concordance test. Med Educ. 2001;35(5):430-436.
17. Fournier J-P, Thiercelin D, Pulcini C, Alunni-Perret V, Gilbert E, Minguet JM, et al. Évaluation du raisonnement clinique en médecine d'urgence : les tests de concordance des scripts décèlent mieux l'expérience clinique que les questions à choix multiples à contexte riche. Pédagogie Médicale. 2006;7(1):20-30.
18. Kelly W, Durning S, Denton G. Comparing a script concordance examination to a multiple-choice examination on a core internal medicine clerkship. Teach Learn Med. 2012;24(3):187-193.
19. Nseir S, Elkalioubie A, Deruelle P, Lacroix D, Gosset D. Accuracy of script concordance tests in fourth-year medical students. Int J Med Educ. 2017.25;8:63-69.
20. Ben Hamida A. Mémoire. Le test de concordance de script comme outil d'évaluation en néonatologie. M´édecine humaine et pathologie. 2015
21. Gagnon R, Charlin B, Lambert C, Carrière B, der Vleuten V. Script Concordance Testing: More Cases or More Questions? Adv Health Sci Educ 2009; 14:1382-96.
22. Gagnon R, Charlin B, Coletti M, Sauve E, Van Der Vleuten C. Assessment in the context of uncertainty: how many members are needed on the panel of reference of a script concordance test? Med Educ. 2005;39(3):284-91.