

Etude exploratoire de l'utilisation du test de concordance de script pour l'évaluation des apprentissages de résidents de médecine interne au Liban.

An Exploratory Study of the Use of Script Concordance Test for the Assessment of Internal Medicine Residents' Learning in Lebanon.

Fady HADDAD¹, Dany GASPARD¹, Marwan NASR²,
Simon ABOU JAOUDE², Elie NEMR²

Résumé *Contexte* : Le Test de concordance de script (TCS) est une méthode récente utilisée dans l'évaluation du raisonnement clinique, dont la validité a été démontrée dans les études nord-américaines et européennes. **But** : Vérifier la validité de cette propriété du TCS dans un milieu culturel différent au Moyen-Orient et comparer le TCS à une autre méthode d'évaluation utilisée en médecine, les questions à choix multiple (QCM). **Sujets et Méthodes** : Un TCS de 13 cas de médecine interne, comportant 120 items, a été administré à deux groupes de candidats : les internes de sixième année de médecine et les résidents de première année de médecine interne. Une comparaison a été faite entre les résultats des internes ($n = 25$) et ceux des résidents ($n = 21$) par un test de Student. La fidélité du test a été mesurée par le coefficient Alpha de Cronbach. Ensuite, un test de corrélation linéaire a été utilisé pour comparer les résultats du TCS avec ceux d'un test comportant des QCM. **Résultats** : La moyenne obtenue par les internes était de 57,86/100, celle des résidents était de 62,12/100. La différence était significative entre les 2 groupes ($p = 0,0153$). L'indice Alpha de Cronbach était de 0,69. Il n'y a pas eu de corrélation entre le TCS et le test de QCM. **Conclusion** : Dans un contexte culturel et académique différent de ceux où il a déjà été testé, le TCS apparaît être un bon outil d'évaluation du raisonnement clinique, permettant de discriminer les candidats selon leur niveau d'expertise. L'absence de corrélation entre TCS et QCM montre que ces deux tests ne sont pas identiques, mais plutôt complémentaires dans une évaluation globale de la compétence médicale.

Mots clés Test de concordance de script ; raisonnement clinique ; évaluation ; résidents ; internes.

Abstract *Introduction*: The Script Concordance Test (SCT) is a novel method used for the assessment of clinical reasoning. Its validity was demonstrated by North American and European studies. **Goal**: To verify the validity of the SCT in a different cultural environment in the Middle-East and to compare the SCT to an other evaluation method used in medicine, that is the multiple choice assessment method. **Subjects and Methods**: A SCT including 13 internal medicine cases, which comprised 120 items, was distributed to two groups of participants: 6th year interns and 1st year internal medicine residents. Scores obtained from interns ($n = 25$) and residents ($n = 21$) were compared using a Student test. The test reliability was analyzed using a Cronbach alpha coefficient. Scores obtained from the SCT and the multiple choice assessment method were compared using a linear correlation test. **Results**: The mean score value obtained from interns and residents was 57,86/100 and 62,12/100, respectively. The difference between both groups was statistically significant ($p = 0,0153$). Cronbach alpha coefficient was 0,69. No correlation was found between the SCT and the multiple choice assessment method. **Conclusions**: In a different cultural and academic context, the SCT seems to be a good clinical reasoning assessment tool to discriminate participants according to their level of expertise. The lack of correlation between the SCT and the multiple choice assessment method suggests that both tests are distinct and are also complementary to assess the overall medical competency.

Key words Script concordance test; clinical reasoning; assessment; residents; interns.

Pédagogie Médicale 2008;9:135-40

Introduction

Dans sa pratique médicale, le clinicien est confronté à deux genres de problèmes cliniques^{1, 2} : certains sont dits « bien définis » lorsque tous les éléments sont présents et qu'ils peuvent être résolus avec un haut degré de certitude ; d'autres sont dits « mal définis » lorsque tous les éléments utiles pour un diagnostic de certitude sont absents ou discordants. Face à de telles situations et dans le cadre d'un processus de raisonnement clinique de type analytique, l'interrogatoire orienté permet au clinicien de générer des hypothèses diagnostiques et d'inclure ou d'exclure certaines hypothèses selon les éléments nouveaux qui sont recueillis. Le diagnostic final dépend de la qualité de ce processus et de la richesse et de la qualité de l'organisation des connaissances sur lequel s'appuie ce raisonnement, lesquelles sont très habituellement liées à l'expérience personnelle du médecin par rapport à la famille de situations dont relève ce problème particulier.

Plusieurs outils sont nécessaires pour l'évaluation des compétences des futurs médecins. Les questions à choix multiples (QCM) sont appropriées pour l'évaluation du raisonnement face à des « situations bien définies » et leur résultat reflète surtout les connaissances déclaratives^{3, 4}. Les examens cliniques objectifs structurés (ECOS) évaluent surtout l'habileté de recueil des données cliniques et les habilités techniques^{4, 5}. Cependant, leur standardisation et leur applicabilité à un grand nombre de candidats demeurent un obstacle.

Dans une approche intégrée d'évaluation de la compétence clinique, un nouvel outil, le test de concordance de script (TCS), a été développé depuis plusieurs années. Ce test cherche à évaluer de manière plus pertinente le raisonnement clinique en situation d'incertitude et devant des problèmes complexes non consensuels. Il a été validé par plusieurs études nord-américaines et européennes^{3, 6, 7, 9, 10}. L'évaluation des apprentissages étant à la fois une pratique pédagogique et une pratique sociale, il paraît légitime de vérifier si les caractéristiques d'un milieu socioculturel et académique différent affectent les qualités du TCS. La stabilité de ce test a été validée dans des environnements culturels et linguistiques différents (France et Canada)⁷. Cependant, au Moyen-Orient, aucune étude concernant

la validité, l'applicabilité et la pertinence d'un TCS n'a, à ce jour, été menée.

La faculté de médecine de l'Université Saint-Joseph (USJ) de Beyrouth a adopté depuis 1996 un curriculum centré sur l'étudiant ; ainsi l'apprentissage par problème et l'apprentissage du raisonnement clinique (ARC) ont été introduits. Parallèlement, une diversification des outils d'évaluation des apprentissages a été entreprise afin de mieux cibler les différentes composantes de la compétence clinique ; en plus des QCM à niveaux taxonomiques variables, des examens cliniques objectifs et structurés (ECOS) ont été introduits et constituent maintenant une partie importante du dispositif d'évaluation. Concernant le TCS, des projets pilotes sont en cours ; ainsi un TCS formé de 13 cas cliniques de médecine interne a été construit avec deux objectifs, le premier étant d'évaluer la validité de cet instrument dans le milieu médical libanais et le second étant d'argumenter la pertinence qu'il y aurait à ajouter cet outil dans l'évaluation du raisonnement clinique de nos résidents.

Matériel et méthodes

Un examen utilisant le format du TCS a été construit par deux experts (deux professeurs de médecine interne dont l'un a une expertise pédagogique), après avoir défini un certain nombre de situations cliniques fréquemment rencontrées en pratique médicale ; 13 cas de médecine interne, 29 questions et 120 items ont été élaborés suivant une procédure standardisée² concernant le choix des cas et la répartition des questions en trois domaines, à savoir le diagnostic, l'investigation paraclinique et le traitement. Les réponses d'un panel d'experts formé de sept médecins internistes ayant une expérience clinique globale reconnue par leurs pairs ont déterminé le score attribué à chaque item.

L'examen a été administré à deux groupes de candidats pendant une période d'une heure : les internes de sixième année de médecine (n = 25) et les résidents de première année de médecine interne (n = 21). Notre cursus d'études médicales dure sept années, parmi lesquelles les deux dernières, c'est-à-dire la sixième et la septième, forment l'internat. Après les sept ans, les étudiants présen-

1- Département de médecine interne. Université Saint-Joseph, Beyrouth, Liban

2- Département de pédagogie médicale. Université Saint-Joseph, Beyrouth, Liban

Correspondance : Fady Haddad. Département de médecine interne, Hôpital Hôtel-Dieu de France, Ashrafieh, Beyrouth (Liban).

Téléphone : 961 - Télécopie : 961 139 77 72. Mailto : prfghaddad@yahoo.com

tent un concours qui leur permet d'accéder au résidanat où ils choisissent l'une des filières de médecine, chirurgie, anesthésie ou pédiatrie. Ainsi, deux années d'expérience et de pratique clinique séparent ces deux groupes. Trois semaines plus tard, les résidents de médecine interne ont présenté le « *in training exam* ». Cet examen est préparé annuellement par le conseil de spécialité de la médecine interne (*Board of Internal Medicine*) ; il est formé de 340 QCM et couvre tous les domaines de la médecine. La durée de cet examen est de sept heures. Le coefficient alpha de Cronbach a été calculé pour mesurer la fidélité (consistance interne) du test. Le test de Student a permis de comparer les moyennes des résultats des internes et des résidents. Le test de corrélation linéaire a été utilisé pour comparer les résultats du TCS et du « *in training exam* ».

Résultats

Le coefficient alpha de Cronbach pour l'ensemble de l'examen était de 0,69, ce qui atteste une bonne fidélité (consistance interne) du test. La moyenne des résultats des internes était de 57,86 alors que celle des résidents était de 62,12, ce qui constitue une différence significative ($p = 0,0153$) (*Tableau 1*).

Il n'y a pas de corrélation significative entre les résultats au TCS et ceux à l'*in training exam* (coefficient de corrélation linéaire à 0,1932 ; $p = 0,3657$) (*Figure 1*).

Figure 1 :
Comparaison des résultats des résidents de médecine interne au test de concordance de script (TCS) et au « *in training test* »

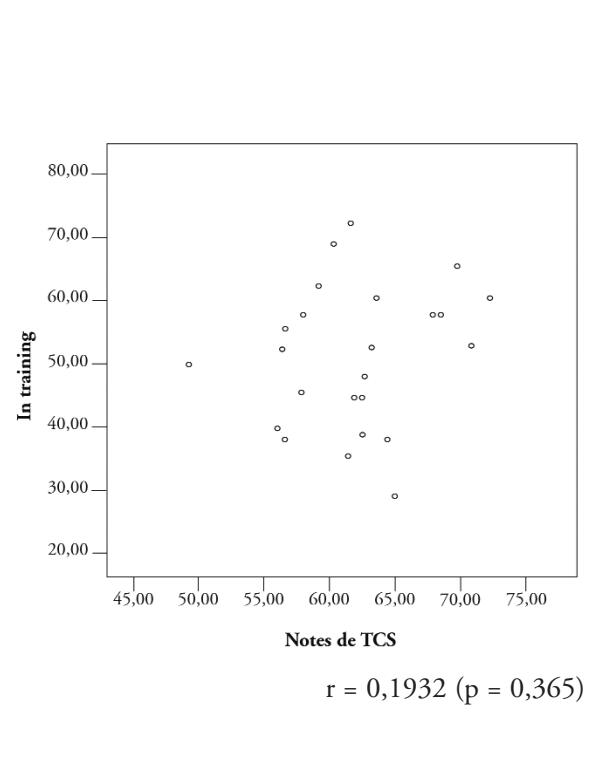


Tableau 1 :
Moyennes des scores obtenus au test de concordance de script par les internes de sixième année de médecine générale et les résidents de médecine interne

| | Effectif | Moyenne | Maximum | Minimum | Ecart-type |
|------------------|----------|---------|---------|---------|------------|
| Internes | 25 | 57,86 | 67,33 | 42,83 | 5,77 |
| Résidents | 21 | 62,12 | 72,29 | 49,61 | 5,58 |

$p < 0,05$ (test de Student)

Discussion

Il est établi que le TCS est utile et pertinent pour l'évaluation du raisonnement clinique dans la mesure où il explore des dimensions différentes du raisonnement clinique que celles qu'explorent les tests précédemment disponibles. Différentes études, dans plusieurs pays européens et nord-américains, montrent que les scores obtenus au TCS sont significativement plus élevés au fur et à mesure que l'expérience médicale grandit, ce qui suggère une bonne validité de construit^{3, 6, 10, 11}.

La présente étude, menée dans une population académique du Moyen-Orient, montre aussi de meilleurs résultats lorsque le test est administré à des étudiants résidents en médecine interne, ayant deux années d'expérience médicale supplémentaire par rapport à des internes de sixième année. Il est à noter que les internes, avant de présenter le TCS, ont bénéficié durant une année de séances d'apprentissage au raisonnement clinique (ARC) et du module de thérapeutique, alors que les résidents terminaient leur troisième année de stages hospitaliers.

Les résultats de cette étude peuvent être comparés avec ceux de l'étude de Sibert *et al.*, qui montraient que la validité de construit d'un TCS d'urologie gardait une bonne stabilité dans deux environnements culturels différents, en l'occurrence la France et le Canada⁷. Les résultats étaient corrélés avec le niveau d'éducation médicale dans les deux milieux. Dans une population du Moyen-Orient, l'amélioration du score avec l'expérience montre que le raisonnement clinique est un processus universel probablement indépendant du milieu culturel. Cependant, il serait plus intéressant de comparer les résultats d'un même TCS administré au Canada, en France et au Liban.

Cette étude a répondu aussi à la question de la nature des outils à incorporer au répertoire des examens d'évaluation au niveau postgradué. L'absence de corrélation entre les résultats du *in training test* et du TCS suggère, en effet, que ce dernier évalue un domaine différent de l'expertise médicale, en l'occurrence probablement davantage en lien avec le processus de raisonnement clinique. Ceci soutient la notion d'une complémentarité des outils d'évaluation^{1, 12, 13}. Par ailleurs, le TCS est réputé rendre mieux compte des situations complexes comparables à celles de la pratique réelle et non seulement des situations dont la prise en charge est strictement établie par des recommandations ou des guides de pratique clinique⁹. Dans cette optique, deux études ont comparé le TCS à d'autres méthodes d'évaluation clinique. Brailovsky et Charlin⁵ ont mené une étude au

Québec en 2001 pour déterminer la valeur prédictive d'un TCS administré avant le résidanat à l'égard des performances après le résidanat. Le TCS a été comparé à trois autres méthodes déjà établies, validées et utilisées au Québec, que sont le *Short answer management problem (SAMP)* (42 cas cliniques suivis chacun de cinq questions rédactionnelles, qui testent la capacité de raisonner et de prendre des décisions concernant les méthodes diagnostiques, investigationnelles, thérapeutiques ou de suivi appropriées), le *Simulated office orals (SOO)* (cinq simulations de cas cliniques de 15 minutes chacune où le rôle du patient est joué par le médecin et dont le but est de tester l'aspect biopsychologique de la médecine et surtout la relation médecin-patient), et l'ECOS (13 stations de simulation de situations cliniques où des capacités spécifiques sont testées chez le candidat, notamment la collecte des données du patient, l'examen clinique et les capacités techniques). Les examens avaient été présentés par les étudiants deux ans après le TCS. Les résultats ont montré une bonne corrélation entre le TCS, le SAMP et le SOO, qui, de part leur construction, constituent une bonne mesure du raisonnement clinique. Une absence de corrélation a été notée entre le TCS et l'ECOS. Ainsi, les auteurs ont pu démontrer que le TCS est une bonne méthode d'évaluation du raisonnement clinique et non des seules capacités techniques, qu'il a une bonne valeur prédictive de l'organisation ultérieure des connaissances chez les étudiants de médecine et qu'il serait complémentaire des ECOS.

La deuxième étude est celle menée par Fournier *et al.*³ à Nice en 2003. Elle avait pour but de comparer le TCS aux QCM en ce qui concerne l'aptitude d'un test à évaluer le niveau d'expérience clinique en médecine d'urgence ; 60 QCM et 30 TCS rédigés selon les recommandations publiées ont été administrés à 20 résidents de première année, 16 étudiants de fin de sixième année et à neuf médecins seniors en poste dans des services d'urgences. Les indices alpha de Cronbach ont été déterminés et ont montré une bonne validité des deux tests. Les résultats des différents groupes ont été comparés pour chacun des deux types de test, puis les tests ont été comparés entre eux. Cette étude a démontré d'une part, que seul le TCS permettrait une bonne prédiction du niveau d'expérience clinique avec des résultats significativement plus élevés dans les groupes consécutifs. Cela n'a pas été le cas pour les QCM où les résidents seniors ont eu des notes supérieures aux médecins. D'autre part, les auteurs ont trouvé que les résultats des deux tests variaient dans le même sens, mais sans corrélation significative entre eux. Ceci suggère que TCS et QCM

testent deux aspects différents et complémentaires de la compétence clinique globale.

Les limites de notre étude sont liées aux effectifs réduits des deux populations étudiées et au nombre restreint d'experts. L'étude faite par Gagnon *et al*⁴ en 2005 a montré qu'un panel de référence composé de 10 à 20 experts avait une bonne représentativité et améliorait la validité des résultats.

En conclusion, le TCS évalue un aspect essentiel de l'expertise médicale qu'est le raisonnement clinique. Il est complémentaire d'autres outils d'évaluation, tels que les QCM et l'examen clinique objectif structuré (ECOS). Nous proposons de considérer son inclusion dans les examens classants des internes et des résidents libanais au cours de leur cursus médical.

Contributions

Fady Haddad a participé à la conception du protocole de recherche, au recueil des données, à l'interprétation des résultats, à l'analyse statistique et à l'écriture du manuscrit. Dany Gaspard a participé au recueil des données, à l'interprétation des résultats, à l'analyse statistique et à l'écriture du manuscrit. Marwan Nasr et Simon Abou Jaoude ont participé à l'interprétation des résultats et à l'écriture du manuscrit. Elie Nemr a participé à la conception du protocole de recherche, à l'interprétation des résultats et à l'écriture du manuscrit.

Références

1. Charlin B, Roy L, Brailovsky CA, Van der Vleuten C. The script concordance test: a tool to assess the reflective clinician. *Teach Learning Med* 2000;12:189-95.
2. Charlin B, Gagnon R, Sibert L, Van der Vleuten C. Le test de concordance de script, un instrument d'évaluation du raisonnement clinique. *Pédagogie Médicale* 2002;3:135-44.
3. Fournier JP, 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;7:20-30.
4. Charlin B, Bordage G, Van der Vleuten C. L'évaluation du raisonnement clinique. *Pédagogie Médicale* 2003;4:42-52.
5. 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;35:430-6.
6. 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;26:501-7.
7. Sibert L, Charlin B, Corcos J, Gagnon R, Grise P, Van der Vleuten C. Stability of clinical reasoning assessment of competence in urology with the Script Concordance Test across two sites from different linguistic, cultural and learning environments. *Med Teach* 2002;24:522-7.
8. 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;41:227-33.
9. Charlin B, Brailovsky CA, Brazeau-Lamontagne L, Samson L, Leduc C. Script Questionnaires: Their Use for Assessment of Diagnostic Knowledge in Radiology. *Med Teach* 1998;20:567-71.
10. Meterissian S, Zabolotny R, Gagnon R, Charlin B. Is the script concordance test a valid instrument for assessment of intraoperative decision-making skills? *Am J Surg* 2007;193:248-51
11. Sibert L, Darmoni S, Dahamna B, Weber J, Charlin B. Online clinical reasoning assessment with the Script Concordance test: a feasibility study. *BMC Med Inform Decis Mak.* 2005;5:18.
12. Charlin B, Brailovsky CA., Leduc C, Blouin D. The Diagnostic Script Questionnaire: A New Tool to Assess a Specific Dimension of Clinical Competence. *Adv Health Sci Educ Theory Pract* 1998;3:51-8.
13. Charlin B, van der Vleuten C. Standardized assessment of reasoning in context of uncertainty: the script concordance approach. *Eval Health Prof* 2004; 27: 304-19.
14. 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;39:284-91.

Manuscrit reçu le 13 mai 2008; commentaires éditoriaux formulés aux auteurs le 27 juin 2008 ; accepté pour publication le 1^{er} juillet 2008.