

PERTINENCE DU TEST DE CONCORDANCE DE SCRIPT DANS L'ÉVALUATION DES COMPÉTENCES EN CARDIOLOGIE

RELEVANCE OF SCRIPT CONCORDANCE TEST IN CARDIAC COMPETENCY ASSESSMENT

L.ABID^{1,3,4,*}, B.HAMMAMI^{2,4}, R.HAMMAMI^{1,4}, S.CHARFEDDINE^{1,4}, A.BAHOUL^{1,4}, T.ELLOUZ^{1,4}, J.ALOULOU^{3,4}, L.GARGOURI^{3,4}, N.HENTATI^{3,4}, D.ABID^{1,4}, S.KAMMOUN^{1,4}

1. Service de cardiologie, Hôpital universitaire Hédi Chaker, Sfax

2. Service de maladies infectieuses, Hôpital universitaire Hédi Chaker, Sfax

3. Direction de la Pédagogie et de Technologies Educatives

4. Faculté de Médecine, Université de Sfax, Tunisie

*E-mail de l'auteur correspondant : leilaabid@yahoo.fr

Résumé

Introduction: Le test de concordance de script (TCS) évalue le raisonnement clinique en contexte d'incertitude.

Sujets et méthodes: Nous avons comparé les résultats au TCS des étudiants ayant effectué un stage de Cardiologie (groupe 1) à ceux n'ayant pas effectué un tel stage (groupe 2). Ces résultats ont aussi été comparés à ceux d'un examen de connaissances de cardiologie préalable aux stages.

Résultats: L'étude a porté sur 44 étudiants dont 22 dans le groupe 1 et 22 dans le groupe 2. Leurs notes à l'examen de connaissances en cardiologie dans le groupe 2 étaient meilleures que celles du groupe 1 ($80,56 \pm 6,37$ vs $65,45 \pm 9,80$, $p < 0,001$). Après le stage de cardiologie, les résultats des 2 groupes au TCS n'étaient pas statistiquement différentes ($53,89 \pm 13,1$ groupe 1 vs $56,04 \pm 7,92$ groupe 2) ($p = 0,51$).

Conclusion : Les résultats du TCS confirment le rôle bénéfique du stage hospitalier de cardiologie dans l'acquisition des compétences cliniques.

Mots clés : Test de concordance de script ; Compétence clinique ; Stage ; Evaluation ; Raisonnement clinique.

Abstract

The Script concordance test (SCT) evaluates clinical reasoning in context of uncertainty.

Aims: Evaluate cardiac training influence on cardiology skills by SCT in students of Sfax faculty of Medicine. We compared SCT results in students who performed a cardiology training (group 1) and those who have not made such a training (group 2). These results were also compared with a review of prior knowledge in cardiology. The study focused on 44 students: 22 (50%) in group 1 and 22 (50%) in group 2. Their marks in the examination of knowledge in cardiology of the group 2 were better than the group 1. Comparable ($80,56 \pm 6,37$ vs $65,45 \pm 9,80$, $p < 0,001$). After the cardiology training, SCT results weren't different between the 2 groups ($53,89 \pm 13,1$ groupe 1 vs $56,04 \pm 7,92$ groupe 2) ($p = 0,51$). In group 1, 16 students (68%) have improved their range when 3 students only of the group 2 did it ($P < 0,001$).

The SCT results confirm the beneficial role of cardiology training on acquiring clinical skills.

Keywords : Script Concordance test; Clinical competence; Evaluation; Training; Clinical reasoning.

ملخص:

مقدمة: اختبار التوافق سكريبت يقيم التفكير السريري في سياق عدم اليقين. الموضوعات والأساليب: قارنا نتائج توافق سكريبت بين الطلاب الذين أكملوا تدريب القلب (المجموعة 1) مع أولئك الذين لم يكن لديهم مثل هذا التدريب (المجموعة 2). كما تمت مقارنة هذه النتائج مع تلك التي وقعت فيها عملية اختبار المعرفة في أمراض القلب قبل القيام بالتدريب. النتائج: شملت الدراسة 44 طالبا، 22 منهم كانوا في المجموعة 1 و 22 في المجموعة 2. وكانت نتائجهم في اختبار المعرفة القلبية في المجموعة 2 أفضل من تلك الموجودة في المجموعة 1 ($80,56 \pm 6,37$ مقابل $65,45 \pm 9,80$, $p < 0,001$). بعد دورة التدريب في قسم أمراض القلب، لم تكن نتائج المجموعتين مع اختبار توافق سكريبت مختلفة إحصائيا ($53,89 \pm 13,1$ في المجموعة 1 مقابل $56,04 \pm 7,92$ في المجموعة 2) ($p = 0,51$). الاستنتاج: تؤكد نتائج اختبار سكريبت الدور المفيد للتدريب الإستشفائي في أمراض القلب في اكتساب المهارات السريرية.

الكلمات المفتاحية: اختبار توافق سكريبت ; الكفاءة السريرية ; التدريب ; التقييم ; المنطق السريري.

INTRODUCTION

Le système d'évaluation de l'étudiant en Médecine à Sfax se base essentiellement sur des examens « classiques » quantifiant essentiellement le degré d'acquisition des connaissances enseignées. Ces examens emploient des questions rédactionnelles, des questions à choix multiples (QCM), des questions à réponses courtes (QROCs), et des cas cliniques. Or, la compétence clinique décrite par Norman [1] regroupe plusieurs dimensions dont le jugement clinique et la capacité à résoudre les problèmes. Bien que le raisonnement clinique soit une composante majeure de la compétence clinique, il existe peu d'outils d'évaluation de cette compétence. A l'inverse de la plupart des outils existants qui explorent les connaissances théoriques factuelles (QCM, QROCS, Cas cliniques..), le test de concordance de script (TCS), développé par Bernard Charlin [2], permet d'évaluer la compétence à raisonner en situation d'incertitude et pour des problèmes complexes qui ne peuvent se résoudre par une simple application de connaissances. Le TCS explore une des étapes du raisonnement clinique, l'interprétation des données cliniques, étape cruciale du processus de raisonnement. Les études ont confirmé que le TCS peut être un outil d'évaluation aussi bien formative que sommative intéressant et prometteur. Toutes les dimensions de la compétence clinique ne peuvent être mesurées par un seul outil d'évaluation. Dans une approche globale de l'évaluation de l'acquisition de la compétence clinique en enseignement médical, il serait intéressant de combiner l'utilisation des TCS, des QCM, des QROCs etc.....

Pour la première fois à Sfax, à travers cette étude, nous avons introduit et testé ce nouvel instrument d'évaluation : le TCS.

L'objectif principal de cette première étude tunisienne est d'évaluer, par un TCS portant sur des domaines particuliers de la cardiologie clinique, les différences de compétences entre deux groupes d'étudiants de deuxième cycle de la faculté de Médecine de Sfax selon qu'ils ont eu ou pas un stage hospitalier de cardiologie.

L'objectif secondaire est de répondre à la question suivante « Est-ce que les étudiants qui ont les meilleures connaissances (meilleures notes en 3^{ème} année) sont ceux qui auront les meilleures compétences (notes du TCS en 6^{ème} année) ? ».

SUJETS ET METHODES

1. Composition et Construction du TCS

Dans notre étude, le TCS a été administré à des étudiants en fin de deuxième cycle des études médicales (6^{ème} année de médecine) qui ont tous réussi leurs examens de cardiologie en 3^{ème} année de Médecine.

Le TCS a été construit selon les recommandations de Dory et al [3]. Nous avons d'abord sélectionné une série de situations cliniques représentatives de la pratique en cardiologie, en termes de: fréquence en pratique clinique, mode de présentation, gravité et possibilités thérapeutiques.

Toutes les situations sélectionnées correspondent à des objectifs d'enseignement de la cardiologie en deuxième cycle à la faculté de médecine de Sfax.

Pour chacune de ces situations, différents paramètres ont été spécifiés:

- les hypothèses pertinentes de diagnostic, d'investigation ou de traitement
- les principales investigations complémentaires à demander pour résoudre le problème
- les informations cliniques, positives ou négatives, qui seraient recherchées pour vérifier les hypothèses.

Ensuite, chaque situation clinique a été rédigée sous la forme de vignette (scénario de quelques lignes) formulée de manière concise et précise. Pour chaque scénario de cas clinique, plusieurs hypothèses ont été proposées sous la forme d'une série d'items (figure n°1).

Le format d'un item dépendait de l'objectif d'évaluation : diagnostic, investigation, thérapeutique. Dans tous les cas, chaque item a comporté trois parties:

- la première partie comprend une hypothèse diagnostique ou une option thérapeutique ou une investigation (figures n°2, 3, 4)
- la deuxième partie présente une information nouvelle (par exemple, une manifestation clinique, un résultat d'exploration complémentaire, etc.)
- la troisième partie est une échelle ordinale de type Likert à cinq points

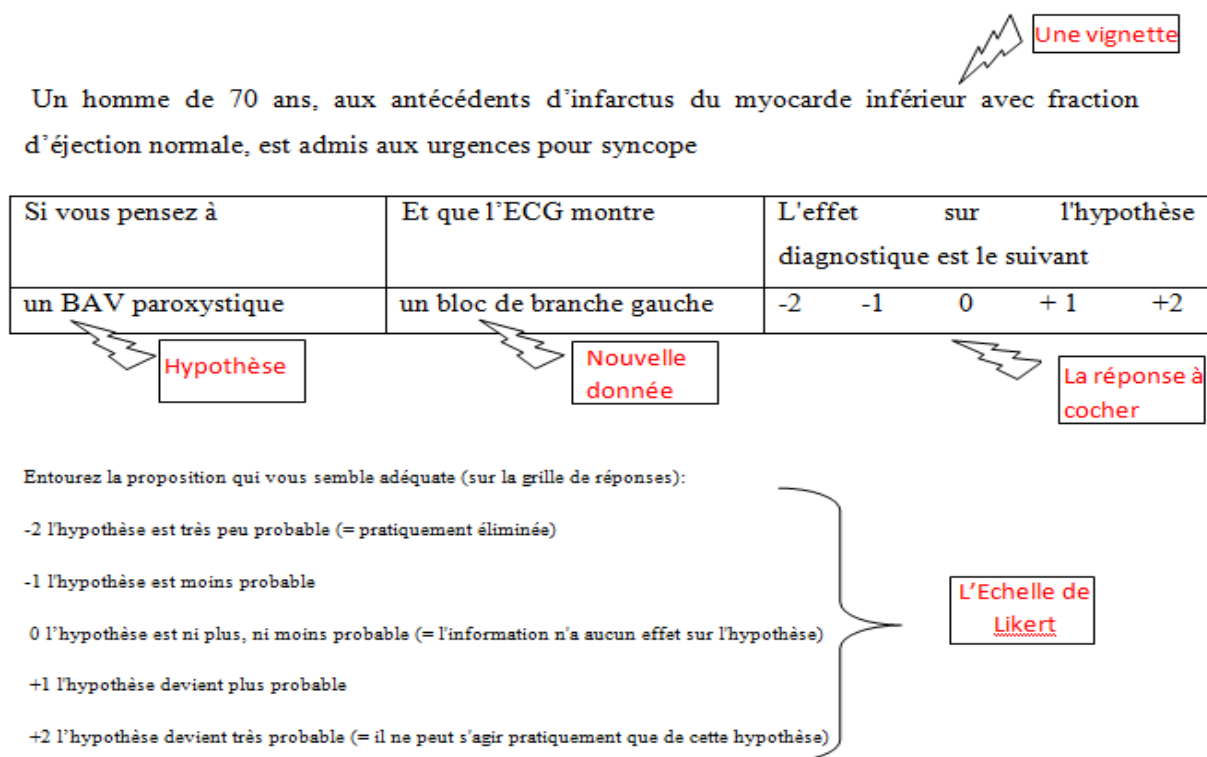


Figure n°1 : Modèle utilisé pour l'explication du principe du test de concordance de script

2. Administration du test

Une fois le test construit par un cardiologue (maître de conférences agrégé), il a été soumis pour avis à des cardiologues expérimentés afin de recueillir leurs appréciations sur la pertinence des questions ainsi que sa faisabilité. Aucun item n'a été éliminé pour non conformité au programme d'enseignement ou encore lorsqu'il a été estimé que la situation présentée ne revêtait pas suffisamment de caractère d'incertitude. Au final, 60 questions (20 vignettes avec 3 items chacune) ont été retenues.

Le TCS a été soumis à un groupe de dix cardiologues (les experts) constituant le panel de référence. Ce groupe d'experts en cardiologie exerçant au CHU Hédi Chaker de Sfax était composé de 2 professeurs, 2 Maîtres de Conférences Agrégés, 4 Assistants Hospitalo-Universitaires et 2 médecins de santé publique spécialistes. Leurs réponses ont permis l'élaboration de l'échelle de notation.

3. Système de notation du TCS

Pour chaque item, la réponse donnait droit

à un crédit qui correspondait au nombre d'experts qui l'avaient choisie (Tableau I).

Le score total pour le test était la somme de tous les crédits obtenus à chaque item. Ce score total était ensuite transformé par une simple règle de trois pour obtenir une note exprimée sur 100.

Le TCS a été secondairement soumis à un groupe de 44 étudiants en médecine de la 6^{ème} année des études médicales de la faculté de Médecine de Sfax (équivalent de la DCEM 4).

Ont été retenus pour ce test uniquement les étudiants ayant effectué l'ensemble de leur cursus médical dans la faculté de médecine de Sfax et n'ayant pas redoublé entre la quatrième et la sixième année de leur cursus universitaire. Ceci a permis de ne retenir que les étudiants qui ont eu les mêmes cours de cardiologie en 3^{ème} année au cours de l'année universitaire 2013–2014. Le principe du test était expliqué avant son administration aux étudiants à l'occasion d'une réunion tenue à cet effet dans le service. Par la suite, les étudiants devaient réaliser le test en 60 minutes (1 minute/item). Une grille de réponse individuelle a été remise à chaque étudiant.

Vignette 1: Un homme de 70 ans, aux antécédents d'infarctus du myocarde inférieur avec fraction d'éjection normale, est admis aux urgences pour syncope

Si vous pensez à	Et que l'ECG montre	L'effet sur l'hypothèse diagnostique est le suivant				
un BAV paroxystique	un bloc de branche gauche	-2	-1	0	+1	+2
une hypotension orthostatique	un BAV du premier degré	-2	-1	0	+1	+2
une dysfonction sinusale	une fibrillation atriale	-2	-1	0	+1	+2

Entourez la proposition qui vous semble adéquate (sur la grille de réponses):

-2 l'hypothèse est très peu probable (= pratiquement éliminée)

-1 l'hypothèse est moins probable

0 l'hypothèse est ni plus, ni moins probable (= l'information n'a aucun effet sur l'hypothèse)

+1 l'hypothèse devient plus probable

Figure n°2: exemple de TCS à visée diagnostique

Vignette 2: Un homme de 65 ans, admis en USIC pour SCAS+

Si vous pensez à	Et que vous apprenez que	Cette nouvelle donnée rend le traitement proposé:				
le traiter par une angioplastie coronaire primaire	la douleur datait de 30 H	-2	-1	0	+1	+2
Le traiter par une thrombolyse	Le patient est sous traitement anticoagulant par voir orale (AVK)	-2	-1	0	+1	+2
Lui prescrire 600 mg de clopidogrel	Il est hospitalisé dans un centre ou l'équipe de la salle de KT est disponible 24h/24h	-2	-1	0	+1	+2

Cette nouvelle donnée rend le traitement proposé:

-2: Contre-indiqué

-1: Peu utile peu pertinent

0: Ni plus ni moins utile

+1: Plus utile plus pertinent

+2: Indispensable

Figure n°3: exemple de TCS à visée thérapeutique

Vignette 17 : patiente âgée de 32 ans consulte pour céphalées en rapport avec une hypertension artérielle.

Hypothèse : Si vous pensez à faire	Nouvelle information: Et que vous trouvez	L'effet sur l'hypothèse initiale est le suivant				
Angio-scanner des artères rénales	La kaliémie normale	-2	-1	0	+1	+2
Un angio-scanner aortique	Les Pouls périphériques présents et symétriques	-2	-1	0	+1	+2
Un Dosage des dérivés méthoxylés urinaires	Des épisodes de palpitations+bouffées de chaleur	-2	-1	0	+1	+2

-2: Absolument contre-indiqué

-1: Peu utile

0: Non pertinent dans cette situation

+1: Utile et souhaitable

+2: indispensable

Figure n°4 : Exemple de TCS à visée investigation

Tableau I : Etablissement des scores en fonction des réponses des experts

	-2	-1	0	+1	+2
Réponse des 10 Experts	0	0	1	7	2
Score des étudiants en Fonction des Réponses des experts	0	0	1/7 x 1 expert ayant choisi cette réponse	Réponse choisi par 7 experts : 7/7	1/7x 2 experts ayant choisi cette réponse
Note de l'étudiant pour l'item	0	0	0.14	1	0.28

4. Analyse statistique

Les données ont été saisies et analysées au moyen du logiciel SPSS version 18.0. Nous avons mené une étude descriptive (calcul des fréquences simples, des fréquences relatives pour les variables qualitatives, des moyennes, et des écarts-types pour les variables quantitatives). Une étude analytique consistant en une comparaison des moyennes, comparaison de pourcentages (séries indépendantes ont été effectuées par le test du chi-deux de Pearson, et en cas de non-validité par le test exact bilatéral de Fisher. Pour les séries appariées ont été effectuées par le test de Mac Nemar, et en cas de non-validité de ce test, par l'utilisation des propriétés de la loi binomiale).

Une Etude de corrélation a été réalisée en utilisant le test de Spearman après avoir vérifié la distribution gaussienne des différents paramètres. En cas de distribution non gaussienne, nous avons utilisé les tests non paramétriques. La fiabilité du test a été étudiée en évaluant le coefficient de fiabilité α de Cronbach. Dans tous les tests statistiques, le seuil de signification a été fixé à 0,05.

RESULTATS

Quarante quatre étudiants de la 6^{ème} année de Médecine de la faculté de Médecine de Sfax ont rempli les conditions d'inclusion dans l'étude. Parmi ces étudiants, un premier groupe de 22 (groupe 1) a effectué un stage de 6^{ème} année de

Médecine d'une durée de 16 semaines dans le service de Cardiologie du centre hospitalier universitaire Hédi Chaker de Sfax durant l'année universitaire 2015 – 2016. Un deuxième groupe (groupe 2) de 22 a eu d'autres stages de Médecine.

1. Comparaison des notes de TCS entre les 2 groupes:

Les connaissances en cardiologie des deux groupes d'étudiants étaient différentes en fin de 3^{ème} année de Médecine. En effet, Une différence statistiquement significative entre les moyennes des notes obtenues à l'examen de Cardiologie en 3^{ème} année pour les deux groupes d'étudiants a été retrouvée. De même, le rang moyen à l'examen de cardiologie de 3^{ème} année était 31,95±8,26 chez les étudiants ayant effectué par la suite un stage de cardiologie, alors que le rang moyen chez ceux qui n'ont pas effectué par la suite de stage de cardiologie était de 13,04±9,07.

Les deux groupes d'étudiants n'étaient pas comparables en termes d'acquisition des connaissances théoriques en cardiologie à la fin des cours de 3^{ème} année, avant les stages de cardiologie de la 6^{ème} année.

Notre test TCS était fiable avec un coefficient α de Cronbach égal à 0,76. Les résultats du TCS réalisé au cours de la 6^{ème} année de Médecine chez ces mêmes étudiants donnaient une note moyenne de 54,97±10,79 sur 100. La moyenne des notes du TCS des étudiants ayant effectué un stage de cardiologie en 6^{ème} année de Médecine (groupe1) n'était pas statistiquement différente de celle des étudiants n'ayant pas effectué de stage de cardiologie (groupe 2) : 53,89±13,1 vs 56,04±7,92 (p=0,51). Aucune différence n'a été également retrouvée si nous prenons en considération séparément les 13 items évaluant la capacité du diagnostic, les 4 items évaluant la capacité de traitement et les 3 items évaluant les capacités d'investigation (Tableau II).

Tableau II : Comparaison des résultats du test TCS des étudiants des deux groupes

Paramètres	Groupe 1	Groupe 2	P
Note en 3 ^{ème}	65,45±9,80	80,56±6,37	<0,001
Classement en 3 ^{ème}	31,95±8,26	13,04±9,07	<0,001
Note TCS global	53,89±13,1	56,04±7,92	0,51
Note TCS diagnostique	20,23±5,47	20,72±3,54	0,72
Note TCS Thérapeutique	6,69±2,03	6,78±1,57	0,87
Note TCS investigation	5,41±1,51	6,11±1,03	0,08

2. Classement des étudiants après l'évaluation par TCS:

Seize des 22 étudiants (72,7%) ayant effectué un stage de cardiologie (groupe 1) ont amélioré leur classement final évalué par le TCS par rapport à leur classement en 3^{ème} année. Au contraire, seuls 3 des 22 étudiants (13,6%) n'ayant pas effectué de stage de cardiologie (groupe 2) ont amélioré leur classement.

Le classement moyen au TCS de cardiologie des étudiants ayant effectué un stage de cardiologie (groupe 1) est de 23,63±15,17 contre 31,95±8,26 lorsqu'ils étaient en 3^{ème} année alors que le classement des étudiants du groupe 2 est passé de 13,04±9,07 en 3^{ème} année à 21,36±10,23 en 6^{ème} année (TCS) (Tableau III).

Tableau III : Classement des étudiants en 3^{ème} année et après évaluation par TCS

Paramètres	Groupe 1	Groupe 2	P
Classement en 3 ^{ème}	31,95±8,26	13,04±9,07	<0,001
Classement en 6 ^{ème} (TCS)	23,63±15,17	21,36±10,23	0,56
Amélioration du classement (entre 3 ^{ème} et la 6 ^{ème})	16 (72,7%)	3 (13,6%)	P<0,001

Autrement dit, le classement des étudiants du groupe 1 a connu une amélioration de $8,31 \pm 13,3$ places alors que le classement des étudiants du groupe 2 a connu une régression de $8,31 \pm 9,97$ places (ANOVA P-value $p < 0,001$).

3. Etude des corrélations entre les notes de 3^{ème} année et celles du TCS:

Pour tout le groupe, il existait une corrélation entre les notes de cardiologie (3^{ème} année) et celles de TCS (6^{ème} année) ($p=0,031$, $r=0,325$). Pour le groupe ayant effectué un stage, aucune corrélation n'a été retrouvée entre la note finale de TCS et la note de l'examen de cardiologie de 3^{ème} année ($p=0,163$). Par contre pour le groupe d'étudiants qui n'ont pas effectué de stage, nous avons trouvé que les notes de cardiologie et les notes de TCS étaient corrélées ($p=0,02$, $r=0,469$).

DISCUSSION

Le TCS est un outil d'évaluation basé sur la théorie cognitive des scripts [4]. Cette théorie suggère que lorsque les médecins sont soumis à un problème clinique, ils mobilisent des réseaux de connaissances préétablis (les scripts) stockés dans la mémoire à long terme, pour comprendre la situation et agir en fonction des buts diagnostic, d'investigation ou de traitement [2]. Ces scripts sont utilisés quotidiennement en pratique clinique. Ils se mettent en place lorsque les étudiants se retrouvent confrontés à leur première situation clinique, puis se développent et s'affinent tout au long de leur carrière. Tandis que les étudiants gagnent en expérience et en connaissances médicales spécialisées, ils développent des scripts mieux organisés et deviennent capables d'en manipuler plusieurs en même temps [4]. Donc le TCS ne peut évaluer que des étudiants possédant un corpus de connaissances suffisant et « digéré » et ne peut donc être proposé en début de formation [5].

Dans notre étude, le TCS a été administré à des étudiants en fin de deuxième cycle des études médicales (6^{ème} année de médecine) qui ont tous réussi leurs examens de cardiologie en 3^{ème} année de Médecine.

Notre étude a permis d'infirmer l'hypothèse selon laquelle les étudiants qui ont les meilleures connaissances sont ceux qui auront les meilleures compétences. En effet, les étudiants du groupe 2 avaient une moyenne de note de 3^{ème} année dépassant de façon significative celle du groupe 1.

Cette différence s'est estompée après la pratique de stage. Ce résultat pourrait s'expliquer par la théorie des scripts selon laquelle, en situation clinique, les médecins mobilisent leurs réseaux de connaissances préétablies ou scripts [6]. Ils n'accorderont pas tous la même importance à un signe clinique ou à un examen complémentaire pour aboutir à un même diagnostic. L'acquisition d'une telle compétence se fait en grande partie en terrain de stage. Notre étude a objectivé une amélioration du classement TCS des étudiants ayant bénéficié d'un stage de cardiologie par rapport à ceux qui n'ont pas effectué un tel stage.

Il a été mentionné que les étudiants adaptent leur mode d'apprentissage en fonction des dispositifs d'évaluation utilisés. Il est alors permis de formuler l'hypothèse que la diffusion de l'utilisation du TCS en cardiologie, qui a pour objectif d'évaluer le raisonnement clinique, puisse permettre d'optimiser les activités d'apprentissage et ramener nos étudiants « au lit du patient » [7]. De même, la réalisation d'un TCS spécifique dans le champ de la cardiologie en début et en fin de stage pourrait constituer un excellent moyen d'évaluer la progression et donc l'utilité réelle pour l'étudiant de son stage.

1. Qualités psychométriques du TCS bien établies:

Le TCS est un outil d'évaluation du raisonnement clinique possédant de bonnes qualités psychométriques (validité, fiabilité interne, faisabilité). Ainsi, le TCS permet de discriminer les candidats dans différents domaines tels que la médecine générale [8], la médecine d'urgence [9], la chirurgie [10], les spécialités médicales [11] et l'anesthésie-réanimation [12]. Notre TCS s'applique à la cardiologie.

Le TCS permet l'évaluation de grands groupes d'étudiants de façon standardisée avec une correction objective et automatisée [13]. La majorité des études étaient menées en Amérique du Nord et en Europe. Néanmoins, on a pu observer que le TCS tend à être utilisé partout dans le monde, comme au Liban [14], Iran [15], en Thaïlande [16], au Maroc [17] et au Brésil [18]. En Tunisie, à notre connaissance, c'est la première étude pilote s'intéressant au TCS.

Le TCS est très souvent utilisé sous forme papier/crayon mais il se prête bien à l'administration en ligne où il offre de grands avantages tant sur le plan logistique que sur l'ouverture des champs d'application.

L'utilisation de TCS en ligne fournit un rapport coût-efficacité intéressant permettant d'inclure les candidats sans les difficultés logistiques habituellement rencontrées par la méthode papier/crayon. Ainsi il permet aux candidats de passer le test en dehors de la faculté, et d'évaluer de grands groupes d'étudiants [11]. On peut intégrer au TCS en ligne l'utilisation du multimédia (image, vidéo, audio) au sein des scénarios, ce qui rend l'évaluation plus proche de la pratique clinique réelle [19]. Notre TCS avait un format papier comportant 8 pages. Nous avons constaté une réticence de la part des étudiants pour le remplir. La forme électronique aurait pu les encourager.

Nous avons constaté de nombreuses publications sur les TCS. Récemment, Lubarsky et al. [20] ont travaillé sur la validité des TCS dans une revue de la littérature. Les preuves accumulées concernant la validité de construit et la cohérence interne sont fortes. Dory et al. [8], dans une revue plus récente proposent une synthèse des recommandations à la construction et mise en place des TCS. Selon Gagnon et al [21], 25 vignettes scénarios contenant chacune 3 à 4 questions permettraient d'atteindre une fiabilité interne du test entre 0,75 et 0,86. Une centaine d'items est en général formulée initialement. Cela permet d'en conserver suffisamment (environ 75), après relecture et élimination des items peu performants. Cette analyse permet de produire un test plus court et plus fiable. Chaque item doit être indépendant des autres. Il faut environ une heure pour un examen contenant 60 à 90 items.

La fiabilité interne du test est évaluée par le coefficient alpha de Cronbach. Il doit être supérieur ou égal à 0,8 pour fournir une bonne fiabilité interne à l'étude, nécessaire pour les examens à forts enjeux. Notre test comportait 60 items. Le coefficient alpha de Cronbach de notre TCS était de 0,76. Ce qui était en faveur de la fiabilité de notre test.

Le panel d'experts doit regrouper 10 à 20 experts pour obtenir une cohérence interne satisfaisante. La fiabilité du test augmente avec la taille du panel. Un panel incluant plus de 20 experts n'apporte qu'une légère augmentation de la fiabilité en comparaison de la difficulté à recruter un tel panel [19]. Une telle approche se discute pour les examens à forts enjeux dans lesquels la fiabilité est un élément clé. Le panel doit avoir une bonne expérience clinique globale du domaine testé [22]. Il peut être intéressant de choisir des experts

spécialisés au sein du panel pour améliorer la validité du test, ce d'autant que le domaine évalué est large (chirurgie générale, médecine interne) [20]. Notre panel comportait 10 experts spécialistes en cardiologie.

Dans notre étude le coefficient α de Cronbach est acceptable. Tout d'abord, notre panel d'experts était de 10, ce qui est conforme aux recommandations [23], mais peut-être que 15 à 20 experts auraient permis d'améliorer la validité interne de notre étude [24]. De même, le rédacteur du TCS a fait partie du panel d'experts. Nous avons bien respecté la règle des 3 items en moyenne par vignettes et nous avons au total 60 items dans nos TCS. Dans de nombreux articles, les auteurs ont utilisé des TCS comprenant 90 items, ce qui n'aurait pas été possible vu la difficulté d'élaborer ces items et surtout qu'il s'agit de notre première expérience. Toutefois, nous avons retrouvé certains articles où malgré un nombre d'items de TCS conséquents ($n = 120$) le coefficient α de Cronbach n'était que de 0,69 et dans un autre travail, un coefficient α de Cronbach égal à 0,83 avec seulement 26 items [11].

Plusieurs études ayant testé l'effet de différentes méthodes de notation, donnent des résultats contradictoires, et la méthode de notation actuelle s'est avérée satisfaisante dans toutes ces études en termes de qualités psychométriques des scores obtenus et d'adéquation avec les bases théoriques du TCS [8].

2. Comparaison du TCS aux outils standards d'évaluation:

Le TCS a souvent été comparé aux QCM. Selon les études, la corrélation entre ces deux outils est faible [13] ou absente [22]. TCS et QCM varient dans le même sens d'une manière générale mais testent deux aspects différents et complémentaires de la compétence clinique. Cette faible corrélation souligne le rôle des connaissances médicales de base, en plus de l'expérience clinique, dans le développement et l'activation des scripts pour guider le processus de raisonnement clinique. Dans notre étude, les scores obtenus aux TCS étaient significativement corrélés à ceux obtenus en 3^{ème} année par les étudiants. Cependant ces scores étaient significativement différents. Pour le groupe ayant effectué un stage, aucune corrélation n'a été retrouvée entre la note finale de TCS et la note de l'examen de cardiologie de 3^{ème} année. Par contre pour le groupe d'étudiants qui n'ont pas effectué de stage, nous avons trouvé que les notes de

cardiologie et la note de TCS étaient corrélées. L'explication la plus probable est que les TCS et les moyens d'évaluation classique en 3^{ème} année de la faculté de médecine de Sfax (QROC, QCM, cas cliniques) ne testent pas les mêmes connaissances. Les TCS testent le raisonnement clinique dans un contexte d'incertitude et les QROC testent des connaissances factuelles en situation clinique.

Ce résultat suggère que les TCS et ces moyens classiques sont complémentaires et que l'implémentation des TCS dans les examens des études médicales en plus des méthodes classiques permettrait de mieux classer les étudiants.

Dans notre étude, plusieurs explications sont avancées pour justifier certains mauvais scores TCS en dépit de la pratique du stage. La présence des étudiants lors des séances d'entraînement aux TCS n'a pas été vérifiée. Certains étudiants n'ont peut-être pas assisté à ces séances, expliquant leur mauvaise note au TCS. Enfin, les étudiants avec de très mauvais scores aux TCS n'avaient pas répondu à tous les TCS, probablement par faute de temps car les items sans réponse sont les 10 derniers. Il aurait été intéressant de recueillir l'opinion et le ressenti des étudiants par un questionnaire ou mieux, des entretiens semi-directifs sur un groupe d'étudiants tiré au sort surtout que nous avons constaté une réticence de la part des étudiants pour adhérer à l'évaluation par le TCS.

CONCLUSION

Nous pouvons formuler l'hypothèse que la diffusion de l'utilisation du TCS en cardiologie et dans d'autres spécialités, dans le but d'évaluer le raisonnement clinique, puisse permettre d'optimiser les activités d'apprentissage et de ramener nos étudiants « au lit du patient ». De même, la réalisation d'un TCS spécifique dans le champ de la cardiologie en début et en fin de stage pourrait constituer un excellent moyen d'évaluer la progression et donc l'utilité réelle pour l'étudiant de son stage. La performance au TCS est liée à l'expérience clinique et donc aux stages pratiques. On peut espérer que l'emploi des TCS renforce l'intérêt des étudiants pour leur stage, leur participation aux activités pratiques, et dans un deuxième temps l'investissement des médecins dans la formation au lit du malade.

Le TCS, moyen valide pour évaluer le raisonnement clinique et la capacité de synthèse des étudiants devrait faire partie, dans un futur proche, des méthodes d'évaluation au sein de notre faculté. Il est intéressant de construire une banque

de TCS. Une collaboration étroite entre l'unité de standardisation au sein de la direction de Pédagogie et des Technologies Educatives de la faculté de médecine de Sfax et les enseignants hospitalo-universitaires est indispensable.

RÉFÉRENCES

- [1] Sibert L, Grand'Maison P, Charlin B, Grise P. Evaluation of clinical competence in urology : innovative approach based on performance observation. *Prog Urol*.1997 ;7 :581-589.
- [2] 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: 189-95.
- [3] Dory V, Gagnon R, Vanpee D, Charlin B. How to construct and implement script concordance tests: insights from a systematic review. *Med Educ* 2012, 46:552-63.
- [4] Charlin B, Tardif J, Boshuizen HP. Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. *Acad Med* 2000 ;75:182-90.
- [5] Lebeau J, Pagonis D. Le test de concordance de script: comment évaluer le raisonnement médical en situation d'incertitude? *Rev Stomatol Chir Maxillofac*. 2006;107:327-9.
- [6] Charlin B, Boshuizen HP, Custers EJ, Feltovich PJ. Scripts and clinical reasoning. *Med Educ*. 2007;41:1178-84.
- [7] Gibot S, Bollaert PE. Le test de concordance de script comme outil d'évaluation formative en réanimation médicale. *Pédagogie Médicale* 2008 ;9:7-18.
- [8] Dory V, Charlin B, Vanpee D, Gagnon R. Multifaceted assessment in a family medicine clerkship: a pilot study. *Fam Med*. 2014 ;46:755-60.
- [9] Boulouffe C, Doucet B, Muschart X, Charlin B, Vanpee D. Assessing clinical reasoning using a script concordance test with electrocardiogram in an emergency medicine clerkship rotation. *Emerg Med J*. 2014 ;31:313-6.
- [10] 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.
- [11] Marie I, Sibert L, Roussel F, Hellot MF, Lechevallier J, Weber J. The script concordance test: a new evaluation method of both clinical reasoning and skills in internal medicine. *Rev Med Interne*. 2005 ;26:501-7.
- [12] Ducos G, Lejus C, Sztark F, Nathan N, Fourcade O, Tack I, et al. The Script Concordance Test in anesthesiology: Validation of a new tool for assessing clinical reasoning. *Anaesth Crit Care Pain Med*. 2015;34:11-5.
- [13] Humbert AJ, Johnson MT, Miech E, Friedberg F, Grackin JA, Seidman PA. Assessment of clinical reasoning: A Script Concordance test designed for pre-clinical medical students. *Med Teach*. 2011;33:472-7.
- [14] Haddad F, Gaspard D, Nasr M, Abou Jaoude S, Nembr E. Etude exploratoire de l'utilisation du test de concordance de script pour l'évaluation des apprentissages de résidents de médecine interne au Liban. *Pédagogie médicale*. 2008 ;9,135-140.
- [15] Amini M, Moghadami M, Kojuri J, Abbasi H, Abadi AA, Molaei 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; 4:418.
- [16] Khonputsa P, Besinque K, Fisher D, Gong WC. Use of script concordance test to assess pharmaceutical diabetic care: a pilot study in Thailand. *Med Teach* 2006; 28:570-3.

PERTINENCE DU TEST DE CONCORDANCE DE SCRIPT DANS L'EVALUATION DES COMPETENCES

- [17] Houssaini TS, Bono W, Tachfouti N, Maillard D. Pertinence d'un test de concordance de script dans l'évaluation des compétences en Néphrologie des étudiants du deuxième cycle de la faculté de Médecine de Fès. Les annales Méd Thérapeutique 2009; 1 :4-10.
- [18] Roberti A, Roberti Mdo R, Pereira ER, Costa NM. Script concordance test in medical schools in Brazil: possibilities and limitations. Sao Paulo Med J. 2016;134:116-20
- [19] Ruiz JG, Tunuguntla R, Charlin B, Ouslander JG, Symes SN, Gagnon R, et al. The script concordance test as a measure of clinical reasoning skills in geriatric urinary incontinence. J Am Geriatr Soc. 2010;58:2178-84.
- [20] Lubarsky S, Dory V, Duggan P, Gagnon R, Charlin B. Script concordance testing: From theory to practice: AMEE Guide No. 75. Medical teacher 2013 ; 35:184-93
- [21] Gagnon R, Charlin B, Coletti M, Sauvé 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.
- [22] Fournier JP, Demeester A, Charlin B. Script concordance tests: guidelines for construction. BMC Med Inform Decis Mak. 2008;8:18.
- [23] Petrucci AM, Nough T, Boutros M, Gagnon R, Meterissian SH. Assessing clinical judgment using the Script Concordance test: the importance of using specialty-specific experts to develop the scoring key. Am J Surg. 2013;205:137-40.
- [24] Jouquan J. L'évaluation des apprentissages des étudiants en formation médicale initiale. Pédagogie Médicale 2002;31:38-52.