

Questions de planification clinique : un nouvel outil pour évaluer la capacité des étudiants en médecine à identifier les éléments-clés discriminants d'un diagnostic différentiel

Clinical Planning Questions: developing a new tool for assessing medical-student skills in identifying key elements of differential diagnoses

Jean-Sébastien RENAUD¹, François RATTÉ¹, Julie F. THÉRIAULT¹, Anne Melissa ROY¹, Luc CÔTÉ¹

¹ Faculté de médecine, Université Laval, Québec, Canada

Manuscrit soumis le 27 mai 2015 ; commentaires éditoriaux formulés aux auteurs le 8 mars et le 5 juillet 2016 ; accepté pour publication le 6 juillet 2016.

Mots-clés

Raisonnement clinique ; évaluation ; formation médicale initiale

Résumé – Contexte : Les outils actuels d'évaluation du raisonnement clinique n'évaluent pas spécifiquement une habileté majeure pour le clinicien novice, soit l'identification des éléments-clés spécifiques des diagnostics différentiels. Nous avons répondu à ce besoin en développant une nouvelle modalité d'évaluation : les questions de planification clinique (QPC). **Objectif :** Présenter cette modalité d'évaluation et en documenter les qualités psychométriques. **Méthodes :** À l'automne 2012, une QPC a été intégrée à l'examen écrit du cours Démarche clinique 1 (DC1) ($n = 222$) et une autre à celui du cours Démarche clinique 3 (DC3) ($n = 216$). Nous avons comparé les qualités psychométriques (difficulté, discrimination, validité prédictive) des QPC à celles des autres modalités d'évaluation écrites du raisonnement clinique utilisées dans ces deux examens. Les résultats à l'examen pratique à la fin de DC1 et de DC3 ont servi de critère pour évaluer la validité prédictive. **Résultats :** Dans les deux examens, la QPC se situe à un niveau de difficulté modéré selon le barème utilisé. L'indice de discrimination de la QPC est de 0,39 dans l'examen de DC1 et de 0,15 dans celui de DC3. La QPC n'est pas corrélée de manière significative aux résultats aux examens pratiques de DC1 et de DC3, à l'instar de la grande majorité des autres types de questions. **Conclusion :** Les QPC apparaissent comme une nouvelle modalité d'évaluation écrite du raisonnement clinique qui est facile à concevoir, complémentaire aux autres modalités et dont les qualités psychométriques sont similaires ou supérieures.

Keywords

Clinical reasoning;
assessment;
undergraduate
medical training

Abstract – Background: Current tools in assessing clinical reasoning do not specifically assess a major skill for novice clinicians: identifying key elements of differential diagnoses. We responded to this need by developing a new assessment method: Clinical Planning Questions (CPQ). **Purpose:** To introduce the assessment method and establish its psychometric properties. **Methods:** In the fall of 2012, CPQs were included in the written exams of two courses: Démarche Clinique 1 (DC1) ($n = 222$) and Démarche clinique 3 (DC3) ($n = 216$). We compared the psychometric properties (difficulty, discrimination, predictive validity) of CPQs to those of other written clinical reasoning assessment methods used in the two exams. The results for the practical exam at the end of DC1 and DC3 were used as a criterion for assessing predictive validity. **Results:** In both exams, CPQs had a moderate level of difficulty. The discrimination index of the CPQs was 0.39 in the DC1 exam and 0.15 in the DC3 exam. The CPQs were not significantly correlated with the results in the DC1 and DC3 practical exams, like in most other types of questions. **Conclusion:** CPQs are seen as a new written clinical reasoning assessment method that is easy to design, complements other methods and has comparable psychometric properties.

Introduction

Problématique

Le raisonnement clinique fait référence aux processus de pensée permettant au clinicien de traiter les informations qu'il a recueillies dans un contexte spécifique de résolution de problèmes et ce, dans le but de prendre des décisions appropriées. Son exercice requiert une base de connaissances et repose sur divers types de processus de raisonnement complexes, dont les processus cognitifs analytiques et non analytiques^[1].

Récemment, à partir d'une recherche action participative, Charlin *et al.*^[2] ont proposé une modélisation des processus cognitifs du raisonnement clinique en médecine. Il apparaît que les cliniciens mobilisent et utilisent des connaissances spécifiques au problème à résoudre, ces connaissances pouvant concerner des champs de nature différente (par exemple, biologique, psychologique, sociale). Avec l'expérience, les cliniciens développent un réseau de connaissances bien organisées reposant non seulement sur les liens qu'ils font entre les éléments discriminants qu'ils associent à une maladie, mais aussi sur les liens avec les représentations de maladies qu'ils ont encodées en mémoire (*illness scripts*)^[3-5]. De plus, dans leur pratique, les cliniciens sont confrontés à des présentations diverses d'un même diagnostic, ce qui

leur permet de raffiner leurs schémas cliniques et donc de faire des rappels plus rapides et efficaces. À la base de ce raisonnement, le clinicien doit donc bien se représenter le problème, d'où l'importance de connaissances organisées en réseau^[6] et la nécessité d'interpréter les données cliniques en leur donnant une valeur (un sens) en fonction du contexte dans lequel il se trouve. La démarche de raisonnement clinique est donc spécifique^[7].

Cette capacité à reconnaître les éléments discriminants d'un problème et à leur accorder une valeur est une étape essentielle de l'apprentissage du raisonnement clinique et de la prise de décision clinique. C'est ce qui a conduit au développement du concept d'éléments-clés spécifiques^[8,9] et au développement de nouvelles modalités d'évaluation sur la prise de décisions cliniques par le Conseil Médical Canadien dans les années 90^[10,11].

La capacité des étudiants à identifier les éléments-clés discriminants d'un diagnostic différentiel (par exemple, symptômes et signes) est primordiale. En effet, les étudiants en mesure d'identifier ces éléments discriminants sauront plus facilement recueillir les informations pertinentes en fonction des hypothèses diagnostiques soulevées^[1]. Il s'agit donc d'une étape cruciale à maîtriser dans l'apprentissage du raisonnement clinique, d'où la nécessité non seulement de l'enseigner mais de l'évaluer pendant la formation médicale.

Or, les outils actuels d'évaluation du raisonnement clinique^[12, 13] n'évaluent pas spécifiquement la capacité des étudiants en médecine à identifier eux-mêmes, sans choix de réponses, les données cliniques pertinentes permettant de discriminer entre les hypothèses diagnostiques probables. Pour pallier ce manque, nous avons développé, au cours du programme de doctorat en médecine à l'Université Laval (Québec, Canada) une nouvelle modalité d'évaluation reposant sur l'utilisation de questions de planification clinique (QPC). Notre expérience avec les QPC a jusqu'à maintenant été positive. Elles sont relativement rapides à développer et à corriger. Elles sont aussi bien reçues par les étudiants. Par contre, nous n'avons aucune donnée psychométrique pour argumenter leur utilisation.

Les questions de planification clinique

C'est le besoin de rendre explicite le processus cognitif du raisonnement clinique qui a principalement motivé le développement des QPC. Les auteurs considèrent comme important que les étudiants développent leur capacité à se faire une représentation clinique juste et réaliste, tout en comparant les éléments discriminants qui permettront de justifier chacune des hypothèses diagnostiques. L'idée de procéder par comparaison d'un diagnostic avec un autre ressemble à ce que le clinicien fait quotidiennement. Par ailleurs, comme l'évaluation est un puissant moteur de l'apprentissage^[14], notamment l'évaluation sommative^[15], nous avons opté pour l'élaboration d'un moyen d'évaluation afin d'atteindre notre objectif et avons d'abord testé les QPC en contexte sommatif, bien qu'il serait approprié de les utiliser en contexte formatif.

Les QPC sont utilisées dans les cours portant sur l'apprentissage du raisonnement clinique (décrits dans la section suivante). Elles s'adressent à des étudiants au niveau préclinique, c'est-à-dire qui débutent leur formation en médecine et qui n'ont pas encore fait l'expérience de la pratique clinique. Il s'agit donc d'étudiants dits novices. Elles ne visent pas à juger la qualité de l'ensemble du raisonnement clinique. Elles ciblent plutôt une partie essentielle de ce dernier, soit l'évaluation de la capacité des étudiants à identifier

les éléments-clés discriminants d'un diagnostic différentiel. Les hypothèses diagnostiques à comparer et contraster sont d'ailleurs fournies à l'étudiant de manière à circonscrire uniquement leur capacité à identifier les éléments-clés qui permettront de les différencier lors de l'histoire de cas ou de l'examen physique. Cette « mise en lumière » d'une partie du raisonnement clinique rend la tâche certes moins authentique et moins appropriée pour l'évaluation d'étudiants plus avancés dans leur formation et disposant d'une plus grande expérience clinique. Pour ces derniers, une évaluation holistique est probablement préférable^[16]. En revanche, elle évite probablement la surcharge cognitive chez les novices et aide à développer la capacité à comparer et contraster les diagnostics différentiels entre eux, ce qui favorise une récolte de données cliniques pertinentes et discriminantes. Travailler cette capacité chez l'étudiant est en accord avec les principes de psychologie cognitive qui favorisent l'apprentissage du raisonnement clinique^[1, 17]. De plus, même si cette tâche a un caractère plus artificiel, il demeure qu'il est nécessaire qu'elle soit maîtrisée tôt par l'étudiant afin qu'il puisse éventuellement démontrer un raisonnement clinique efficace en situation authentique. En d'autres termes, c'est un savoir-faire essentiel à maîtriser par l'étudiant, car il devra le mobiliser pour résoudre des problèmes cliniques en situation authentique plus tard dans sa formation et sa pratique.

Les QPC sont des questions à réponses courtes construites à partir d'une mise en situation clinique succincte se limitant à la présentation de données cliniques de base telles que l'âge, le sexe et la raison de consultation du patient. Trois hypothèses diagnostiques probables sont proposées à l'étudiant. Ce dernier doit choisir trois éléments-clés spécifiques de l'anamnèse et de l'examen physique pour chaque hypothèse diagnostique proposée et inscrire ses réponses dans un tableau comparatif. Aucun choix de réponse n'est fourni pour les éléments-clés à identifier. Les QPC exigent de l'étudiant qu'il se questionne sur les données qui lui seront essentielles dans la planification de son anamnèse et de son examen physique afin de comparer et contraster les trois hypothèses diagnostiques suggérées dans le contexte clinique proposé. C'est la raison pour laquelle cette

Monsieur Verreault est âgé de 84 ans. Il se présente à l'urgence, car il est dyspnéique depuis ce matin. Avant de le rencontrer, vous organisez votre pensée.

À partir des trois hypothèses diagnostiques proposées, complétez le tableau en identifiant les éléments du questionnaire et de l'examen physique qui sont les plus discriminants (spécifiques) pour distinguer chacune des trois conditions pathologiques. Si un élément est commun aux trois diagnostics (ex. : dyspnée), ce n'est pas un élément discriminant. Inscrivez une seule réponse par case.

Dyspnée aigüe			
Diagnostic	Insuffisance cardiaque décompensée	Pneumonie	Embolie pulmonaire
Éléments discriminants à l'histoire :	1.	1.	1.
	2.	2.	2.
	3.	3.	3.
Éléments discriminants à l'examen physique :	1.	1.	1.
	2.	2.	2.
	3.	3.	3.

Fig. 1. Exemple de Question de planification clinique.

tâche se veut stimulante au plan cognitif (*probing-challenging task*). Un exemple est proposé sur la figure 1

Les scénarios cliniques sont les plus brefs possibles et correspondent à des raisons de consultation fréquentes en médecine, dont le diagnostic différentiel est cohérent par rapport à la réalité clinique. Les raisons de consultation sont inspirées de situations cliniques réelles rencontrées en contexte d'urgence, d'hospitalisation ou de clinique externe. Les scénarios cliniques doivent également être en lien avec les objectifs spécifiques du cours. Par exemple, la raison de consultation pourrait être la toux chez un adulte et le diagnostic différentiel pourrait être une pneumonie, une exacerbation de maladie obstructive chronique ou une insuffisance cardiaque décompensée.

Après avoir rédigé le scénario clinique, la clé de correction est établie par les cliniciens-enseignants responsables du cours à partir de leur expérience clinique et de la littérature sur les diagnostics différentiels. Une fois l'évaluation complétée, un échantillon aléatoire de 10 % des copies d'étudiants est tiré. Les réponses issues de cet échantillon sont analysées par le ou les cliniciens-enseignants. Si certaines de ces réponses sont jugées acceptables mais qu'elles ne faisaient pas partie de la grille de correction initiale, elles sont ajoutées. Enfin, une correction dichotomique est faite, c'est-à-dire qu'un point est attribué pour chaque bonne réponse et aucun point n'est donné pour une mauvaise réponse ou encore pour une non-réponse. Ainsi, dans l'exemple précédent, le résultat d'un étudiant peut varier entre 0 et 18 points.

Utiliser une correction dichotomique permet une correction simple et efficace. On pourrait néanmoins considérer d'ajuster la pondération en fonction des éléments les plus discriminants, par exemple en soumettant chaque QPC à un panel d'experts. Par contre, ceci aurait l'inconvénient de rendre plus fastidieux leur rédaction et fort probablement de freiner leur utilisation.

Objectif

L'objectif de cet article est de présenter en quoi consistent les QPC et d'en documenter les qualités psychométriques. Nous avons jugé opportun de procéder le plus tôt possible à cet examen des qualités psychométriques, car cette démarche permet de cumuler des preuves fort utiles pour juger la validité d'une évaluation^[18], la validité étant considérée comme la caractéristique la plus importante de toute évaluation^[19].

Méthodes

Contexte et participants

À l'Université Laval, dans le cadre du programme de doctorat en médecine, les étudiants du pré-externat (années 1 à 3) participent à une série de cinq cours intitulés Démarche clinique 1 (DC1) à Démarche clinique 5 (DC5), à raison d'un par trimestre. Ces cours visent le développement du raisonnement clinique, des habiletés de communication médecin-patient et de l'examen physique. Chaque cours de cette série utilise notamment un examen écrit combiné à un examen pratique de fin de semestre. Aux fins du présent projet, une QPC a été intégrée, à l'automne 2012, à l'examen écrit du cours DC1 et une autre à celui du cours DC3. L'examen écrit du cours DC1 a été administré à une cohorte de 222 étudiants de première année et celui de DC3 à une cohorte de 216 étudiants de deuxième et troisième années. Les résultats de ces étudiants à l'examen pratique de fin de semestre ont également été recueillis afin de les mettre en relation avec les résultats à l'examen écrit.

Examens

Les examens écrits des cours DC1 et DC3 retenus pour l'étude avaient tous deux pour but d'évaluer chez les étudiants la démarche de raisonnement clinique. La QPC proposée aux étudiants des deux cohortes était différente quant à son contenu clinique, car elle était en lien avec les systèmes physiologiques étudiés dans chacun des deux cours. L'examen écrit de DC1 portait sur des cas cliniques relatifs aux systèmes cardiaque et pulmonaire. Il était composé de 33 questions, dont 28 questions à choix multiples (QCM), trois questions à réponse courte (QRC), un problème de raisonnement clinique (PRC) et une QPC. L'examen écrit du cours DC3 portait quant à lui sur des cas cliniques relatifs au système neurologique et à la psychiatrie. Ce dernier était constitué de 30 questions, dont 20 QCM, sept QRC, une question de type « vrai ou faux », un PRC et une QPC. De manière à assurer la validité de contenu des deux examens, toutes les questions, incluant les QPC, ont été rédigées par deux cliniciens-enseignants spécialistes du raisonnement clinique.

L'examen pratique de DC1 et de DC3 est une tâche d'intégration des habiletés cliniques. L'étudiant doit, à partir d'une raison de consultation, procéder à l'anamnèse et à l'examen physique pertinent. Le raisonnement clinique est intrinsèque à cette démarche. Il est objectivé par le résumé de la situation clinique, l'explication de la démarche diagnostique, et par l'élaboration et la justification du diagnostic différentiel de ce cas. La note de cet examen reflète l'ensemble de la démarche diagnostique.

Analyses

Afin de comparer les qualités psychométriques des QPC à celles des autres modalités d'évaluation écrites du raisonnement clinique utilisées dans les deux examens, nous avons calculé leurs indices classiques d'items et leur validité prédictive. En plus de nous renseigner sur le degré de cohérence interne des deux examens écrits, l'analyse d'items nous a permis d'estimer l'indice de difficulté (c'est-à-dire, la moyenne) et de discrimination (c'est-à-dire, la corrélation item-total corrigée) des QPC et des

Tableau I. Statistiques descriptives et indice de fidélité de l'examen écrit et de l'examen pratique des cours Démarche clinique 1 et 3

Cours	Examen	Coefficient Alpha de Cronbach	Moyenne	Écart-type	Min.	Max.
DC1	Écrit	0,61	84,74	5,63	65,75	96,50
	Pratique	0,62	80,74	8,18	51,6	96,20
DC3	Écrit	0,33	76,83	6,22	52,50	89,00
	Pratique	0,68	77,70	9,92	46,30	98,40

DC1 : Cours Démarche clinique 1

DC3 : Cours Démarche clinique 3

Min. : Valeur minimale observée

Max. : Valeur maximale observée

autres types d'items, c'est-à-dire les QCM, les PRC, les QRC et les questions à réponse « vrai ou faux ». Le barème d'interprétation de l'indice de difficulté a été élaboré à partir de l'expérience des professeurs et en tenant en compte du fait que le seuil de réussite à l'examen était de 60 % : Très facile = 90 % à 100 % ; Facile = 80 % à 89 % ; Modéré = 70 % à 79 % ; Difficile = 60 % à 69 % ; Très difficile = 0 % à 59 %. Le barème d'interprétation de l'indice de discrimination a quant à lui été construit en s'inspirant des seuils suggérés par Nunnally et Bernstein^[20], Schmeiser et Welch^[21] ainsi que Laveault et Grégoire^[22] : Bon = 0,30 à 1,00 ; Modéré = 0,20 à 0,29 ; Faible = 0,10 à 0,19 ; Très faible = 0,01 à 0,09 ; Non discriminant = -1,00 à 0. Par la suite, nous avons vérifié la validité prédictive des QPC et des autres types de questions en utilisant les résultats à l'examen pratique de fin de semestre. Pour ce faire, dans les cours DC1 et DC3, nous avons calculé la force d'association entre le résultat à chaque question de l'examen écrit et celui à l'examen pratique de fin de semestre à l'aide du coefficient de corrélation de Spearman^[23].

Résultats

Analyse d'items

Les statistiques descriptives et l'indice de fidélité de l'examen écrit et de l'examen pratique des cours DC1 et DC3 sont présentées au tableau I. Le

coefficient alpha de Cronbach de chacun des examens se situe entre 0,33 et 0,68.

Les tableaux II et III présentent les résultats de l'analyse de l'indice de difficulté des items aux examens écrits. Dans l'examen de DC1, l'indice de difficulté de la QPC est de 73,99 et celui des autres types de questions varie entre 46,20 et 100,00. Selon le barème utilisé, la QPC est d'un niveau de difficulté modéré et fait partie du quartile des questions les plus difficiles dans cet examen. Dans l'examen de DC3, l'indice de difficulté de la QPC est de 74,77 et celui des autres questions varie entre 46,23 et 100,00. Comme à l'examen de DC1, la QPC est d'un niveau de difficulté modéré et fait partie du quartile des questions les plus difficiles.

Les tableaux II et IV présentent aussi les résultats de l'analyse de l'indice de discrimination des items aux examens écrits. Dans l'examen de DC1, l'indice de discrimination de la QPC est de 0,39 et celui des autres questions varie entre -0,03 et 0,35. La QPC occupe donc le premier rang sur les 33 questions de l'examen en termes de pouvoir discriminant. Selon le barème utilisé, elle fait partie des 12 % des questions à cet examen possédant un bon pouvoir discriminant. Dans l'examen de DC3, l'indice de discrimination de la QPC est de 0,15 et celui des autres questions varie entre -0,04 et 0,25. La QPC occupe ici le neuvième rang sur les 30 questions en termes de pouvoir discriminant. Selon le barème utilisé, elle possède dans cet examen-ci un faible pouvoir discriminant, mais fait partie du 40 % des questions ayant le meilleur pouvoir discriminant.

Tableau II. Indices de difficulté et de discrimination aux examens écrits de Démarche clinique 1 et 3, selon le type de questions

Cours	Type de question	Nbre	Indice de difficulté	Indice de discrimination
DC1	QPC	1	73,99	0,39
	PRC	1	84,05	0,25
	QRC	3	Entre 89,06 et 92,57	Entre 0,19 et 0,35
	QCM	28	Entre 46,20 et 100,00	Entre -0,03 et 0,34
DC3	QPC	1	74,77	0,15
	PRC	1	46,23	0,17
	Vrai ou faux	1	94,67	0,16
	QRC	7	Entre 59,38 et 91,67	Entre -0,01 et 0,25
	QCM	20	Entre 63,00 et 100,00	Entre -0,04 et 0,24

DC1 : Cours Démarche clinique 1

DC3 : Cours Démarche clinique 3

QPC : Question de planification clinique

PRC : Problème de raisonnement clinique

QRC : Question à réponse courte

QCM : Question à choix multiple

Nbre : Nombre de questions

Tableau III. Distribution du niveau de difficulté des items des examens écrits de Démarche clinique 1 et 3

Niveau de difficulté	Valeurs	Démarche clinique 1		Démarche clinique 3	
		Nbre	%	Nbre	%
Très facile	90 % à 100 %	16	48	19	63
Facile	80 % à 89 %	10	33	4	13
Modéré	70 % à 79 %	3	10	3	10
Difficile	60 % à 69 %	2	7	2	7
Très difficile	0 % à 59 %	2	7	2	7

Nbre : Nombre de questions

Tableau IV. Pouvoir de discrimination des items des deux examens écrits de Démarche clinique 1 et 3

Pouvoir discriminant	Valeurs	Démarche clinique 1		Démarche clinique 3	
		Nbre	%	Nbre	%
Bon	0,30 à 1,00	4	12	0	0
Modéré	0,20 à 0,29	8	24	3	10
Faible	0,10 à 0,19	11	33	9	30
Très faible	0,01 à 0,09	6	18	11	37
Non discriminant	-1,00 à 0	4	12	7	23

Nbre : Nombre de questions

Validité prédictive

Le tableau V présente, pour les deux examens écrits, l'indice de validité prédictive (coefficient de

corrélation de Spearman) selon le type de question. Dans l'examen de DC1, l'indice de validité prédictive de la QPC est de 0,13 ($p = 0,076$) et celui des autres questions varie entre -0,06 ($p = 0,371$) et 0,12

Tableau V. Indice de validité prédictive selon le type de questions aux examens écrits de Démarche clinique 1 et 3

Cours	Type de question	Nbre	Corrélation de Spearman
DC1	QPC	1	0,13
	PRC	1	0,04
	QRC	3	Entre 0,03 et 0,05
	QCM	28	Entre -0,06 et 0,12
DC3	QPC	1	0,06
	PRC	1	0,09
	Vrai ou faux	1	0,06
	QRC	7	Entre 0,03 et 0,20
	QCM	20	Entre -0,03 et 0,13

DC1 : Cours Démarche clinique 1

DC3 : Cours Démarche clinique 3

QPC : Question de planification clinique

PRC : Problème de raisonnement clinique

QRC : Question à réponse courte

QCM : Question à choix multiple

Nbre : Nombre de questions

($p = 0,072$), avec une valeur médiane de 0,091. Dans l'examen de DC3, l'indice de validité prédictive de la QPC est de 0,06 ($p = 0,369$) et celui des autres questions varie entre -0,03 ($p = 0,692$) et 0,20 ($p = 0,003$), avec une valeur médiane de 0,06.

Discussion

L'objectif de cet article était de présenter les QPC et d'en documenter les qualités psychométriques. Les QPC permettent d'évaluer de manière ciblée un élément crucial du raisonnement clinique de l'étudiant, soit sa capacité à identifier les éléments cliniques qui permettront de discriminer entre des hypothèses diagnostiques probables. Elles ne visent donc pas à poser un jugement sur la qualité de l'ensemble du raisonnement clinique. En ce sens, elles se situent en complément aux autres moyens existants. Il est d'ailleurs reconnu qu'il est préférable de recourir à plus d'un moyen pour évaluer un processus aussi complexe que le raisonnement clinique^[12]. Rappelons aussi que les QPC n'évaluent pas la capacité à générer des diagnostics différentiels puisque ces derniers sont fournis à l'étudiant. Les QPC comportent différents avantages,

notamment celui de ne pas influencer l'étudiant avec un choix de réponse pour les éléments discriminants à identifier. Aussi, ce modèle de question est simple à concevoir, les données sont limitées, les consignes sont simples à comprendre et le nombre de réponses attendues est clairement défini. Tout en étant faciles à interpréter, ces questions sont susceptibles d'amener l'étudiant dans un processus de réflexion s'apparentant à celui du clinicien devant le même type de situation dans sa pratique.

D'un point de vue psychométrique, les QPC apparaissent être un moyen viable d'évaluation écrite d'une partie du raisonnement clinique. Les résultats des analyses réalisées montrent que les QPC présentent des qualités psychométriques comparables ou supérieures aux autres modalités d'évaluation écrite du raisonnement clinique que nous avons utilisées dans nos deux examens, soit les problèmes de raisonnement clinique, les questions à choix de réponses, les questions à réponse courte et les questions de type « vrai ou faux ». Dans le contexte de nos deux examens, les QPC présentaient un niveau de difficulté modéré. Ce niveau de difficulté est généralement souhaitable étant donné qu'il augmente la probabilité d'obtenir de bons indices de discrimination^[24]. Le pouvoir de discrimination des QPC était acceptable,

variant entre faible et bon, démontrant qu'elles arrivent à départager les étudiants forts des étudiants faibles. De plus, elles se situaient toujours dans le créneau des 40 % de questions les plus discriminantes.

Quant à leur validité prédictive, les QPC n'étaient pas associées de manière statistiquement significative aux résultats aux examens pratiques de fin de trimestre, tout comme la quasi-totalité des questions aux deux examens écrits. Plusieurs hypothèses pourraient expliquer cette corrélation nulle. D'abord, les QPC et les autres questions aux examens écrits évaluent de manière isolée différentes parties du raisonnement clinique, généralement des savoir-faire, alors que les examens pratiques évaluent la capacité des étudiants à mobiliser efficacement ces différents savoir-faire pour résoudre des problèmes cliniques. Il est fort possible que la maîtrise des savoir-faire, pris individuellement, ne soit pas corrélée à la capacité de les mobiliser efficacement. En d'autres termes, l'examen écrit et l'examen pratique évaluent des niveaux différents de la compétence clinique selon la pyramide de Miller^[25] et ceci risque d'avoir nuit à l'importance des corrélations observées. Par ailleurs, la spécificité de cas peut avoir joué un rôle dans ces résultats^[12]. Enfin, les apprentissages réalisés entre l'examen écrit et l'examen pratique ont sans doute contribué à diminuer l'importance de la corrélation entre ces deux évaluations.

Limites

Avant de recommander l'utilisation des QPC sans réserve, leurs qualités psychométriques devront être davantage étudiées, car la présente étude comporte certaines limites. Premièrement, il s'agit d'une mise à l'épreuve des QPC limitée à deux examens. La généralisabilité des résultats reste donc toujours à démontrer. Deuxièmement, la performance d'un clinicien en situation de résolution de problèmes cliniques est dépendante du type de cas cliniques, ce qui traduit l'impact de la spécificité des cas^[7]. Pour cette raison, il faut s'attendre à ce que des questions d'examen se rapportant à des cas cliniques différents soient moins fortement corrélées entre elles, affectant à la baisse la cohérence interne. En ce sens, il pourrait être

préférable de juger la fidélité d'une évaluation du raisonnement clinique en estimant sa stabilité temporelle plutôt que sa cohérence interne, ce qui n'était pas possible dans le contexte des cours de Démarche clinique. Pour cette même raison, les comparaisons entre les types de questions ont probablement été contaminées par le fait que chaque question reposait sur un cas clinique différent. Utiliser un même cas clinique pour plusieurs types de questions faciliterait donc leur comparaison. Troisièmement, afin de mieux documenter leur validité, d'autres preuves devront être recueillies, et ce, surtout relativement au contenu et au processus de réponse^[14]. Il serait notamment nécessaire de recourir à une méthodologie de nature qualitative, par exemple en utilisant la méthode de pensée à voix haute (*think aloud*)^[26], afin d'évaluer le processus de réponse aux QPC et de mieux comprendre les processus cognitifs impliqués. Ce type d'étude n'était par contre pas possible dans notre contexte d'évaluation sommative.

Conclusion

Cet article visait à présenter les QPC, une nouvelle modalité d'évaluation écrite des ressources liées au raisonnement clinique, et d'en documenter les qualités psychométriques. Les QPC permettent d'évaluer l'un des constituants fondamentaux majeur pour les étudiants novices en médecine, soit l'identification des éléments-clés spécifiques des diagnostics différentiels. Les analyses réalisées montrent que les QPC ont des qualités psychométriques au minimum comparables à celles d'autres modalités d'évaluation écrite du raisonnement clinique. En somme, les QPC apparaissent comme une nouvelle modalité d'évaluation écrite du raisonnement clinique facile à concevoir, simple à comprendre pour l'étudiant, viable et complémentaire aux autres modalités existantes. Néanmoins, la validation des QPC demeure à parfaire et d'autres études devront être menées en ce sens. Enfin, il serait important d'explorer le potentiel pédagogique des QPC dans l'enseignement du raisonnement clinique, par exemple en les utilisant de manière formative plutôt que sommative ou encore dans

l'identification des sources de difficultés chez les étudiants présentant des problèmes de raisonnement clinique.

Contributions

Jean-Sébastien Renaud a contribué à la conception du protocole de recherche, à l'interprétation des résultats, aux analyses statistiques, à l'écriture et à la révision du manuscrit. François Ratté a contribué à la conception du protocole de recherche, au recueil des données, à l'écriture et la révision du manuscrit. Julie F. Thériault a contribué à la conception du protocole de recherche, au recueil des données, à l'écriture et la révision du manuscrit. Anne Melissa Roy a contribué aux analyses statistiques, à l'interprétation des résultats et à l'écriture du manuscrit. Luc Côté a contribué à l'écriture et à la révision du manuscrit.

Déclaration d'intérêts

Aucun auteur ne déclare de conflit d'intérêt en lien avec le contenu de cet article

Approbation éthique

Les projets de cette nature sont considérés par le Comité d'éthique de la recherche avec des êtres humains de l'Université Laval comme un « Projet impliquant uniquement l'évaluation de programme, l'assurance de la qualité ou la revue et l'examen de programmes ou de services », qui ne nécessitent pas d'approbation éthique.

Valorisation scientifique

Une partie de ce travail a fait l'objet d'une communication orale présentée à la Conférence canadienne sur l'éducation médicale (CCEM), qui s'est tenue à Québec (Québec, Canada) du 20 au 23 avril 2013.

Références

1. Nendaz M, Charlin B, Leblanc V, Bordage G. Le raisonnement clinique: données issues de la recherche et implications pour l'enseignement. *Pédagogie Médicale* 2005;6:235-54.
2. Charlin B, Lubarsky S, Millette B, Crevier F, Audétat M-C, Charbonneau A, et al. Clinical reasoning processes: unravelling complexity through graphical representation. *Med Educ* 2012;46:454-63.
3. Charlin B, Tardif J, Boshuizen HPA. Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. *Acad Med* 2000;75:182-90.
4. Charlin B, Boshuizen HPA, Custers EJ, Feltovich PJ. Scripts and clinical reasoning. *Med Educ* 2007;41:1178-84.
5. Schmidt HG, Rikers RMJP. How expertise develops in medicine: knowledge encapsulation and illness script formation. *Med Educ* 2007;41:1133-9.
6. Bordage G. Elaborated knowledge: a key to successful diagnostic thinking. *Acad Med* 1994;69:883-5.
7. Elstein AS, Schulman LS, Sprafka SA. *Medical problem solving: an analysis of clinical reasoning*. Cambridge, MA: Harvard University Press, 1978.
8. Page G, Bordage G. The Medical Council of Canada's key features project: a more valid written examination of clinical decision-making skills. *Acad Med* 1995;70:104-10.
9. Page G, Bordage G, Allen T. Developing key-feature problems and examinations to assess clinical decision-making skills. *Acad Med* 1995;70:194-201.
10. Bowles LT. A worthy search: the development of the key-features concept. *Acad Med*. 1995;70:88-90.
11. Bordage G, Brailovsky C, Carretier H, Page G. Content validation of key features on a national examination of clinical decision-making skills. *Acad Med* 1995;70:276-81.
12. Charlin B, Bordage G, Van Der Vleuten C. L'évaluation du raisonnement clinique. *Pédagogie Médicale*. 2003;4:42-52.
13. Ratté F, Thériault JF. *Démarche clinique : raisonnement clinique, examen physique, entrevue médicale*. Québec: Presses de l'Université Laval, 2012.

14. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet* 2001; 357:945-9.
15. Raupach T, Brown J, Anders S, Hasenfuss G, Harendza S. Summative assessments are more powerful drivers of student learning than resource intensive teaching formats. *BMC Medicine* 2013;11: 1-10.
16. Hodges B, Regehr G, McNaughton N, Tiberius R, Hanson M. OSCE checklists do not capture increasing levels of expertise. *Acad Med* 1999;74: 1129-34.
17. Bowen JL. Educational strategies to promote clinical diagnostic reasoning. *N Engl J Med* 2006;355: 2217-25.
18. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Educational Research Association, 2014.
19. Downing SM, Haladyna TM. Validity and its threats. In: Downing SM, Yudkowsk R, editors. *Assessment in health professions education*. New York, NY: Routledge, 2009:21-55.
20. Nunnally J, Bernstein L. *Psychometric theory*. 3rd ed. New York: McGraw-Hill, 1994.
21. Schmeiser CB, Welch CJ. Test development. In: Brennan RL, editor. *Educational measurement*. 4th ed. Westport (CT): Praeger Publishers, 2006:307-53.
22. Laveault D, Grégoire J. *Introduction aux théories des tests en psychologie et en sciences de l'éducation*. 3e ed. Bruxelles: De Boeck, 2014.
23. Spearman C. The proof and measurement of association between two things. *Am J Psychol*. 1904;15: 72-101.
24. Crocker LM, Algina J. *Introduction to classical and modern test theory*. Manson (OH): Cengage Learning, 2006.
25. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65:S63-7.
26. Fonteyn ME, Kuipers B, Grobe SJ. A description of think aloud method and protocol analysis. *Qual Health Res* 1993;3:430-41.

Correspondance et offprints : Jean-Sébastien Renaud, Faculté de médecine Pavillon Ferdinand-Vandry, local 2881-D Université Laval, 1050 avenue de la Médecine, Québec (Québec) G1V 0A6.
Mailto : jean-sebastien.renaud@fmed.ulaval.ca