

Evaluer la dimension d'incertitude du raisonnement clinique

Assessing the uncertainty component of clinical reasoning

D'une manière générale, on considère que les professionnels sont confrontés à deux types de problèmes à résoudre¹. Certains problèmes, dits simples (well defined problems) ou encore bien structurés (well structured problems), sont caractérisés par le fait que l'ensemble des données nécessaires est d'emblée accessible et que la solution peut résulter d'une simple application de connaissances ; ils peuvent être résolus avec un haut degré de certitude et, d'ailleurs, les professionnels experts du domaine concerné s'accordent en général sur la nature de la solution correcte à mettre en œuvre. D'autres problèmes, dits complexes (ill defined problems) ou encore mal structurés (ill structured problems), sont en revanche caractérisés par le fait que les données constitutives ne sont pas toutes disponibles d'emblée et que leur solution n'est pas univoque, impliquant une réflexion d'ordre supérieur qui requiert la mobilisation de connaissances diverses ; de tels problèmes ne peuvent pas être résolus avec un haut degré de certitude et, d'ailleurs, les professionnels experts du domaine concerné sont souvent en désaccord quant à la meilleure solution à mettre en œuvre, y compris a posteriori quand le problème peut être considéré comme ayant été résolu.

De façon spécifique, il est reconnu que les problèmes auxquels sont confrontés les professionnels de santé sont, dans l'immense majorité des cas, de type complexe ou mal structuré. Le raisonnement clinique, considéré au sens générique du terme -c'est-à-dire en y incluant à la fois le processus de raisonnement diagnostique et celui de prise de décisions d'investigation et de traitement-, désigne le processus de résolution de problèmes que les cliniciens mettent en œuvre face à de tels problèmes². Il n'est donc guère surprenant de constater que, face à un même problème de santé, les chemins empruntés pour élaborer les solutions varient selon les cliniciens, y compris au sein d'une même discipline.

Selon Schön³, les institutions d'enseignement supérieur entraînent bien leurs étudiants à la résolution des problèmes de structure simple mais mal à celle des problèmes de structure complexe. La même remarque peut s'appliquer à l'évaluation du raisonnement clinique. Nous disposons de bons outils pour évaluer la résolution de problèmes cliniques de structure simple et de peu d'outils pour évaluer celle des problèmes de structure complexe, notamment pour ceux qui impliquent

la prise de décision en contexte d'incertitude, caractéristique clé de l'exercice professionnel dans les sciences de la santé.

Si ces prémisses sont exactes, un outil d'évaluation du raisonnement clinique qui sonde la capacité à résoudre les problèmes de structure simple devrait être régulièrement en défaut pour déceler les personnes réellement compétentes dans un domaine de raisonnement ; en effet, des étudiants qui ont accumulé beaucoup de connaissances, déclaratives en particulier, peuvent mieux réussir à ces tests que des praticiens, dont les savoirs livresques sont moins disponibles, mais qui sont riches d'expériences cliniques et qui ont, de ce fait, construit de riches et pertinents réseaux de connaissances d'action. La littérature désigne cela sous le terme d'effet intermédiaire de l'évaluation : les cliniciens de niveau intermédiaire, des résidents ou internes en fin de formation par exemple, obtiennent régulièrement à ce type de tests des scores supérieurs à ceux des praticiens expérimentés. À l'inverse, s'il était établi que des praticiens possédant plusieurs années d'expérience obtenaient régulièrement de meilleurs scores à tel ou tel test d'évaluation du raisonnement clinique, cela constituerait un argument fort en faveur de la validité de ces tests, c'est-à-dire leur capacité à bien mesurer ce qu'ils prétendent mesurer, en l'espèce la qualité du raisonnement clinique.

Les questions à choix multiples (QCM) portant sur des cas cliniques, dits à contexte riche, sont tenues pour être actuellement un standard en matière d'évaluation du raisonnement clinique⁴. Elles présentent un cas clinique pour lequel le candidat doit synthétiser l'information disponible, toutes les données nécessaires à la solution du problème étant fournies ; par construction, il n'existe qu'une seule bonne réponse. Ces examens permettent une correction automatisée et peuvent être administrés à de très grands groupes d'étudiants, d'où leur large utilisation dans les examens de certification, aux États-Unis notamment.

Le test de concordance de script (TCS) est un format de test plus récent, qui cherche à prendre en compte la dimension d'incertitude inhérente au processus de raisonnement clinique⁴. Il présente un cas clinique pour lequel toute l'information clinique nécessaire à sa résolution n'est pas fournie ou dans lequel certaines

données sont ambiguës. Dans de tels contextes, nous savons que les cliniciens experts adoptent des chemine-
ments de raisonnement variables. Il ne peut donc y
avoir une seule « bonne réponse ». Cette variabilité est
incorporée dans le processus d'attribution du score, qui
pondère les indices numériques affectés à chacune des
réponses proposées à l'étudiant en fonction des propor-
tions respectives d'experts ayant sélectionné celles-ci.
Une telle procédure permet de dépasser l'appréciation
binaire en « tout ou rien ».

L'étude présentée dans ce numéro par Jean-Pierre
Fournier et ses collaborateurs⁵ compare les résultats
obtenus à ces deux formats de tests par plusieurs popu-
lations de médecins, de niveaux d'expérience différents,
travaillant en salle d'urgence. Les deux tests ont été réa-
lisés avec soin comme en témoigne les excellentes valeurs
de leurs coefficients respectifs de cohérence interne
(mesurées par le test alpha de Cronbach). Les résultats
montrent une disparition de l'effet intermédiaire. Les
médecins les plus expérimentés obtiennent de meilleurs
scores que les médecins de niveau intermédiaire (les
internes). Ce travail est le premier à démontrer une telle
aptitude discriminante dans la littérature internatio-
nale. Par ailleurs, les scores obtenus aux TCS et aux
QCM varient dans le même sens mais sans corrélation
significative. Ce point indique que QCM et TCS
explorent deux aspects distincts mais complémentaires
de la compétence au raisonnement clinique. Ces résul-
tats devront être confirmés par d'autres études et dans
d'autres disciplines mais, s'ils étaient confirmés, ils sug-
géreraient que le TCS peut légitimement revendiquer
une originalité et une valeur ajoutée au sein des outils
disponibles pour l'évaluation du raisonnement clini-
que. Tout en respectant les exigences de la perspective
docimologique dans laquelle il s'inscrit, et à côté d'au-
tres outils éprouvés, qui mesurent déjà les dimensions les
moins incertaines et les moins ambiguës du raisonne-
ment clinique, tels que les QCM à contexte riche, il
pourrait compléter avec profit les moyens disponibles ou
en développement, qui cherchent à appréhender la
dimension d'incertitude propre au processus de raisonne-
ment clinique. Quand on connaît les conséquences
qu'ont les procédés d'évaluation à l'égard des apprentis-
sages des étudiants -et dont rend compte, par exemple,
l'adage « l'évaluation conduit le curriculum » -, l'intro-
duction d'un tel outil serait de nature à orienter favo-
rablement les activités d'enseignement et d'apprentis-
sage, au service d'une meilleure préparation des futurs
professionnels à la dimension d'incertitude du raisonne-
ment clinique et à la réalité de la pratique clinique.

Il reste que les problèmes que soulève le souhait de
prendre authentiquement en compte la dimension
d'incertitude du raisonnement clinique ne sont pas
tous résolus. Dès lors que l'on reconnaît qu'il y a, en
effet et presque toujours, plusieurs manières de résoudre
avec compétence un problème de santé donné, notre
rôle d'enseignant et d'évaluateur devrait être d'aider les
étudiants à être en mesure d'élaborer et de justifier
l'une des solutions raisonnables et acceptables possibles.
Dès lors, un étudiant qui propose et argumente correc-
tement l'une de ces solutions ne devrait pas être péna-
lisé dans l'attribution des scores, au seul motif qu'une
minorité parmi les experts a choisi cette solution. Bien
évidemment, les outils d'évaluation qui permettraient
de dépasser ce point ont d'autres inconvénients ou
limites : lourdeur, correction non automatisable, néces-
sité de s'inscrire dans une démarche plus qualitative
pour apprécier les critères classiques de validité et de
fidélité, etc. A l'évidence, la recherche en éducation des
sciences de la santé a encore des pistes fécondes à explorer
en matière d'évaluation du raisonnement clinique.

Bernard CHARLIN

Mailto:bernard.charlin@umontreal.ca

Références

1. King PM, Kitchener KS. *Developping reflective judgment : Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco (CA): Jossey-Bass, 1994.
2. Nendaz M, Charlin B, LeBlanc V, Bordage G. *Le raisonnement clinique : données issues de la recherche et implications pour l'enseignement*. *Pédagogie Médicale* 2005;6:234-53
3. Schön D. *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books, 1983.
4. Charlin B, Bordage G, Van der Vleuten C. *L'évaluation du raisonnement clinique*. *Pédagogie Médicale* 2003;4:42-52.
5. Fournier J-P, Thiercelin D, Pulcini C 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.