

Quelques questions posées par l'épreuve d'évaluation de la lecture critique d'articles scientifiques

A l'issue du deuxième cycle de leurs études (études pré-gra-duées du cursus nord-américain), tous les étudiants en médecine français doivent désormais passer des épreuves classantes nationales (ECN) anonymes^{1,2}. Conformément à une conception républicaine bien établie du mérite et de l'équité, le classement à ces épreuves conditionne les choix de chaque étudiant; il détermine à la fois la filière de formation spécialisée à laquelle il pourra accéder ainsi que l'université et le centre hospitalier universitaire de rattachement, au sein desquels il recevra sa formation. Fort logiquement, au cours du deuxième cycle, les étudiants orientent leurs activités vers la préparation des ECN. Une partie importante de cette préparation s'effectue en dehors des facultés, au sein d'officines privées, au détriment toléré et parfois encouragé des enseignements réguliers et des stages hospitaliers; en dépit de ce biais, des palmarès officiels des facultés de médecine françaises circulent, extrapolés imprudemment mais avec complaisance à partir du classement des étudiants. Nonobstant ces évolutions préoccupantes, dont il faudra bien un jour évaluer l'impact sur la qualité des apprentissages des étudiants diplômés et donc sur les compétences des médecins formés, l'important budget que les pouvoirs publics consacrent chaque année à l'organisation des ECN (1,35 millions d'euros)³ souligne l'apparent consensus social établi autour de cette modalité d'orientation professionnelle finale des étudiants en médecine en France.

Ces contraintes académiques et sociales étant jusqu'à nouvel ordre acceptées -et leurs conséquences pédagogiques étant implicitement assumées-, il convient au moins que les modalités de l'épreuve satisfassent effectivement les standards de rigueur, garants du principe d'équité qui en fonde la légitimité. L'exigence d'avoir à classer avec fiabilité 4500 étudiants -et probablement 7000 dès 2011⁴- ne laisse guère le choix de convoquer le regard docimologique pour en apprécier les critères. Un tel examen devrait donc être à la fois valide, fidèle et réalisable avec des contraintes acceptables. Consistant actuellement en l'analyse de neuf dossiers cliniques, à propos desquels les étudiants sont interrogés à l'aide de quatre à dix questions à réponse ouverte, il devra comporter en outre, à partir de 2008, une épreuve de lecture critique d'un ou plusieurs articles scientifiques (LCA).

En raison de sa nouveauté, cette dernière épreuve suscite un questionnement et des inquiétudes, tant de la part de certains enseignants que de la plupart des étudiants. Dès 2001, dans ces mêmes colonnes⁴, Jean-Paul Fillastre et Raymond Colin formulaient des réserves quant à la décision d'inclure prématurément une telle épreuve, « alors

que nous ne disposons pas encore de méthodes valides et fidèles pour évaluer cette capacité »; ils mettaient en garde contre les « risques importants qui pourraient altérer l'équité de l'examen national et aboutir à des situations de contentieux ». En 2002, à la suite d'un mouvement de grève national, les étudiants en médecine obtenaient des ministères de tutelle le report de l'épreuve à 2008, compte tenu d'une préparation à l'exercice qu'ils jugeaient insuffisante et surtout très inégale d'une faculté à l'autre. Depuis lors, les différentes facultés se sont tant bien que mal attelées à la tâche; en se fondant sur des approches pédagogiques issues de leurs cultures institutionnelles propres, elles ont mis en œuvre des activités d'enseignement, d'apprentissage et d'évaluation de la LCA selon des modalités variables. Seule une minorité de ces expériences fait l'objet d'une évaluation systématique et, à notre connaissance, aucune initiative coordonnée n'a été mise en place pour en mutualiser les résultats et en tirer des conclusions. Il faut donc saluer les initiatives pionnières en ce domaine et remercier les auteurs qui inscrivent ces démarches exploratoires dans une logique scientifique d'évaluation par leurs pairs, en publiant leurs travaux.

L'article de Christine Louis-Sylvestre et coll.⁵, qui paraît dans ce numéro, fait suite au travail de Francis Roussel et coll.⁶, publié dans le précédent. Les résultats des deux études accréditent l'idée que, dans l'état actuel de la conception de l'épreuve et de préparation des enseignants chargés de la corriger, l'organisation d'une épreuve de LCA dans le cadre des ECN soulève des questions importantes. Certes, les deux études comportent des limites méthodologiques, que leurs auteurs discutent d'ailleurs fort loyalement. Des spécialistes en orthodoxie docimologique pourraient souhaiter qu'elles aient eu recours à des raffinements expérimentaux et statistiques plus exigeants. Il demeure que des enseignants engagés, soucieux de leur responsabilité sociale, ont montré la voie en s'astreignant à une démarche réflexive quant à leurs pratiques d'enseignants et d'évaluateurs. De ce fait, même si la solidité de leur démonstration et la généralisabilité de leurs conclusions ne peuvent être tenues comme définitivement établies, leurs résultats sont suffisamment significatifs pour introduire un doute sérieux quant à la fidélité exigible de la correction de l'épreuve de LCA, même quand on utilise une grille de correction. L'apparente sécurité apportée par l'exigence d'une double correction des épreuves ne peut pas complètement rassurer. En stipulant que « chaque dossier, corrigé en double correction indépendante, est noté de 0 à 100 [...] [et que] la note retenue est la moyenne arithmétique des deux notes »², les dispo-

sitions réglementaires actuelles acceptent de fait l'infidélité de l'instrument mais formulent une solution très hasardeuse à cet inconvénient, en postulant que la moyenne arithmétique de deux mesures erronées fournira une mesure exacte. En prévoyant que " lorsque, pour un même dossier, l'écart de notation constaté entre deux correcteurs est égal ou supérieur à 15 points, une troisième correction est assurée [et que] dans ce cas, la notation qui en résulte est retenue »², ces dispositions acceptent également le risque qu'une copie victime de deux erreurs de mesure se produisant dans le même sens et avec le même ordre de grandeur ne soit pas révisée.

Il n'est pas possible de statuer quant à la validité de l'épreuve puisque son évaluation n'était l'objet d'aucune des deux études rapportées. Il est très positif que les objectifs pédagogiques formulés par le conseil scientifique du centre national des concours d'internat⁷ prennent en compte plusieurs éléments du cadre de référence établi par Nicole Audet et Hélène Leclère, à l'issue d'une recension systématique des écrits, pour rendre compte des habiletés requises pour la LCA en médecine⁸. On peut cependant émettre l'hypothèse qu'il existe de sérieuses incertitudes quant à la validité de construits des épreuves qui seront proposées. Ce type de validité concerne la justesse et la pertinence avec lesquelles l'épreuve renseigne sur les apprentissages relatifs aux différentes composantes de la « construction » théorique des compétences visées. L'une des difficultés principales tient au fait que les aptitudes cognitives requises sont fortement contextualisées à des situations et à des contenus bio cliniques spécifiques. Des compétences transversales et transférables ne peuvent être construites, grâce à des interventions pédagogiques appropriées, qu'après une exposition itérative à une variété de situations particulières⁹. Dès lors, l'évaluation de telles compétences, comme le souligne Geoff Norman en conclusion d'une récente revue systématique de la littérature⁹, soulève de nombreuses questions non résolues qui devraient être abordées dans le contexte d'une démarche de recherche.

L'évaluation des apprentissages est d'abord un acte pédagogique. A cet égard, les enseignants doivent avoir conscience que les éléments qu'ils prennent en compte dans le cadre de leurs évaluations adressent aux étudiants des signaux extrêmement puissants quant aux apprentissages à privilégier. Au-delà du contexte français évoqué, les contributions préliminaires de Francis Roussel et Christine Louis-Sylvestre constituent une invitation expresse à ce qu'un groupe de travail ad hoc s'intéresse aux expériences de formation à la LCA et d'évaluation des apprentissages concernés. Dans un esprit d'expérimentation, acceptant que les procédures soient imparfaites mais cherchant constamment à les optimiser, ce groupe pourrait mutualiser les pratiques innovantes, les travaux de recherche évaluative et, à terme, formuler des recommandations. Dans son rôle de société savante, riche des ressources francophones qu'elle peut mobiliser, la Société internationale francophone d'éducation médicale (SIFEM) pour-

rait, sur sollicitation, apporter une expertise scientifique et contribuer à valoriser ces travaux.

L'évaluation des apprentissages est aussi une pratique sociale, qui permet à l'institution universitaire d'attester le niveau de connaissances et de compétences développées par les étudiants et de rendre publiques les décisions les concernant. Dans le contexte spécifique des exigences et des contraintes des ECN françaises, il appartient aux responsables universitaires et politiques d'apprécier les réponses à apporter à des difficultés et à des questions désormais plus clairement identifiées.

Jean Jouquan
mailto: jean.jouquan@chu-brest.fr

Références

1. Décret n° 2004-67 du 16 janvier 2004 relatif à l'organisation du troisième cycle des études médicales. Journal officiel de la République française du 18 janvier 2004: 1394-1395.
2. Arrêté du 29 janvier 2004 relatif à l'organisation des épreuves classantes nationales anonymes donnant accès au 3^e cycle spécialisé des études médicales. Journal officiel de la République française du 5 février 2004: 2519-2521.
3. Projet de loi de finances pour 2005: enseignement supérieur. Rapport général fait au nom de la commission des Finances, du contrôle budgétaire et des comptes économiques de la Nation sur le projet de loi de finances pour 2005 [Online]. Disponibles sur: <http://www.senat.fr/rap/l04-074-325/l04-074-3251.pdf>
4. Commission Démographie médicale. Rapport présenté par Berland Y. 2005 [On-line]. Disponible sur: <http://www.sante.gouv.fr/>
5. Louis-Sylvestre C, Furham C, Housset B. Difficultés de correction d'une épreuve d'analyse critique d'article scientifique: une étude exploratoire. Pédagogie Médicale 2005; 6: 138-146.
6. Roussel F, Czernichow P, Lavoigne A, Lemeland J-F, Fillastre J-P. Reproductibilité d'une épreuve de lecture critique d'article: évaluation par une étude pilote chez 59 étudiants en médecine. Pédagogie Médicale 2005; 6: 71-78.
7. Conseil scientifique du centre national du concours d'internat. Épreuve écrite de lecture d'article. 2003 [Online]. Disponible sur: <http://www.cnci.univparis5.fr/medecine/CritiqueArticle.doc>
8. Audet N, Leclère H. Les habiletés requises pour la lecture critique en médecine: un cadre de référence issu d'une recension des écrits. Pédagogie Médicale 2001; 2: 206-212.
9. Norman G. Critical thinking and critical appraisal. In: GR Norman, CPM Van der Vleuten, DI Newble (eds). International Handbook of Research in Medical Education. Dordrecht: Kluwer Academic Publishers, 2002: 277-298.