

L'approche de la notion de critère de jugement principal par les étudiants lors de l'épreuve de lecture critique d'article au cours des épreuves classantes nationales 2010

Monsieur,

Depuis 2004, les étudiants en médecine français sont soumis en fin de deuxième cycle à des épreuves classantes nationales (ECN), le rang de classement individuel qui en résulte déterminant l'orientation finale de chaque étudiant pour l'une ou l'autre des filières de formation spécialisée de troisième cycle (internat). Depuis 2009, le format des ECN a été enrichi : l'épreuve d'analyse de neuf dossiers cliniques a été complétée par une épreuve de lecture critique d'article (LCA), qui cherche à évaluer la capacité d'analyse critique de l'information scientifique du futur médecin^[1-3]. L'ajout de cette épreuve a un effet notable sur le classement individuel des étudiants mais pas sur le classement des facultés qui en est déduit^[4].

Nous rapportons notre expérience en tant que membres d'un jury chargé de la correction de l'épreuve de LCA lors de l'édition 2010 des ECN, en faisant part de notre étonnement face aux réponses des étudiants quant à la notion de critère de jugement principal dans le cadre d'une étude clinique.

En 2010, l'article soumis à l'analyse critique des étudiants rapportait un essai randomisé – CIBIS II, comparant dans une étude multicentrique, contrôlée, contre placebo, l'efficacité du bisoprolol sur la mortalité chez des patients insuffisants cardiaques symptomatiques de stades III et IV de la classification NYHA^[5]. La quatrième question était la suivante : quel est le critère de jugement principal de l'étude ?

Quels sont les éléments qui vous permettent de juger de sa validité ?

La grille de correction prenait en compte les réponses suivantes pour qualifier ce critère principal de jugement : mortalité toutes causes confondues ; critère objectif ; critère fréquent dans l'insuffisance cardiaque. Dès les premières minutes de la correction, les huit membres du jury ont été interloqués par certaines réponses originales concernant le critère principal (mortalité). D'un commun accord, nous avons décidé de colliger toutes les réponses originales, déplacées et humoristiques. Nous avons corrigé 300 copies en quatre binômes. La plupart des étudiants ont formulé une réponse correcte et le critère de jugement principal était adéquatement défini par « mortalité toutes causes confondues ». Les étudiants précisaient que ce critère est objectif, fréquent, pertinent, fiable, clinique, unique et défini *a priori*. Nous avons relevé 100 réponses originales et différentes, certaines étant citées dans plusieurs copies, qui sont présentées ci-dessous selon une typologie relative à certains aspects qualitatifs de ce critère :

- critère simple ou unique : critère *on/off*, binaire, dur, non négociable, réel et constant, non composite, censé, définitif, compris par tous, irréversible, pas de faux positif.
- critère mesurable : critère validé, facilement « mesurable » (mort/vivant), mesurable

sur une échelle, pas de variation inter- et intra-observateur.

- critère spécifique : universel, international, « indolore », non invasif, reproductif, facilement observable, consensuel, « bénéfique » pour le patient, acceptable pour le patient, « sans danger » pour le patient.
- critère réglementaire : éthique, validé par la science, validé par le comité d'éthique et de protection des personnes.

Certaines réponses nous ont paru très originales : « la mortalité est reproductible dans le temps et par d'autres équipes médicales » ; « le diagnostic de décès est consensuel sur le plan international » ; « la mortalité est un critère inamovible et constant, une fois réalisée » !

Au cours de l'apprentissage de la LCA, l'étudiant doit assimiler des notions nouvelles concernant la méthodologie de la recherche clinique, l'analyse des résultats et l'analyse statistique. Les apprentissages visés concernent donc à la fois une méthodologie spécifique et un répertoire terminologique adapté. Dans l'article proposé, le critère principal de jugement, « mortalité toutes causes confondues », concernait un événement clinique unique qui permet d'évaluer l'efficacité d'un médicament dans l'insuffisance cardiaque symptomatique. Ce critère robuste conditionnait l'analyse statistique du travail. Une majorité des étudiants a répondu simplement que ce critère était pertinent, objectif et fréquent dans l'insuffisance cardiaque. Nous avons observé que certains étudiants proposaient une liste de termes qualitatifs. Ils ont en fait compris que la correction s'appuyait sur l'identification, au sein de leurs réponses, des mots-clés qui constituent la « grille de correction » ; ils sont ainsi tentés d'en allonger la liste, parfois au détriment de leur pertinence, en espérant ainsi réussir la question. De fait, le jury a eu l'impression qu'un certain nombre d'étudiants essayaient de donner un nombre maximal de mots-clés concernant l'item critère de jugement principal. La conséquence est que certaines réponses paraissent décalées ou non pertinentes dans ce contexte. Le problème est qu'au milieu des

réponses décalées, les réponses justes étaient présentes, ce qui justifiait une bonne notation.

Nos réflexions comportent plusieurs limites. Elles ne résultent que d'un seul jury de quatre binômes de correcteurs et concernent de ce fait seulement 300 copies. Elles s'appuient par ailleurs sur un recueil des données (les réponses non conformes étudiants) par chaque membre du jury, qui n'avait pas été anticipé et qui, dès lors, a été en partie aléatoire ; néanmoins, une discussion collégiale nous a permis de discerner les réponses inadaptées.

Des stratégies de réponses centrées sur l'énoncé d'un maximum de mots-clés sont désormais explicitement promues parmi les étudiants préparant la LCA [6]. Seuls les mots-clés contenus dans la grille donnent des points. L'étudiant emmagasine une série de mots-clés qu'il cherchera à placer dans les réponses. Le jury a eu la certitude que cette attitude a prévalu pour certains étudiants qui n'ont probablement pas compris la robustesse de la mortalité, toute cause confondue, comme critère principal d'une étude portant sur un médicament dans l'insuffisance cardiaque.

Au-delà de la finalité et du bien fondé des stratégies de réponse des étudiants, nos constats interrogent aussi fondamentalement les qualités métrologiques (validité et fidélité) des procédures de correction des ECN, qui ne sont sans doute pas à la hauteur de ce qu'on pourrait attendre d'un examen aux enjeux individuels aussi importants pour les étudiants.

*Boris MELLONI*¹,

*Diane DESCAMPS*²,

*Laurence FAIVRE*³,

*Rémi MARIANOWSKI*⁴,

*Jacques MANSOURATI*⁴,

*Jean-Noël BASTIE*³,

*Jean-Hughes DALLE*²

*Damrong RATANASAVANH*⁴

¹ Université de Limoges, France

² Université Paris VII, France

³ Université de Bourgogne, Dijon, France

⁴ Université de Bretagne occidentale, Brest, France

Mailto : boris.melloni@unilim.fr

Références

1. Durieux P. La lecture critique n'est pas destinée à des universitaires déconnectés de la réalité. *Presse Med* 2009;38:1006-7.
2. Jouquan J. La lecture critique d'article au-delà du contexte franco-français... Pour éviter que l'arbre ne cache la forêt ! *Pédagogie Médicale* 2009;10:77-82.
3. Lorette G. La lecture critique fait partie des disciplines fondamentales. *Presse Med* 2009;38:1007.
4. Lorette G, Jolly D, Ambrosi P, Dreyfuss D, Magne J-L, Kohler C, Duhaut P, Gérard J-L. Influence de la lecture critique d'articles sur les classements des étudiants aux épreuves classantes nationales 2009. *Presse Med* 2010;39:e118-25.
5. CIBIS II Investigators and Committees. The cardiac Insufficiency Bisoprolol Study II (CIBIS II) a randomised trial. *Lancet* 1999;353:9-13.
6. Sroussi M, Bessiere F. Mots-Clés. Les points des grilles ECN. *Milon-La-Chapelle : S-éditions*, 2^e édition, 2011.