

Reproductibilité de la correction d'une épreuve de lecture critique d'article : évaluation par une étude pilote chez 59 étudiants en médecine

Francis ROUSSEL¹, Pierre CZERNICHOW², Alain LAVOINNE³,
Jean-François LEMELAND⁴, Jean-Paul FILLASTRE⁵

Résumé *Contexte* : L'introduction d'une épreuve de lecture critique d'articles dans l'examen classant national en 2008 implique de former les étudiants et de mettre au point une procédure reproductible de correction. Actuellement la formation est très variable entre les différentes facultés et la reproductibilité de ce type d'épreuve est inconnue. **But** : Évaluer la reproductibilité de la correction d'une épreuve de lecture critique d'articles dans des conditions proches de celles de l'examen national classant. **Sujets et méthodes** : 59 étudiants volontaires de 2^e et 3^e années du deuxième cycle des études médicales ont passé une épreuve surveillée de trois heures, sur un article original en français, comportant 10 questions validées et un résumé de 250 mots au plus. Trois enseignants ayant préparé le dossier de l'épreuve ont corrigé en aveugle les questions et le résumé selon 16 items, sans grille de correction. L'accord entre les trois paires de correcteurs a été apprécié par les coefficients de corrélation intra classe (CCI). **Résultats** : En moyenne, l'accord entre les correcteurs était modéré ou médiocre : CCI = 0,68 (intervalle de confiance à 95 % : 0,52 - 0,80), CCI = 0,71 (0,55 - 0,81) et CCI = 0,48 (0,26 - 0,65) selon la paire. Cette concordance était variable suivant les items (très mauvaise pour deux questions). Elle était meilleure pour les questions que pour le résumé pour 2 des 3 paires de correcteurs. **Conclusion** : Cette méthode de correction de l'épreuve de lecture critique d'articles est peu reproductible. L'impréparation actuelle des étudiants et des enseignants est probablement un facteur aggravant.

Mot clés Lecture critique d'articles ; évaluation ; reproductibilité inter correcteurs ; examen classant national.

Abstract *Context*: At french medical universities, a new national sorting exam (NSE) has been implemented, which will include from 2008 onwards a test of scientific papers critical analysis (SPCA). This implies further educational courses for students and a reliable marking process. Currently, educational development is highly variable between universities, and the reliability of the marking process remains unknown. **Aim**: To evaluate the reliability of SPCA test marking, in conditions close to those of the NSE. **Subjects and methods**: 59 volunteer medical students from 4th or 5th year at the Rouen university medical school completed a supervised examination of three hours, on an original article in French, including 10 validated open questions and an abstract of less than 250 words. Three teachers, who prepared the assessment, blindly marked the questions and the abstract, according to 16 items, without a correction grid. The agreement among the three pairs of correctors was evaluated according to the intraclass correlation coefficients (ICC). **Results**: The overall mean ICCs (95% confidence interval) for the 3 pairs were moderate or poor: 0.68 (0.52 - 0.80), 0.71 (0.55 - 0.81) and 0.48 (0.26 - 0.65). The marking agreement was better for the questions than for the abstract for 2 out of the 3 pairs of teachers. It was variable with the items, with a very poor score for two of them. **Conclusion**: The reliability of this SPCA test marking method was considered poor. The present lack of students and faculty preparation is probably a worsening factor.

Key words Papers critical analysis; assessment; reliability; national ranking exam.

Pédagogie Médicale 2005; 6: 71-78

1- Département d'Anatomie et de Cytologie Pathologiques - Faculté de Médecine - 76000 Rouen - France

2- Département d'Epidémiologie et de santé publique ; même adresse

3- Laboratoire de Biochimie ; même adresse

4- Département de Microbiologie ; même adresse

5- Service de Néphrologie ; même adresse

Correspondance : Francis Roussel - Pavillon Jacques Delarue - CHU de Rouen - 76031 Rouen Cedex - France

Tél.: 02 32 88 80 67 - mailto: Francis.Roussel@chu-rouen.fr

Introduction

La formation continue des médecins implique la lecture régulière de travaux scientifiques ; les médecins doivent, en effet, pouvoir mettre à jour leur documentation professionnelle, ce qui implique en particulier d'analyser et juger de façon autonome la valeur scientifique des articles publiés, ainsi que leur caractère applicable à d'autres situations.

L'enseignement de la lecture critique d'article (LCA) vise à compléter leur formation initiale dans ce domaine ; elle substitue partiellement l'acquisition d'un savoir (connaissances) par celle d'un savoir-faire (sélectionner soi-même des connaissances valides).

Cet enseignement, qui doit préparer les futurs médecins à leur formation continue, est différent de l'expertise scientifique requise en amont, pour sélectionner un manuscrit en vue d'une éventuelle publication scientifique¹ (processus éditorial).

Contexte

Une épreuve de LCA des étudiants est prévue dans le cadre de l'examen classant national (ECN) à partir de 2008², selon des modalités définies par le Centre national du concours de l'internat (CNCI). Cette nouvelle épreuve de trois heures, comptant pour 20 % des points, devrait concerner un article original, de portée générale, écrit en langue française ; elle devrait associer six à dix questions sur la méthodologie du travail, d'une part, et la rédaction d'un résumé de moins de 250 mots, d'autre part.

La plupart des facultés de médecine ont mis en place de façon empirique dès 2002 des formations à cette épreuve ; toutefois, ni les étudiants, ni les correcteurs ne sont entraînés à de telles épreuves nouvelles, alors qu'une telle expérience est nécessaire pour évaluer l'acquisition³ de cette habileté. L'organisation d'une épreuve de LCA dans un examen national à visée normative semble préoccupante dans un tel contexte, ce d'autant que la reproductibilité de la correction d'une épreuve de LCA n'est pas connue.

Le but de cette étude-pilote est d'apprécier la reproductibilité intercorrecteurs d'une correction d'épreuve de LCA organisée suivant les spécifications actuellement connues de cette épreuve, sans précaution méthodologique particulière, dans les conditions actuelles de formation des étudiants et de préparation des correcteurs, afin de repérer les difficultés potentielles de ce type d'épreuve.

Sujets et méthodes

Etudiants

A la faculté de médecine de Rouen, un enseignement des méthodes épidémiologiques est réalisé depuis 1998 en première année du deuxième cycle des études médicales (DCEM1) sous forme d'un autoapprentissage des principes de méthode, suivi de séances d'enseignement dirigé de 90 minutes, utilisant des articles originaux comme supports ; cet enseignement est validé par une épreuve terminale d'analyse d'article de 30 minutes, comportant 4 à 5 questions concernant un résumé d'article⁴. Les étudiants ont été sollicités par voie d'affichage et 59 volontaires ont passé l'épreuve en septembre 2002 : 25 et 34 étaient respectivement en deuxième et troisième années du DCEM, soit respectivement 25 et 31 % des inscrits des promotions concernées.

Ces étudiants ont été soumis à une épreuve dans les conditions prévues pour l'ECN, mais sans formation supplémentaire, en particulier pour la rédaction d'un résumé.

Dossier de LCA

Un article original a été choisi par un groupe local de cinq enseignants comme support à une épreuve de LCA : il s'agissait d'un essai randomisé bicentrique comparant deux modes de dépistage de l'hépatite C en médecine générale, selon que les patients avaient reçu ou non une information préalable⁵. Un dossier comportant l'article (sans le résumé ni la discussion), et dix questions (*Tableau 1*), y compris les réponses-type attendues (*cf. annexe 1*) et un résumé type (*cf. annexe 2*), a été construit⁶ ; il a fait l'objet d'un travail commun des enseignants au cours de 5 réunions. Ce dossier a été ensuite validé par un groupe d'experts de LCA auprès du CNCI au niveau national.

Epreuve

L'épreuve a été organisée dans les conditions d'un concours : surveillance, absence d'échanges entre les étudiants, temps limité à 3 heures, remise d'une copie rendue anonyme. Les étudiants avaient pour consigne : a) de répondre aux questions posées concernant l'article ; b) de rédiger un résumé selon les standards admis de la rédaction scientifique. La consigne précisait explicitement que le résumé devait comprendre des paragraphes respectivement dédiés aux objectifs, aux méthodes, aux résultats et à la conclusion.

Tableau 1 :
Questions utilisées pour l'épreuve d'évaluation des aptitudes
à la lecture critique d'article

Items	Points maximum	
1- Pour cet essai randomisé visant à démontrer l'efficacité d'une intervention, précisez : _____	11	
		dont
a) la nature précise de l'intervention qu'on cherche à évaluer _____	3	
b) le ou les critères de jugements utilisés _____	5	
c) pourquoi y a t-il « deux unités d'analyse » _____	3	
2- Cet essai justifie-t-il un avis du CCPPRB? Justifiez votre réponse. _____	7	
3- Pourquoi a-t-on stratifié l'essai selon le mode d'exercice? _____	5	
4- Les auteurs considèrent que le résultat de l'essai est négatif : quelles raisons pourraient, en théorie, expliquer ce résultat? _____	9	
5- Quel est l'effet du sexe masculin tel qu'il résulte des tableaux 4 et 5? _____	5	
6- D'après les résultats de cette étude, en quoi les toxicomanes se distinguent-ils des patients à ALAT > N vis-à-vis du dépistage de l'hépatite C? _____	7	
7- Les résultats obtenus dans cet essai vous paraissent-ils applicables à l'ensemble des médecins généralistes en France? Justifiez votre réponse. _____	5	
8- Les résultats obtenus dans cet essai vous paraissent-ils applicables à l'ensemble de la population française adulte? Justifiez votre réponse. _____	5	
9- D'après les résultats dont vous disposez, peut-on comparer la clientèle des deux groupes de médecins? Justifiez votre réponse. _____	8	
10- Proposez un titre alternatif pour cet article _____	8	
Résumé _____	30	
		dont
Objectifs _____	5	
Méthodes _____	10	
Résultats _____	10	
Conclusion _____	5	

CCPPRB : Comité consultatif de protection des personnes dans la recherche biomédicale

ALAT : alanine amino transférase

Tableau 2 :
Evaluation de la fidélité inter correcteur au cours de la correction de l'épreuve de lecture critique d'article

Coefficients de corrélation intra-classe obtenus par 3 paires de correcteurs pour 59 étudiants selon les items, avec leurs intervalles de confiance à 95 %.

	Paires de correcteurs		
	XY	XZ	YZ
Questions			
Q1a	0,84 (0,75 - 0,90)	0,21 (-0,04 - 0,44)	0,19 (-0,06 - 0,42)
Q1b	0,66 (0,49 - 0,78)	0,30 (0,05 - 0,51)	0,32 (0,08 - 0,53)
Q1c	0,31 (0,06 - 0,52)	0,25 (0,00 - 0,47)	0,00 (-0,25 - 0,25)
Q2	0,90 (0,84 - 0,94)	0,88 (0,80 - 0,93)	0,87 (0,79 - 0,92)
Q3	0,84 (0,75 - 0,90)	0,21 (-0,04 - 0,44)	0,19 (-0,06 - 0,42)
Q4	0,52 (0,31 - 0,68)	0,38 (0,14 - 0,58)	0,33 (0,08 - 0,54)
Q5	0,68 (0,52 - 0,80)	0,44 (0,21 - 0,62)	0,36 (0,12 - 0,56)
Q6	0,60 (0,41 - 0,74)	0,38 (0,14 - 0,58)	0,56 (0,36 - 0,71)
Q7	0,82 (0,71 - 0,89)	0,68 (0,52 - 0,80)	0,56 (0,36 - 0,71)
Q8	0,69 (0,53 - 0,80)	0,62 (0,44 - 0,76)	0,56 (0,36 - 0,71)
Q9	0,74 (0,60 - 0,84)	0,34 (0,10 - 0,55)	0,10 (-0,15 - 0,35)
Q10	0,63 (0,44 - 0,76)	0,60 (0,40 - 0,74)	0,45 (0,22 - 0,63)
Sous-total	0,82 (0,65 - 0,86)	0,57 (0,32 - 0,69)	0,32 (0,03 - 0,49)
Résumé			
R1	0,72 (0,57 - 0,82)	0,49 (0,27 - 0,66)	0,57 (0,37 - 0,72)
R2	0,37 (0,13 - 0,57)	0,40 (0,17 - 0,59)	0,56 (0,36 - 0,71)
R3	0,46 (0,23 - 0,64)	0,47 (0,24 - 0,64)	0,50 (0,29 - 0,67)
R4	-0,07 (-0,31 - 0,19)	0,11 (-0,15 - 0,35)	0,45 (0,22 - 0,63)
Sous-total	0,29 (0,27 - 0,66)	0,28 (0,25 - 0,65)	0,62 (0,57 - 0,82)
Ensemble de l'épreuve	0,70 (0,52 - 0,80)	0,72 (0,55 - 0,81)	0,48 (0,26 - 0,65)

Q 1 à 10 : questions 1 à 10 . R1 à 4 : rédaction des 4 paragraphes du résumé

Correction

Trois correcteurs (X, Y et Z), enseignants à la faculté, appartenant au groupe ayant conçu le dossier, ont corrigé chacun les 59 copies de façon indépendante, à l'aide d'un barème sur 100 points, répartis entre les dix questions et le résumé (*tableau 1*).

Analyse statistique

Les scores portés par chaque correcteur à chaque étudiant ont fait l'objet d'une double saisie. Les scores moyens des 3 correcteurs aux mêmes copies pour les questions, d'une part, et pour le résumé, d'autre part, ont été comparés par un test t apparié. Les coefficients de corrélation intra-classes⁷ (CCI) et leur intervalle de confiance à 95 % (IC 95 %) ont été calculés pour les trois couples de correcteurs pour chaque question, pour l'ensemble des questions, pour le résumé et enfin pour l'ensemble de l'épreuve. Les CCI ont été comparés par étude du recouvrement des IC 95 %. Le niveau de concordance a été qualifié selon les références admises. La moyenne des CCI par correcteur a été calculée. La relation entre la reproductibilité des corrections et le niveau des copies a été testée par un coefficient de corrélation entre la moyenne des CCI et celle des scores obtenus par les étudiants.

Résultats

Le scores moyen (\pm écart type) sur 20 donné par les 3 correcteurs à l'ensemble de l'épreuve aux 59 étudiants était de 8,7 (\pm 2,0) avec une dispersion modérée (minimum 4,5 ; maximum : 14,6). Le scores moyen était de 8,2 (\pm 2,2) pour les questions et de 9,9 (\pm 2,6) pour le résumé. L'écart moyen entre les scores du résumé et des questions de chaque étudiant était de 1,7 (\pm 2,6), écart supérieur à zéro ($p < 10^{-6}$).

Les CCI des questions, du résumé et de l'ensemble des épreuves figurent dans le *Tableau 2* pour les 3 paires de correcteurs. Pour l'ensemble de l'épreuve, les résultats variaient de 0,48 (IC 95 % : 0,26 – 0,65) pour la paire YZ à 0,72 (IC 95 % : 0,55 – 0,81) pour la paire XZ.

Sur les 48 CCI (16 items pour 3 paires), aucun n'était très bon, 8 étaient bons, 14 modérément bons, 16 médiocres et 10 mauvais selon les qualificatifs usuels.

La correction des questions était plus reproductible que celle des résumés pour 2 des 3 paires de correcteurs.

La reproductibilité (moyenne des CCI des trois paires de correcteurs) n'était pas significativement liée avec le score moyen obtenu par les étudiants ($r = - 0,35$).

Discussion

Ce travail pilote comporte des limites notables : les étudiants étaient volontaires, issus d'une seule faculté et peu nombreux ; ils n'étaient concernés par l'ECN que 2 à 3 ans plus tard, délai permettant des apprentissages complémentaires ; leur habilité⁸, telle qu'elle a été évaluée dans ce travail, ne correspond pas à celle correspondant à la date prévue pour eux pour cet examen. L'absence d'enjeu pour eux modifiait les conditions de l'épreuve et peut-être leur comportement. L'analyse ne portait que sur un seul article, corrigé par des enseignants d'une seule faculté, la plupart n'ayant pas une expérience notable de corrections d'épreuves de LCA.

Ces étudiants n'avaient pas encore fait l'objet d'une préparation spécifique de la LCA en vue de l'ECN, mais ils avaient reçu une formation méthodologique en DCEM1, formation s'appuyant en bonne partie sur l'analyse d'articles scientifiques. Toutefois, cette situation correspond probablement à celle d'autres facultés en France actuellement ; elle est donc opportune pour juger des performances qui pourraient être associées à une telle épreuve, sans précautions particulières.

Dans ces conditions, ce groupe d'étudiants de 2^e cycle n'a pas obtenu la moyenne à cette épreuve de LCA respectant les conditions définies par le CNCI ; l'épreuve est donc apparue difficile. Le score moyen obtenu au résumé était paradoxalement meilleur que pour les questions, alors que les étudiants n'avaient pas été formés spécifiquement à cet exercice⁹.

L'épreuve est également apparue peu discriminante, d'après la dispersion des scores. En outre, sa correction a été difficile, comme le suggère la faible reproductibilité inter-juges. La reproductibilité du processus de correction pouvait être influencée par l'épreuve elle-même, par les items qui la composaient, par les étudiants ou par les évaluateurs. La valeur moyenne des CCI des questions et du résumé étaient proches, correspondant à un accord inter-juges modéré ; ce résultat médiocre n'était pas le fait de l'une ou l'autre des deux parties de l'épreuve. La reproductibilité de la correction variait fortement selon les items : elle était médiocre ou mauvaise pour plus de la moitié d'entre eux, mais surtout la dispersion des CCI était importante (extrêmes : - 0,07 à 0,90).

Certaines difficultés de correction imprévues ont été rencontrées pour certains items, mettant en difficulté les trois correcteurs (Q1c et conclusion du résumé) ; on peut s'interroger sur le libellé de telles questions, voire leur contenu réel.

Il apparaît légitime de n'utiliser en situation de concours que des questions ayant été testées en conditions réelles, non seulement en terme de difficulté de réponse, mais encore en terme de fiabilité de correction.

La reproductibilité de la correction n'était pas liée avec les scores des étudiants, bien que ce résultat n'ait été établi que sur un effectif restreint. Ce point suggère que cette reproductibilité n'était pas dépendante du niveau des étudiants. Cependant, les étudiants n'avaient probablement pas toutes les habiletés requises dans le cadre de référence⁸, comme celle d'évaluer la pertinence de l'article, la validité des résultats, leur utilité pratique. Ils n'avaient eu aucun apprentissage de la rédaction d'un résumé, alors que l'entraînement améliore les résultats¹⁰.

La médiocre reproductibilité des corrections peut enfin provenir des correcteurs eux-mêmes, bien que ceux-ci aient directement contribué à l'élaboration du dossier, en particulier des réponses-type ; ce point suggère l'importance d'utiliser une grille de correction pour ce type d'épreuve ; il serait intéressant de vérifier qu'une telle grille permettrait bien de réduire la variabilité inter-correcteur. L'épreuve de LCA, telle qu'elle est envisagée actuellement, est restreinte aux articles originaux ; cependant, le résumé, dont ils ne devraient pas disposer dans les conditions prévues pour l'ECN, joue un rôle important pour la formation des médecins en exercice ; en outre, ces derniers sont fréquemment confrontés à d'autres types de publication, comme des revues systématiques ou des recommandations de bonne pratique ; inversement, la rédaction d'un résumé en vu d'une publication n'est pas une activité habituelle pour un médecin en exercice ; les habiletés jugées nécessaires pour la pratique professionnelle ultérieure ne sont pas donc appréciées sur des bases très réalistes ; cependant les avantages et inconvénients des différents constituants prévus pour l'ECN (dossiers, LCA) ont été peu étudiés. Plus généralement, d'autres types d'épreuves que la LCA peuvent être envisagés pour pallier les inconvénients des questions rédactionnelles à développements longs¹¹.

Conclusion

Cette étude pilote suggère la difficulté d'une épreuve de LCA pour les étudiants, la médiocrité de sa reproductibilité inter correcteurs et souligne l'importance de précautions à prendre pour la correction, comme l'utilisation d'une grille de correction et une correction multiple.

Si une telle épreuve doit classer plus de 5 000 candidats au niveau national, une formation spécifique des étudiants et des enseignants, une préparation et un test des épreuves et de leur correction par un travail transversal au sein de toutes les facultés semble indispensable dans un souci d'équité et d'éthique pédagogique.

Remerciements

Les auteurs remercient les 59 étudiants qui se sont prêtés au test, ainsi que M. Richard Medeiros qui a revu la version anglaise du résumé.

Contributions

Francis Roussel et Pierre Czernichow ont participé à l'élaboration du dossier utilisé pour l'étude, corrigé les épreuves, réalisé l'analyse des données de l'étude et rédigé une première version du manuscrit ; ils ont effectué la révision du manuscrit à partir des critiques des lecteurs. Alain Lavoine, Jean-François Lemeland ont participé à l'élaboration du dossier utilisé pour l'étude, et contribué à la relecture et l'élaboration de la version du manuscrit qui a été soumise à la revue. Jean-Paul Fillastre a participé à l'élaboration du dossier utilisé pour l'étude, corrigé les épreuves, et contribué à la relecture et l'élaboration de la version du manuscrit qui a été soumise à la revue.

**Annexe 1 :
Réponses type utilisées pour la correction des 59 copies par les correcteurs**

- 1a. Information des patients par affiche et dépliants en salle d'attente.
- 1b. Pourcentage de médecins prescripteurs, pourcentage de sérologies prescrites, pourcentages des sérologies demandées par les malades, pourcentages des sérologies effectuées, pourcentages des sérologies positives.
- 1c. Parce qu'il s'agit d'un essai avec randomisation de groupes (clientèles de médecins) et non de patients : l'analyse porte sur les médecins (prescripteurs ou pas) et sur les patients auxquels une sérologie a été prescrite.
2. Non. Pas d'intervention faisant courir un risque particulier aux patients ; l'offre d'une sérologie fait partie des soins de routine. Il existe un bénéfice direct.
3. On a pensé que le mode d'exercice pouvait influencer le critère de jugement. On a pensé que le mode d'exercice pouvait influencer le recrutement. On a voulu s'assurer qu'il ne perturberait pas l'analyse (deux tirages au sort sont effectués : un dans chaque strate, un chez les médecins exerçant seul, un autre chez ceux exerçant en groupe). Ou bien : le mode d'exercice a été considéré comme un facteur de confusion possible.
4. Effectifs insuffisants (défaut de puissance) ; biais dans la sélection des patients (répartition différente des groupes à risque dans les clientèles) ; comportement particulier des toxicomanes (12 sérologies positives mathématiquement manquantes pour 25 trouvées au total) ; biais dans la mesure des prescriptions des sérologies (données manquantes, perdus de vue) ; inefficacité de l'intervention si les autres raisons sont éliminées.
5. Les hommes réalisent 2 fois plus souvent la sérologie prescrite que les femmes, tous facteurs égaux par ailleurs, alors qu'il n'y a pas de lien entre le sexe et la positivité de la sérologie.
6. Les toxicomanes réalisent moins souvent la sérologie (OR = 0,04) que les patients ALAT > N, tous facteurs égaux par ailleurs, alors qu'ils ont un risque supérieur d'être positifs (OR = 19,8).
7. Non. Les médecins sont volontaires, donc non représentatifs des 2 régions concernées. Leur fréquence est d'ailleurs très différente. Ils proviennent de 2 régions particulières.
8. Non. Il s'agit de consultants de médecins généralistes, donc il existe une sélection par le recours aux soins. Ils proviennent de 2 régions particulières.
9. Non. Seules les caractéristiques des patients ayant fait l'objet d'une prescription de sérologie ont fait l'objet d'un recueil de données : on peut donc comparer : les caractéristiques des médecins, les caractéristiques des malades ayant eu une prescription de sérologie ; mais on ne peut pas comparer les caractéristiques de l'ensemble des malades.
10. Efficacité d'une information en salle d'attente sur le dépistage de l'hépatite C : essai randomisé chez 184 médecins généralistes du Doubs et de l'Île de France.

Annexe II :

Résumé utilisé pour la correction des 59 copies par les correcteurs (248 mots)

Objectif : Evaluer l'efficacité d'une information par affiche et dépliants en salle d'attente de médecine générale sur le dépistage de l'hépatite C.

Méthodes : 184 médecins volontaires du Doubs (n=119) et d'Ile de France (n=65) ont été répartis par tirage au sort stratifié sur le mode d'exercice (isolé ou en groupe) : groupe 1 (n=90) : recherche de facteur de risque d'hépatite C et offre éventuelle d'une sérologie sur l'initiative du médecin ; groupe 2 (n= 94): id et information préalable sur l'hépatite C des patients (affiche et dépliants) en salle d'attente. Le pourcentage de prescripteurs et le nombre de sérologies ont été comparés respectivement par les tests du khi-carré et de Mann-Whitney. Enfin les facteurs associés à la réalisation des sérologies prescrites, et à la séropositivité ont été recherchés par régression logistique.

Résultats : En 15 mois 55 médecins du groupe 1 (61%) et 48 du groupe 2 (51%) (non significatif) ont prescrit 323 et 294 sérologies respectivement (non significatif). Les patients du groupe 2 ayant eu une sérologie étaient plus âgés ; leurs facteurs de risque étaient similaires mais plus nombreux. La réalisation des sérologies prescrites était plus fréquente chez les 60 ans et plus, et les hommes, moins fréquente chez les toxicomanes, et ceux ayant un tatouage. La séropositivité n'était liée qu'aux patients toxicomanes. Son contrôle était plus fréquemment demandé par les patients informés, mais pas plus fréquemment réalisé et pas plus fréquemment positif.

Conclusion : L'information n'a pas amélioré le rendement du dépistage de l'hépatite C.

Références

1. Fillastre JP, Colin R. *Analyse critique ou lecture critique des articles médicaux. Quelle cible choisir pour l'enseignement et l'évaluation? Pédagogie Médicale 2001 ; 2 : 197-198.*
2. *Conseil scientifique du centre national du concours d'internat. Epreuve écrite de lecture d'article. 2003 [on-line]. Disponible sur : <http://www.cnci.univ-paris5.fr/medecine/CritiqueArticle.doc>*
3. Jouquan J. *L'évaluation des apprentissages des étudiants en formation médicale initiale. Pédagogie Médicale 2002 ; 3 : 38-52.*
4. Tavoracci MP, Ladner J, Kerdiles M, Hallais C, Josset V, Merle V et al. *Un nouveau dispositif de formation en santé publique dans les études médicales : mise en place et mesures d'impact. Pédagogie Médicale 2005 ; 6 : 15-24.*
5. Roudot-Thoraval F, Monnet E, Mercet P, Bastie A, Dhumeaux D, Miguet JP. *Stratégies de dépistage de l'hépatite virale C en Médecine générale. Résultats d'un essai randomisé bicentrique. Gastroenterol Clin Biol. 2000 ; 24 : 1037-1041.*
6. Beaucage C, Viger Y. *Epidémiologie appliquée : une initiation à la lecture critique de la littérature en sciences de la santé. Boucherville (QC) : Gaëtan Morin éditeur, 1996.*
7. Fermanian J. *Mesure de l'accord entre deux juges: cas quantitatif. Rev Epidem et santé publ 1984 ; 32 : 408-413.*
8. Coté L, Turgeon L. *Comment lire de façon critique les articles de recherche qualitative en médecine? Pédagogie Médicale 2002 ; 3 : 81-90.*
9. Audet N, Leclère H. *Les habilités requises pour la lecture critique en médecine : un cadre de référence issu d'une recension des écrits. Pédagogie Médicale 2001 ; 2 : 206-212.*
10. Taylor R, Reeves B, Ewings P, Binnes S, Keast J, Mears R. *A systematic review of the effectiveness of critical appraisal skills training for clinicians. Med Educ 2000 ; 34 : 120-125.*
11. Stern DT, Linzer M, O'Sullivan PS, Weld L. *Evaluating medical residents' literature-appraisal skills. Acad Med 1995 ; 70 : 152-154.*

Manuscrit reçu le 11 mars 2004 ; commentaires éditoriaux formulés aux auteurs le 5 octobre 2004 ; accepté pour publication le 21 décembre 2004.