

EVALUATION D'UNE GRILLE DE CRITERES DE QUALITE DE RESSOURCES PEDAGOGIQUES DISPONIBLES SUR L'INTERNET

A. Hamon, S.J. Darmoni, F. Le Duff, J.C. Dufour, M. Joubert, P. Le Beux, M. Fieschi, J. Weber et J. Bénichou.

Unité de Biostatistique, Pavillon Pillore, Hôpital Charles Nicolle, CHU de Rouen, 76031 Rouen Cedex

Contexte :

L'évaluation de la qualité de l'information médicale délivrée par Internet est actuellement devenue une nécessité. Dans le cadre du projet *Université Médicale Virtuelle Francophone*, une grille de neuf critères a été développée afin de permettre de juger de la qualité de sites Internet proposant des documents médicaux pédagogiques.

Objectif :

Le but de cette étude est d'évaluer la concordance des jugements émis par trois observateurs différents sur la base de ces neuf critères.

Matériel :

Trois informaticiens médicaux ont choisi, sur la base de codes éthiques européens et américains déjà validés, les neuf critères suivants : Source, Existence d'un comité éditorial, Citations des sources originales, Cible du site, Mise à jour, Pertinence des hyper-liens, Navigabilité, Design et Mécanismes pour la rétroaction. Pour chaque critère une grille a été élaborée. Quatre niveaux de qualité (très bon, bon, mauvais, très mauvais) ont été retenus. Vingt quatre sites francophones proposant des documents pédagogiques ont été sélectionnés en France, Belgique, Suisse et Canada. Chaque site a alors été évalué de façon indépendante par chacun des trois observateurs.

Analyse statistique :

Pour chaque critère, nous avons tout d'abord cherché à établir s'il existait un accord substantiel pour chaque paire d'observateurs. Plusieurs indices statistiques ont été utilisés : le coefficient Kappa, le coefficient de corrélation de Spearman et la statistique Gamma. Puis l'étude de l'accord global entre les trois observateurs a été effectuée en estimant deux autres indices : le coefficient Kappa généralisé et la statistique W de Kendall.

Résultats :

Les différentes approches utilisées pour juger de l'accord observé nous ont permis de dégager trois groupes parmi les neuf critères. Un premier groupe, contenant deux critères (Cible du site et Mécanismes pour la rétroaction), est caractérisé par un très bon accord, tant entre les observateurs deux à deux que sur l'ensemble des trois. Pour trois autres critères (Citations des sources originales, Navigabilité et Design), l'accord reste assez élevé, bien que les classements de l'un des trois observateurs ne soient pas tout à fait en accord avec ceux des deux autres. Enfin pour les quatre derniers critères (Source, Existence d'un comité éditorial, Mise à jour et Pertinence des hyper-liens) l'accord entre les juges est très faible.

Conclusion :

Cette étude montre qu'il est possible d'utiliser les cinq critères montrant un accord substantiel (Citations des sources originales, Cible du site, Navigabilité, Design et Mécanismes pour la rétroaction) pour l'évaluation de la qualité de sites Web. Pour les quatre autres critères (Source, Existence d'un comité éditorial, Mise à jour et Pertinence des hyper-liens) de nouveaux travaux sont nécessaires afin d'analyser les divergences entre observateurs et modifier leurs définitions en conséquence.