

QUESTION OU DILEMME PÉDAGOGIQUE

Les systèmes de notation METRIQ-8 et ALiEM AIR pour la qualité des publications de blogues en médecine d'urgence présentent des preuves de validité interne, mais n'ont pas été testés à l'externe dans une population générale de cliniciens et de stagiaires.

Référence

Brent Thoma, Stefanie S. Sebok-Syer, Isabelle Colmers-Gray, Jonathan Sherbino, Felix Ankel, N. Seth Trueger, Andrew Grock, Marshall Siemens, Michael Paddock, Eve Purdy, William Kenneth Milne, Teresa M. Chan & the METRIQ Study Collaborators (2018). Quality Evaluation Scores are no more Reliable than Gestalt in Evaluating the Quality of Emergency Medicine Blogs: A METRIQ Study. *Teaching and Learning in Medicine*, 30:3, 294-302. DOI: 10.1080/10401334.2017.1414609

En Quoi Est-Ce Pertinent Pour La Formation En Médecine d'Urgence ?

L'érudition éducative a considérablement changé depuis l'ère numérique. Alors que les résidents en médecine d'urgence (MU) utilisent de plus en plus l'accès libre et gratuit à l'éducation médicale (Free Open Access Medical Education—FOAM), des inquiétudes ont été soulevées quant à la qualité et le niveau de la révision par les pairs de ces ressources. Déterminer des systèmes de notation robustes pour les publications de blogues en MU peut aider les médecins d'urgence à assurer que le contenu qu'ils lisent est de qualité éducative supérieure.

Niveau de Preuve

Niveau de preuve III ou IV

Niveau d'Apprentissage

Évaluation

Conception de l'Étude

Une étude transversale sur la notation de publications de blogues au moyen d'un échantillon de commodité de 309 participants (121 étudiants en médecine, 88 résidents en MU et 100 assistants en MU), recueilli au moyen d'une méthodologie multimodale au sein d'une communauté de pratique virtuelle.

Sources de Financement

Collège royal des médecins et chirurgiens du Canada
Association canadienne des médecins d'urgence

Contexte

Les formulaires d'inscription et les sondages ont été distribués en ligne et complétés par des participants dans 27 pays différents. 83,1 % des répondants provenaient du Canada ou des États-Unis.

Résumé de l'Étude

309 étudiants en médecine, résidents en médecine d'urgence (MU) et assistants en MU ont été recrutés au moyen d'un échantillonnage de commodité via des pages FOAM ou lors d'une conférence. Ils ont ensuite été enseignés, au moyen de vidéos d'une minute, comment noter des blogues médicaux avec les outils de

Résumé de l'Étude (continué)

notation METRIQ-8 et ALiEM AIR. Vingt publications récentes de blogues ont été sélectionnées aléatoirement pour une évaluation avec ces deux outils, ainsi qu'une évaluation de la qualité au moyen de la « gestalt » de l'évaluateur. Puisqu'il a été soulevé que la fatigue pourrait possiblement fausser les résultats pendant les 90-120 minutes requises pour évaluer l'ensemble des vingt publications, les organisateurs de l'étude ont également randomisés par bloc l'ordre des publications de blogue à évaluer. De plus, ils ont présenté les outils de notation un à la fois, évaluant cinq publications de blogues avec la gestalt et METRIQ-8 ou AIR, ensuite cinq avec la gestalt et l'autre outil, ensuite les dix derniers avec les trois méthodes. Ceci était pour assurer que les utilisateurs n'étaient pas dépassés au début avec l'utilisation des deux nouveaux outils simultanément.

Les coefficients de corrélation de Pearson ont été calculés pour mesurer le degré de corrélation entre les résultats METRIQ-8, AIR et de la gestalt. La fiabilité au sein des groupes ainsi que les différents types de répondants (étudiant, résident ou assistant) ont également été explorés. Les corrélations de Pearson étaient fortes entre les résultats des deux outils et de la gestalt ($>0,90$) qui maintenaient de la force, lorsque décomposés par type de répondant ($>0,87$), signifiant que les résultats de la gestalt étaient semblables aux résultats des outils de notation, peu importe le niveau de formation médicale. De plus, la corrélation intraclasse était juste (0,3-0,4), signifiant que les étudiants, les résidents et les assistants ont tous évalué les publications de blogues de manière semblable au sein de leurs groupes.

CONCLUSION

Les instruments structurés comme les outils de notation METRIQ-8 et ALiEM AIR ne réduisent pas la variabilité de l'évaluation de la qualité des publications de blogues médicaux comparativement à la gestalt. Cependant, la forte corrélation avec la gestalt signifie que des impressions de qualité peuvent être systématiquement classées pour aider à appuyer les comparaisons et les discussions entre les évaluateurs.