

# Degré d'importance des articles : Les processus d'évaluation des articles publiés sont subjectifs

Les opinions exprimées dans cet éditorial sont celles de l'auteur et ne représentent pas nécessairement celles de l'Association médicale canadienne ou ses filiales.

**M**algré la complexité des formules qui visent à déterminer la qualité des publications médicales, le processus d'évaluation est fondamentalement subjectif. Pour cet éditorial, nous avons cherché sans succès des corrélations entre les scores d'importance attribués par des experts et le taux de citation enregistré sur une durée de 20 ans pour des articles tests provenant de deux revues réputées. D'autres méthodes d'évaluation de l'importance sont de plus en plus utilisées. À l'instar des agriculteurs au marché, les chercheurs devront désormais faire la promotion de leurs produits après avoir fait tout en leur pouvoir pour qu'ils soient de la meilleure qualité possible.

On a beau mettre au point des protocoles pour améliorer la recherche médicale, la détermination de l'importance d'un rapport de recherche repose toujours également sur l'intuition et le respect des lignes directrices<sup>1</sup>. Ce problème touche les articles tant avant leur publication qu'après. Aucun rédacteur en chef ne souhaite rejeter un rapport pour les mauvaises raisons, comme ce fut le cas pour l'article soumis à la revue *Science* par Bruce Glick et Timothy Chang, étudiants aux cycles supérieurs à l'Ohio State University, à la suite d'un imbroglio survenu dans leur laboratoire : Chang avait par erreur utilisé les poulets de Glick, dont les bourses de Fabricius avaient été excisées par Glick, pour enseigner aux étudiants comment faire augmenter la production d'anticorps, ce qui leur a permis de confirmer le rôle de cet organe dans la synthèse des anticorps. Le rédacteur en chef de *Science* n'a aucunement remis en question la méthodologie utilisée; il a plutôt refusé l'article sous prétexte que son contenu n'intéresserait pas le lectorat de la revue. Sans y réfléchir davantage, les auteurs ont respecté l'avis du rédacteur et ont publié leur article dans *Poultry Science*<sup>2</sup>.

Il y a deux décennies, plusieurs revues de renom ont élaboré des politiques pour la publication accélérée de travaux jugés particulièrement importants. Les auteurs et les pairs évaluateurs pouvaient suggérer la publication accélérée de certains articles, mais le dernier mot revenait aux rédacteurs en chef. Cette compétence et cette responsabilité se devaient toutefois être testées. William Ghali, de l'Université de Calgary, et ses collaborateurs du Canada et de la Suisse ont relevé le défi<sup>3</sup>. Dans le cadre d'une expérience rigoureusement conçue, ils ont invité 42 experts de partout dans le monde à évaluer, dans cinq domaines,

12 articles jugés particulièrement importants et 12 articles témoins sélectionnés selon la revue, la pathologie ou l'intervention d'intérêt, le domaine général et l'année de publication. Bien que le score moyen accordé aux articles jugés importants ait été légèrement plus élevé, mais statistiquement significatif, que celui accordé aux autres articles, les experts ont choisi des articles témoins au lieu des articles jugés importants dans cinq des 12 paires, ce qui a poussé les auteurs de l'étude à conclure que le processus de sélection était incohérent.

Selon le rédacteur en chef de l'une des revues, les articles jugés importants avaient été cités deux fois plus que les articles témoins entre le moment où Ghali et ses collaborateurs avaient décidé de mener leur étude et celui où ils avaient publié leur rapport, ce qui justifiait leur sélection<sup>4</sup>. Le taux de citation semble être une mesure objective de l'importance d'un article, mais c'est le facteur d'impact (FI) qui est considéré comme la mesure la plus juste de la réussite d'une revue. Le FI, une formule brevetée qui appartient à Clarivate Analytics, est calculé en prenant le nombre

**Tableau 1. Évaluation par les experts et taux de citation des articles sélectionnés par les équipes de rédaction des revues pour publication accélérée, comparativement à des articles appariés publiés selon la méthode régulière**

Paire d'articles	Score des experts*		Nombre de citations 1998–2018†	
	Accélérée	Régulière	Accélérée	Régulière
A	6,5	6,7	131	402
B	7	5,1	645	185
C	4,1	4,2	973	107
D	5,6	5,8	35	42
E	5,5	5	522	453
F	6	4,8	161	101
G	7	7	465	365
H	7	7,2	972	804
I	6,3	5,5	1 457	773
J	6	5,5	256	245
K	5,1	6,1	359	328
L	8,2	5,4	4 894	556
Moyenne ± écart type	6,2 ± 1,1	5,7 ± 0,9	906 ± 1 323	363 ± 251
Test t jumelé bilatéral	$p = 0.13$		$p = 0.16$	

\*Les données sur les scores des experts proviennent de Ghali et collaborateurs<sup>3</sup>.

†Le nombre de citations provient d'une consultation de Web of Science, le 15 février 2018.

de citations, par année civile, des articles publiés par la revue au cours des deux années précédentes et en le divisant par le nombre d'articles citables publiés au cours des deux années précédentes. Le FI du *Journal canadien de chirurgie* (JCC) augmente lentement : il est passé de 0,5 en 2006 à 1,9 en 2016.

On peut utiliser les deux décennies écoulées depuis l'expérience de Ghali et de ses collaborateurs pour voir où se situent les articles sélectionnés par les équipes de rédaction par rapport aux articles témoins des experts (tableau 1). Le taux moyen de citation des articles sélectionnés est maintenant trois fois plus élevé que celui des articles témoins. Dans deux cas seulement, les articles témoins ont surpassé l'article sélectionné auquel ils avaient été jumelés. La variation du taux de citation des divers articles est considérable, et elle est plus grande que la différence entre les articles appariés. Aucune corrélation entre les scores des experts et le nombre de citations n'est apparente (corrélation de Spearman = 0,36;  $p = 0,08$ ). Des signaux qui portent à confusion semblent émerger des données. Les articles sélectionnés et les articles témoins sont appariés sur le plan de la qualité, ce dont témoigne le taux de citation. L'avantage dont profitent les articles sélectionnés pourrait justifier le jugement des rédacteurs en chef, ou être le reflet de la priorité accordée au moment de publier. Ce qui est remarquable, c'est la variabilité des taux de citation, qui semble témoigner de la valeur du contenu : la paire citée le moins souvent portait sur un problème touchant le tiers monde, et l'article cité le plus souvent, sur les problèmes cardiaques liés à une pilule amaigrissante<sup>3</sup>.

En 1955, le rédacteur en chef de la revue *Science* croyait qu'une étude sur les poulets n'intéresserait pas son lectorat. Il n'avait pas réalisé qu'il s'agissait là d'un moment décisif en immunologie, dont les retombées positives se font encore sentir aujourd'hui. Glick et Chang ont continué chacun de leur côté, confirmant tous deux leur découverte lors d'expériences complémentaires. Même si leur article demeure le plus souvent cité de la revue *Poultry Science*, leur importance dans la découverte des éléments fondamentaux de l'immunité humorale n'a jamais été reconnue. Glick est décédé en 2009 et a laissé un souvenir impérissable à d'innombrables étudiants de premier cycle et à

29 étudiants des cycles supérieurs en tant que mentor exemplaire. Même si la revue *Science* a accepté, en 1969, de publier l'un des quelque 200 articles scientifiques rédigés par Glick, le rejet de 1955 a coûté à ce dernier un prix Nobel.

André Picard, du *Globe and Mail*, a présenté le point de vue d'un journaliste médical sur le débat entourant les choix éditoriaux, et a affirmé que le problème ne touche pas strictement les milieux universitaires<sup>5</sup>. Les journalistes se réjouiront peut-être du fait que les publications savantes commencent à exploiter les médias sociaux. Comme la plupart des autres revues, le JCC se sert des médias sociaux pour promouvoir le travail des auteurs. L'Altmetrics (mesure d'impact alternatives) est un nouvel outil pour déterminer la réussite d'une revue dans ce contexte. Dans sa critique raisonnée des processus déterminant l'importance de la recherche dans la presse de tous les jours et le monde de l'édition savante, Picard reconnaît que « tous les médias cherchent la controverse »<sup>5</sup>. Il est clair que la controverse et la valeur du contenu ont joué un rôle substantiel dans le nombre de citations générées par les articles testés il y a deux décennies. À l'avenir, nous devons faire très attention pour éviter d'aggraver le biais lorsque les médias sociaux exercent leur pleine puissance.

**Vivian C. McAlister**

Corédacteur en chef, *Journal canadien de chirurgie*

**Intérêts concurrents** : Aucun déclaré.

DOI: 10.1503/cjs.004118

## Références

1. McAlister V. Toward a "New School" of surgical research. *Can J Surg* 2017;60:220.
2. Glick B, Chang TS, Jaap RG. *Poult Sci* 1956;35:224-5.
3. Ghali WA, Cornuz J, McAlister FA, et al. Accelerated publication versus usual publication in 2 leading medical journals. *CMAJ* 2002; 166:1137-43.
4. Stanbrook MB. Not fast enough? *CMAJ* 2002;167:738.
5. Picard A. Getting on track: how scientific journals and mainstream journalists could do a better job of communicating with the public. *CMAJ* 2002;166:1153-4.