

Rétractation

Retraction

P.-F. Dequin

© SRLF et Springer-Verlag France 2011

« (...) un bien acquis par fraude ne profite jamais longtemps. »

(*Sophocle, Œdipe à Colone*. Trad. Robert Pignarre.

In : Théâtre complet. Garnier-Flammarion, 2000, p 289)

Les faits

Dans le numéro de décembre 2009 du journal *Anesthesia and Analgesia* était publié un essai clinique de Boldt et al. [1]. Chez 50 patients de chirurgie cardiaque soumis à une circulation extracorporelle (CEC) avaient été comparées, par randomisation, deux solutions d'amorçage de la CEC : un hydroxyéthylamidon (HEA) de poids moléculaire moyen et de faible taux de substitution, dilué dans une solution électrolytique dite équilibrée, et de l'albumine à 5 %, diluée aux deux tiers dans du soluté salé isotonique. Les auteurs rapportaient au bénéfice de l'HEA, l'absence d'acidose induite, une moindre réponse inflammatoire, une meilleure préservation des paramètres de la crase sanguine, l'absence d'augmentation de marqueurs sériques infracliniques d'altération tubulaire rénale. Un éditorial accompagnait l'article [2], insistant sur les limites de l'étude, mais se plaçant aussi dans une perspective raisonnable d'utilisation des HEA de dernière génération dans cette indication, sous réserve, bien sûr, d'essais contrôlés de plus forte puissance.

Le 28 octobre 2010, Steven Shafer, éditeur en chef d'*Anesthesia and Analgesia*, publiait sur le site Internet du journal une note de rétractation de l'article de J. Boldt [3]. Le même texte était publié dans le numéro de décembre 2010 de l'édition papier du journal [4].

Que s'est-il passé ? Dans un premier temps, l'éditeur a reçu des lettres de lecteurs mettant en doute la réalité des dosages de cytokines rapportés, en raison de leur variabilité

trop faible ; la même remarque a été faite par les éditeurs quant à la variabilité de l'excès de base. En mai 2010, une demande a été adressée à la chambre fédérale des médecins de Rhénanie-Palatinat, organisme compétent en matière d'éthique pour l'institution de J. Boldt. En octobre, cet organisme a rapporté ses premières conclusions : contrairement à ce qui était mentionné dans l'article, le travail ne lui avait pas été soumis ; aucun consentement écrit n'a été recherché auprès des patients ; aucune randomisation n'a eu lieu et aucun questionnaire de suivi à distance n'a été utilisé. C'est sur ces bases qu'a été prise la décision de rétractation.

Simple (mais grave) manquement à l'éthique ou invention de données ?

L'enquête se poursuit, car un doute sérieux persiste sur la réalité même de l'étude, J. Boldt n'ayant pu fournir aucun des documents sources habituels [3,4]. Il a été exclu le 25 novembre de l'hôpital de Ludwigshafen [5], et son titre professoral est remis en cause. La société allemande d'anesthésie et de soins intensifs a publié un communiqué condamnant sévèrement son comportement [6]. L'incapacité d'un chercheur à fournir ses documents sources a souvent été un argument majeur pour conclure à une fraude [7-12] ou la juger très probable [13], quel qu'en soit le motif invoqué : dossiers égarés lors d'un déménagement, effacement du disque dur, intervention de termites... Les exemples de cohortes complètement inventées ne manquent pas [10,11].

Conséquences pour la communauté scientifique

Des trois niveaux d'action dont dispose un éditeur pour revenir sur un article publié (correction, avertissement sur des doutes ou *expression of concern*, rétractation), le dernier signifie très directement au lecteur de ne pas tenir compte de cet article et de ne pas le citer [14]. À l'époque où n'existaient que les versions papiers des journaux, l'information

P.-F. Dequin (✉)

Service de réanimation médicale polyvalente,
université François-Rabelais,
CHU Bretonneau, F-37044 Tours cedex 09, France
e-mail : dequin@med.univ-tours.fr

sur la rétractation ne circulait qu'imparfaitement, et on a montré que certains articles rétractés continuaient à être cités [15]. Sur les versions électroniques de ces articles, il est maintenant facile d'ajouter de façon très visible la mention *article retracted*. Le délai médian de rétractation dans les cas de fraude est de trois ans [16] : il n'a fallu ici que dix mois ; sans doute *Anesthesia and Analgesia* a-t-il été très sensibilisé par la récente affaire Reuben [11].

Bien évidemment, un doute existe désormais sur la valeur scientifique des autres travaux de J. Boldt. L'enquête aboutit parfois à la rétractation de la majorité des articles d'un auteur convaincu de fraude (Darsee, Reuben, Schön). Parfois, seul un article est formellement rétracté, malgré des doutes très sérieux quant à la probité des autres [17], soit que certains éditeurs ou institutions ne puissent ou ne veuillent diligenter d'enquête, soit parce que certaines enquêtes n'aboutissent que difficilement : il a fallu 13 ans d'enquête au *British Medical Journal* pour évoquer la falsification des données d'un article de Singh, et ses travaux n'ont pas été formellement rétractés [13]. Il est bien sûr théoriquement possible que, dans le parcours scientifique d'un chercheur, la fraude ne soit qu'un accident tardif ou isolé. Mais si la démarche scientifique implique le débat contradictoire sur l'interprétation des résultats ou la discussion des études, elle repose aussi implicitement sur la confiance en la réalité des faits et données rapportés. Même rare (4 pour 1000 des essais cliniques publiés en langue anglaise en 2005 ont été rétractés, et pas tous pour fraude), la fraude dûment constatée ne peut qu'ébranler cette confiance. Elle atteint à ce titre toute la communauté scientifique ; son irruption dans le débat public, à une époque où de façon très irrationnelle la confiance des sociétés vis-à-vis de toute institution est facilement remise en question, pourrait avoir une influence très négative sur notre capacité à réaliser des essais cliniques.

Au-delà de ces considérations générales, une grande partie du corpus scientifique sur l'utilisation en réanimation des HEA (58 des 200 articles de J. Boldt) et notamment des études réfutant leur toxicité rénale est entachée de doute. Aucune conclusion ne sera possible sans enquête approfondie [17] et sans nouvelles études fiables ; mais c'est indiscutablement une atteinte grave pour tous les chercheurs impliqués dans le domaine.

Événement rare (mais récurrent)

Même si elles font beaucoup parler d'elles, les fraudes caractérisées demeurent rares. Elles touchent tous les domaines scientifiques, de l'anthropologie à la géologie en passant par la physique des nanoparticules et l'immunologie. La médecine n'est malheureusement pas épargnée. Les motivations des fraudeurs sont connues : elles peuvent être financières, soit pour l'industriel promoteur [18], soit pour

l'auteur qui tire des revenus de son statut d'expert en partie frauduleusement acquis [11]. La publication transparente des conflits d'intérêts est de ce fait une exigence justifiée. La volonté de convaincre est également une motivation connue : il est ainsi admis que Gregor Mendel a modifié ses résultats expérimentaux pour convaincre de la justesse de sa théorie génétique ; celle-ci n'est pourtant pas remise en cause [19]. Il est possible aussi que Cyril Burt ait réalisé avec rigueur sa première publication, les suivantes n'en étant que la reproduction avec une population plus vaste mais inventée [20]. Mais la motivation sans doute la plus fréquente est le *publish or perish* qui avait été invoqué lors de l'affaire Darsee [7] ; elle aboutit à confondre l'ombre et la proie, l'article publié — à quel prix ! — et l'exigence d'une recherche réalisée avec rigueur. Elle est sans doute aussi la rançon d'une évaluation encore trop quantitative de la recherche.

Comment tombent les fraudeurs ?

Si C. Burt publiait seul [20] et si E. Alsabti s'inventait parfois des coauteurs [21], la plupart des travaux publiés sont affaires d'équipes réelles. L'association de nombreux coauteurs a pu être une technique de facilitation de la fraude [22], mais expose bien sûr à la rétractation d'un coauteur plus sourcilieux que les autres comme dans l'affaire Hwang [23,24]. C'est aussi rappeler les règles fondamentales d'*autorship* : ne pas accepter de signer un article qu'on ne peut pas revendiquer en totalité, au risque de devoir se rétracter individuellement aux premiers doutes publics, comme dans l'affaire Sudbø [10,25] ; pour un directeur d'équipe, ne pas valider sans vérification les articles soumis par un membre de l'équipe, comme le fit E. Braunwald avec J. Darsee [7] ou B. Batlogg avec H. Schön [12]. Pour le lecteur, se méfier des essais d'envergure publiés, de façon très atypique, par un seul auteur (Chandra, Alsabti) et des auteurs extraordinairement prolifiques (Alsabti, Schön, Singh, Chandra...).

Certaines impossibilités matérielles sont assez facilement repérables : Breuning ou Cruz, par exemple, furent suspectés de fraude, car le nombre de patients inclus pendant une période donnée dépassaient largement le recrutement de leurs institutions [8,26] ; Sudbø parce que le responsable de la cohorte du Norwegian Radium Hospital savait qu'elle ne comportait pas, loin s'en faut, les 454 patients cancéreux prétendument inclus dans son étude [25].

Les données numériques et l'étude statistique sont un piège à fraudeurs : les coefficients de corrélation de Burt étaient identiques à la décimale au fur et à mesure que sa cohorte s'accroissait [20] ; l'analyse par des biostatisticiens des données rapportées par Chandra prouva assez rapidement leur invraisemblance [9] ; la trop faible variabilité des dosages cytokiniques de Boldt a lancé le processus

aboutissant à la rétractation [1,4]. Des tests statistiques permettent de dépister ces falsifications de données (étude de la variabilité, recherche de déséquilibres à l'inclusion, diagrammes de dispersion, détection de la préférence pour certains chiffres, détection de clusters de données, mesures de similarité...) : inventer des données échappant à ces tests, si tant est que ce soit possible, demandera beaucoup plus de temps et d'énergie que réaliser les études avec rigueur... Certains journaux (*Annals of Internal Medicine* [*JAMA*]) imposent désormais une expertise statistique indépendante.

Les manquements à l'éthique (absence de saisie d'un comité d'éthique ou non-recherche de consentement) peuvent révéler des fraudes (Boldt, Reuben) et sont bien sûr condamnables en eux-mêmes (même si l'étude est par ailleurs conduite et rapportée avec rigueur). Pour les réanimateurs exerçant en France, il faut insister ici sur la nécessité de rechercher les consentements de poursuite des patients inclus avec consentement des proches ou selon une procédure d'urgence ; pour les patients n'ayant pas recouvré leur compétence en fin de séjour, de ne pas oublier de les rechercher ultérieurement ; et bien que ce ne soit pas prévu par la loi, de proposer après consentement d'urgence de rechercher un consentement de poursuite des proches, préférable à pas de consentement du tout.

Partie cachée de l'iceberg ?

Enfin, finalement très visible, la fraude a sans doute de ce fait un impact assez faible sur les pratiques — une fois les études falsifiées écartées de nos raisonnements. Beaucoup moins visibles mais aussi beaucoup plus fréquents, et donc d'impact plus grand, sont les articles biaisés par un « embellissement des données ». Les méthodes en sont nombreuses : modification du critère de jugement entre la conception de l'essai et la rédaction de l'article (62 % dans l'étude de Chan et al. [27]), calcul d'effectif a posteriori [28], exclusion de patients non répondeurs, non-respect du double insu, essais de supériorité négatifs présentés comme des essais d'équivalence, conclusion sur des critères de jugement secondaires malgré une puissance insuffisante... Cet embellissement des données peut être intentionnel ; il peut aussi résulter d'une méconnaissance des méthodes de réalisation d'une étude ou de l'interprétation des données. Il aboutit à diminuer, mais de façon occulte, le niveau de preuve des études ainsi « embellies ».

« Ce n'est pas de la découverte, mais de la recherche elle-même, de ses exigences artisanales, de l'attitude d'esprit qu'elle réclame, de l'information et des échanges qu'elle comporte, que naît l'étrange vertu qui crée l'authenticité de la connaissance médicale. »

Jean Hamburger, *La puissance et la fragilité, 20 ans après*. Flammarion, 1990.

Références

- Boldt J, Suttner S, Brosh C, et al (2009) Cardiopulmonary bypass using a high dose of a balanced hydroxyethyl starch versus an albumin-based priming strategy. *Anesth Analg* 109:1752–62 (retracted article)
- Nussmeier NA, Searles BE (2009) The next generation of colloids: ready for “prime time”? *Anesth Analg* 109:1715–17
- Shafer SL (2010) <http://www.aeditor.org/NoticeofRetraction.pdf> (consulté le 18 janvier 2011)
- Shafer SL (2010) Notice of retraction. *Anesth Analg* 111:1567
- <http://www.morgenweb.de/service/archiv/artikel/692644194.html> (consulté le 18 janvier 2011)
- <http://idw-online.de/pages/de/news399181> (consulté le 18 janvier 2011)
- Culliton BJ (1983) Coping with fraud: the Darsee case. *Science* 220:31–5
- Breuning SE, Ferguson DG, Davidson NA, Poling AD (1983) Effects of thioridazine on the intellectual performance of mentally retarded drug responders and non responders. *Arch Gen Psychiatry* 40:309–13. Retraction in Freedman DX (1988) *Arch Gen Psychiatry* 46:634
- Chandra RK (2001) Effect of vitamin and trace-element supplementation on cognitive function in elderly subjects. *Nutrition* 17:709–12. Retraction in Meguid MM (2005) *Nutrition* 21:286
- Sudbø J, Lee JJ, Lippman SM, et al (2005) Non-steroidal anti-inflammatory drugs and the risk of oral cancer: a nested case-control study. *Lancet* 366:1359–66 (retracted article)
- Reuben SS, Buvenandran A, Katz B, Kroin JS (2008) A prospective randomized trial on the role of perioperative celecoxib administration for total knee arthroplasty: improving clinical outcomes. *Anesth Analg* 106:1258–64. Retraction in Shafer SL (2009) *Anesth Analg* 108:1350
- <http://www.larecherche.fr/content/recherche/article?id=11458> (consulté le 24 janvier 2011)
- White C (2005) Suspected research fraud: difficulties of getting at the truth. *BMJ* 331:281–8
- Sox HC, Rennie D (2006) Research misconduct, retraction and cleansing the medical literature: lessons from the Poehلمان case. *Ann Intern Med* 144:609–13
- Budd JM, Slevert ME, Schultz TR (1998) Phenomena of retraction: reasons for retraction and citations to the publications. *JAMA* 280:296–7
- Nath SB, Marcus SC, Druss BG (2006) Retractions in the research literature: misconduct or mistakes? *Med J Aust* 185:152–4
- Smith R (2005) Investigating the previous studies of a fraudulent author. *BMJ* 331:288–91
- Curfman GD, Morrissey S, Drazen JM (2006) Expression of concern: Bombardier et al, Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. *N Engl J Med* 2000 343:1520–8. *N Engl J Med* 2005 353:2813–4
- <http://www.nih.gov/about/director/ebiomed/mendel.htm> (consulté le 24 janvier 2011)
- Fletcher R (1991) *Science, ideology and the media: the Cyril Burt scandal*. Transaction, London, 447 p
- Broad WJ (1980) Would-be academician pirates papers. *Science* 208:1438–40
- <http://www.sciencepresse.qc.ca/scandales/gupta.htm> (consulté le 24 janvier 2011)
- Hwang WS, Ryu YJ, Park JH, et al (2004) Evidence of a pluripotent human embryonic stem cell line derived from a cloned blastocyst. *Science* 303:1669–74. Retraction in: Kennedy D (2006) *Science* 311:335

24. Chneiweiss H (2006) Cloningate? La publication scientifique et le clonage thérapeutique face à la mystification Hwang. *Med Sci* 22:218–22
25. Gerber P (2006) What can we learn from the Hwang and the Sudbø affairs? *Med J Aust* 184:632–5
26. Roberts I, Smith R, Evans S (2007) Doubts over head injury studies. *BMJ* 334:392–4
27. Chan AW, Hróbjartsson A, Haahr MT, et al (2004) Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 291: 2457–65
28. Charles P, Giraudeau B, Dechartres A, et al (2009) Reporting of sample size calculation in randomised controlled trials: review. *BMJ* 338:b1732