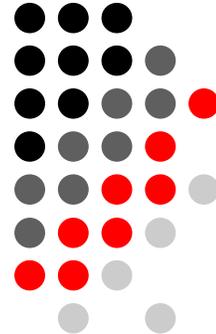




Centre interuniversitaire
de recherche sur la science
et la technologie

Note de recherche
2020-05



L'effet SIGAPS : La recherche médicale française sous l'emprise de l'évaluation comptable

Yves Gingras et Mahdi Khelifaoui



Pour nous joindre

Téléphone : 514.987-4018

Adresse électronique : cirst@uqam.ca

site : www.cirst.uqam.ca



Adresse postale

CIRST

Université du Québec à Montréal

C.P. 8888, succ. Centre-ville

Montréal (Québec)

H3C 3P8

Adresse civique

CIRST

8^e étage

Université du Québec à Montréal

Pavillon Paul-Gérin-Lajoie

1205, rue Saint-Denis

Montréal, Québec



ISBN-13 978-2-923333-83-0

Notes biographiques

Yves Gingras est professeur au Département d'histoire de l'UQAM, directeur scientifique de l'Observatoire des sciences et des technologies (OST) et membre régulier du Centre interuniversitaire de recherche sur la science et la technologie (CIRST). Il a publié de nombreux ouvrages dont les plus récents sont, *Histoire des sciences* (Presses universitaires de France, 2018), *Sociologie des sciences* (Presses universitaires de France, 2017), *L'impossible dialogue. Sciences et religions* (Boréal et Presses universitaires de France, 2016); *Science and Religion. An impossible dialogue* (Polity Press, 2017) *Les dérives de l'évaluation de la recherche. Du bon usage de la bibliométrie* (Raisons d'agir Éditions, 2014); *Bibliometrics and Research Evaluation. Uses and Abuses*, (MIT Press, 2016). Il a aussi dirigé l'ouvrage collectif *Controverses: accords et désaccords en sciences humaines et sociales* (CNRS Éditions, 2014).

Mahdi Khelifaoui est chercheur postdoctoral à l'Université d'Ottawa et membre associé au Centre interuniversitaire de recherche sur la science et la technologie (CIRST). Il a publié plusieurs travaux en bibliométrie, notamment dans *Scientometrics*, *Quantitative Science Studies* et *Research Evaluation*.

Remerciements

Nous tenons à remercier Hervé Maisonneuve Michel Dubois et Emilien Schultz pour leurs commentaires et suggestions sur une version préliminaire de ce texte.

Ce texte paraîtra dans la revue *Zilsel. Sociologie, histoire, anthropologie et philosophie des sciences et des techniques*.

Résumé

Cette recherche a pour but de mettre en évidence les effets pervers générés par l'introduction du système SIGAPS (Système d'interrogation, de gestion, et d'analyse des publications scientifiques) sur la production scientifique française en médecine et en sciences biomédicales. Cet outil bibliométrique de gestion et de financement de la recherche présente un exemple emblématique des dérives que peuvent générer les méthodes d'évaluation de la recherche reposant sur des critères purement comptables. Dans cette note, nous présentons d'abord le fonctionnement de SIGAPS, pour ensuite expliquer précisément en quoi les méthodes de calcul des « points SIGAPS », fondés sur les facteurs d'impact des revues et l'ordre des noms des co-auteurs, posent de nombreux problèmes. Nous identifions notamment les effets du système SIGAPS sur les dynamiques de publications, les choix des lieux de publications, la langue de publication et les critères de recrutement et de promotion des chercheurs. Finalement, nous montrons que l'utilisation du système SIGAPS ne répond pas bien à tous les critères de ce que l'on pourrait appeler une « éthique de l'évaluation » qui devrait respecter certaines règles, comme la transparence, l'équité et la validité des indicateurs.

Table des matières

INTRODUCTION	1
1. LE SYSTÈME DE FINANCEMENT SIGAPS	4
2. LES EFFETS DU SYSTÈME DE FINANCEMENT SIGAPS	7
2.1 La multiplication des publications	8
2.2 Les fluctuations du facteur d'impact	15
2.3 La langue de publication	18
2.4 Choix des co-auteurs	20
3. LES DÉRIVES DU SYSTÈME SIGAPS	21
CONCLUSION : ÉTHIQUE ET DÉRIVE SYSTÉMIQUE	25
BIBLIOGRAPHIE	27

Introduction

L'évaluation de la recherche est aujourd'hui caractérisée par le recours de plus en plus systématique à des méthodes quantitatives, s'appuyant surtout sur la mesure du nombre de publications et de citations reçues par les chercheurs. Désormais largement répandues dans le monde académique, elles ont accompagné depuis une quinzaine d'années, et à des degrés divers selon les pays, la montée de la gestion technocratique et néolibérale des universités et des institutions scientifiques (Barats et al., 2018; Angermuller et van Leeuwen, 2019). Ces méthodes reposent sur l'usage d'indicateurs bibliométriques souvent mal construits ou mal utilisés (Gingras, 2014a), tels que « l'indice h » et les « facteurs d'impacts » des revues. Elles ont été acceptées avec une surprenante résignation par la communauté scientifique, surtout dans le monde biomédical, et parfois même accueillies avec enthousiasme par certains chercheurs et gestionnaires qui croient en leur plus grande « objectivité ». Les chercheurs en début de carrière, adoptant une posture pragmatique – et souvent peu critique – vis-vis de ces indicateurs, les considèrent désormais comme faisant partie des « règles du jeu » académique (Nicholas et al., 2018, 2020). Le retour de bâton s'est toutefois fait ressentir avec la multiplication des cas d'inconduites et de manquements à l'éthique scientifique un peu partout à travers le monde (Chevassus-au-Louis, 2016; Biagioli et Lippman, 2020).¹

La Chine, par exemple, avait promu ce type d'évaluations comptables pendant des années, en octroyant des primes faramineuses aux chercheurs les plus présents dans des revues de langue anglaise dites à « facteur d'impact élevé ». Mais elle a finalement décidé d'abandonner ce système, tant il a engendré de comportements déviants, dont l'utilisation de données frauduleuses, la manipulation des processus d'évaluations par les pairs, voire même le développement d'un marché noir de co-signatures sur des articles acceptés dans des revues « prestigieuses » (Hvistendhal, 2013; Mallapty, 2020). En Italie, une récente étude bibliométrique

¹ Les critiques répétées envers ces évaluations de type comptable ont cependant eu quelques effets, puisque certains organismes de recherche interdisent désormais l'usage par les comités d'évaluation des indicateurs les plus simplistes, comme le facteur d'impact des revues et « l'indice h » des chercheurs, et insistent plutôt sur l'analyse de la qualité des projets et des publications. C'est par exemple le cas du Conseil de recherches en sciences naturelles et en génie du Canada qui, dans son guide destiné aux évaluateurs de ses programmes de financement, exclut de sa définition de l'impact scientifique « les indicateurs quantitatifs, tels que le facteur d'impact ou l'indice h ». Voir : https://www.nserc-crsng.gc.ca/_doc/Reviewers-Examineurs/Complete-Manual-ManualEvalCompleet_eng.pdf

a mis en lumière les conséquences délétères d'une loi, promulguée en 2010, qui rendait les embauches et les promotions des chercheurs italiens tributaires de leurs « performances » en termes de publications et de citations reçues (Baccini, De Nicolao et Petrovich, 2019). Censés mettre fin aux pratiques de copinages, ces critères d'évaluation quantitatifs, dits « objectifs », ont dans les faits conduit à une croissance anormale du taux d'autocitations nationales, c'est-à-dire des citations faites par les chercheurs d'un pays aux publications de leur pays.² Ainsi, loin d'inciter les chercheurs italiens à l'« excellence » et à produire une science de meilleure qualité, la loi de 2010 les a simplement poussé à citer beaucoup plus souvent leurs propres publications ou celles de leurs collègues, afin de gonfler artificiellement leur « impact scientifique » et celui de leur département ou institution de rattachement.

La controverse récente entourant le nombre de publications jugé anormalement élevé du Dr. Didier Raoult (Gingras et Khelifaoui, 2020), et l'attention médiatique portée au financement de son Institut, a eu pour conséquence de faire remonter à la surface les effets indésirables d'un système de financement de la recherche médicale française implanté en 2008, le Système d'Interrogation, de Gestion et d'Analyse des Publications Scientifiques. Désigné par l'acronyme SIGAPS, ce système inconnu du grand public (avec très peu de références dans la presse française) a été développé au Centre hospitalier universitaire (CHU) de Lille en 2002, d'abord comme un logiciel de cartographie et de suivi systématique de l'activité scientifique des chercheurs de l'établissement. En 2006, le ministère de la Santé lui a confié pour mission de l'implanter dans 70 autres institutions hospitalières, dont les 32 CHU de France, ce qui fut complété en 2008. En 2016, 116 institutions de santé de différents types, dont les CHU, des centres hospitaliers, des Centres de lutte contre le cancer, des établissements à but non lucratif ou privés, étaient inclus dans le système SIGAPS.

Le déploiement du système SIGAPS à l'échelle nationale s'est accompagné d'une transformation importante de son utilisation et de sa finalité. D'outil de cartographie, de gestion et d'analyse des publications à l'échelle des établissements, il a été transformé en outil de gestion administrative de l'enveloppe budgétaire dédiée à la recherche. L'adoption de SIGAPS par le ministère de la Santé visait notamment à accompagner la réforme des modalités de

²Voir également Khelifaoui et. al. (2020) qui analysent l'évolution historique de cette tendance pour les principaux pays producteurs de science, mesurée à l'aide des autoréférences nationales et non des autocitations.

financement des missions d'enseignement, de recherche, de référence et d'innovation (MERRI), adoptée en 2007. Cette réforme divisait le financement de l'enseignement et de la recherche médicale en trois parts : une part fixe, une part variable et une part modulable dans laquelle fut inclus un financement fondé sur les données générées par SIGAPS. Alors que les budgets de recherche, pour les CHU notamment, étaient auparavant alloués sur une base forfaitaire et proportionnelle à la taille de l'établissement, l'instauration du nouveau système SIGAPS a profondément changé la donne, en introduisant un financement basé sur la « performance », modèle qualifié par certains chercheurs de « tarification à l'acte bibliométrique » (Mancini *et al.*, 2009).

En effet, le financement de la recherche étant de plus en plus lié à des critères quantitatifs de « performance », SIGAPS est, de fait, également devenu indirectement un système d'évaluation des chercheurs et des institutions, car les budgets SIGAPS alloués dépendent de la cotation des articles, lesquels sont bien sûr signés par des chercheurs. Les chercheurs des CHU ont d'ailleurs fini par faire eux-mêmes le lien entre leurs publications et le financement SIGAPS. Ainsi, une enquête récente menée au sein de l'Assistance Publique - Hôpitaux de Paris, principal CHU de France et premier bénéficiaire de l'enveloppe SIGAPS, relevait que 56% des chercheurs interrogés regrettaient que ces fonds, qui reviennent à l'institution, ne leur soient pas versés directement pour les aider à financer leurs recherches et publications (Duracinsky, Péretz et Chassany, 2017). Mais le signe le plus visible du glissement de l'outil vers l'évaluation individuelle des chercheurs, qui sera documenté plus loin dans l'article, est que plusieurs CHU utilisent désormais les données de publication individuelles compilées par SIGAPS dans les processus d'embauche ou de promotion des chercheurs.

L'objet de cet article est de mettre en évidence les effets pervers générés par un tel système. Il ne s'agit pas ici de retracer la genèse du système SIGAPS, ce qui exigerait des entrevues et des recherches en archives, mais plutôt de partir de l'existence du système pour analyser ses effets sur divers aspects de la recherche, tant au niveau individuel qu'institutionnel. Nous présenterons d'abord le fonctionnement de SIGAPS, pour ensuite expliquer précisément en quoi le calcul des « points SIGAPS », fondé sur les facteurs d'impact des revues, pose de nombreux problèmes. Nous montrerons que l'utilisation du système SIGAPS n'est pas conforme à ce que l'on pourrait appeler une « éthique de l'évaluation » qui devrait respecter

certaines règles élémentaires, comme la transparence, l'équité et la validité des indicateurs. Mal construits et sélectionnés de manière arbitraire, ces indicateurs incorporés dans des algorithmes de calcul font rapidement oublier aux chercheurs leur absence de fondement, alors qu'ils sont pourtant les premiers visés par ces évaluations purement numériques de leurs activités de recherche.

1. Le système de financement SIGAPS

Le système de financement SIGAPS permet d'attribuer un score aux chercheurs et aux institutions hospitalières françaises en fonction de leurs publications savantes. Le nombre de points accordés est une fonction de la « qualité » de la revue où est parue la publication et de la position du chercheur sur la liste des co-auteurs de cette publication. La « qualité » des revues est elle-même évaluée à partir d'une liste qui les classe en catégories (A, B, C, D, E, NC) sur la base de leur « facteur d'impact », obtenu auprès de Clarivate Analytics, firme qui produit la base de données Web of Science (WoS), laquelle sert de référence pour les calculs des facteurs d'impact. Cet indicateur est obtenu en divisant le nombre de citations faites pendant deux ans à une revue « X » par toutes les revues indexées dans le WoS (incluant la revue « X »), par le nombre de publications parues dans la revue « X » durant la même période. Les facteurs d'impact des revues pour une année donnée permettent à SIGAPS de réaliser un classement des revues, selon différentes spécialités médicales et biomédicales, en cinq catégories (A, B, C, D, E). Ainsi, pour une spécialité donnée, les revues se trouvant au-dessus du 90^e percentile sont classées « A », les revues se trouvant entre le troisième quartile et le 90^e percentile sont classées « B », les revues se trouvant entre la médiane et le troisième quartile sont classées « C », les revues se trouvant entre le premier quartile et la médiane sont classées « D », alors que les revues classées en dessous du premier quartile sont classées « E ». Une sixième catégorie de revues, dite « NC » pour non classé, vient s'ajouter. Elle comprend certaines revues indexées dans la base de données médicale PubMed, mais qui ne sont pas incluses dans le WoS.

Le facteur d'impact constitue une moyenne qui caractérise la revue et non pas les articles eux-mêmes, lesquels peuvent être plus ou moins cités que la moyenne des articles de la revue. De plus, cette moyenne est calculée sur deux ans alors que l'impact des publications savantes

se fait généralement sentir sur une durée de 5 à 6 ans en sciences médicales et de 7 à 8 ans en sciences naturelles et en ingénierie (Larivière et Sugimoto, 2018, p. 112). Ce décalage temporel est évidemment dû au fait qu'il faut d'abord lire l'article, ensuite écrire un texte le citant, le soumettre à une revue et attendre qu'il soit publié après avoir été évalué. Pour cette raison, il est évident qu'un facteur d'impact calculé sur 5 ans modifierait de manière importante les rangs des revues (Glanzel et Moed, 2002). La mesure retenue pour évaluer la « qualité » des publications dans le système SIGAPS n'est donc pas basée sur la publication elle-même, évaluée en termes de citations réellement obtenues, mais sur la notoriété de la revue dans laquelle l'article est publié. En fait, en termes statistiques, le facteur d'impact ne mesure que la probabilité qu'un article soit cité dans le futur, sachant la revue dans laquelle il est publié. Ainsi, un article paru dans une revue classée « C » peut s'avérer être plus cité qu'un autre paru dans une revue classée « B ».

Le problème posé par la classification disciplinaire des revues en fonction de leur facteur d'impact prend une ampleur supplémentaire dans le cas des revues à caractère multidisciplinaire. En effet, le nombre de citations reçues par un article peut varier grandement en fonction des spécialités, qui se distinguent les unes des autres par le nombre de chercheurs qui y sont actifs ainsi que par leurs rythmes de publications. Une excellente revue associée à une spécialité où les chercheurs sont peu nombreux et les rythmes de publications lents aura nécessairement un facteur d'impact inférieur à une revue de même calibre, mais associée à une spécialité où les chercheurs, et donc les potentiels auteurs citant, sont beaucoup plus nombreux. Or, il peut aussi se trouver qu'une revue de moindre qualité associée à une spécialité où les chercheurs sont très nombreux, ait quand même un facteur d'impact plus élevé qu'une très bonne revue associée à une spécialité de taille plus petite. Ce problème se règle normalement par la classification disciplinaire des revues, ce que fait SIGAPS, mais il demeure dans le cas des revues multidisciplinaires, qui sont associées à plusieurs spécialités. En effet, dans le système SIGAPS, une revue multidisciplinaire est classée par rapport à l'ensemble des revues des différentes spécialités à laquelle elle est associée, ce qui va nécessairement pénaliser les chercheurs qui sont actifs dans des spécialités de petites tailles. Des chercheurs français ont identifié ce problème dans SIGAPS dès 2009, en citant le cas éclairant d'une revue de statistiques médicales : « Par exemple, le journal *Statistics in Medicine* ($IF_{2006} = 1,7$) est affecté à 5 disciplines différentes et s'il est plutôt bien classé dans les disciplines

« Informatique Médicale » (rang 4/22) et « Statistiques & Probabilité » (10/83) il l'est beaucoup moins dans les autres disciplines, notamment « Médecine, recherche et expérimentation » (39/76) dans laquelle il se retrouve en concurrence avec des journaux tels que *Nature Medicine* ($IF_{2006} = 28,6$). Au final, du fait de son appartenance multiple, le journal sera catégorisé C, malgré une bonne reconnaissance chez les biostatisticiens » (Mancini et al., 2009, p. 175).

Le classement global des revues est complété en attribuant, de manière arbitraire, un pointage à chaque catégorie : 8 points pour une revue classée « A », 6 points pour une revue classée « B », 4 points pour une revue classée « C », 3 points pour une revue classée « D », 2 points pour une revue classée « E » et 1 point pour une revue classée « NC ». En plus de la valeur attribuée à la revue et, par extension, à l'article, SIGAPS tient également compte de la position des chercheurs dans la liste des co-auteurs. Nous savons en effet que, dans le domaine des sciences biomédicales, les articles ont un très grand nombre d'auteurs. On peut alors penser que toutes les contributions se valent et simplement diviser le nombre de points par le nombre d'auteurs. Mais le « sens commun savant » penche plutôt pour l'idée que les auteurs signataires n'ont pas tous contribué de manière égale et qu'il faut donc distinguer la valeur de chacun en fonction de sa position dans la liste co-auteurs. Les manières de décider de l'ordre des noms varie beaucoup selon les disciplines (Pontille, 2016). Dans le système le SIGAPS, défini pour les sciences biomédicales, il a ainsi été fixé qu'occuper la première ou la dernière position de la liste des co-auteurs d'un article vaut 4 points, la seconde position vaut 3 points, la troisième ou l'avant-dernière position valent 2 points, toute autre position valant 1 point. Même si l'ordre des noms peut faire consensus, il demeure que le simple fait de quantifier ces positions, risque de renforcer la course au « premier auteur » – le dernier étant le plus souvent le responsable du laboratoire – et donc les tensions, sinon les conflits ouverts, entre collaborateurs.

Le score SIGAPS final d'une publication est donc le produit de deux variables : le nombre de points obtenus par la position de l'auteur multiplié par le nombre lié à la catégorie de la revue (A, B, C, D, E, NC). Pour donner un exemple simple, occuper la première ou la dernière position des co-auteurs d'un article publié dans une revue classée « A », comme *Nature* ou *The Lancet*, rapporte 4 points pour la position de l'auteur, multipliés par 8 points pour le

classement de la revue, soit 32 points dans SIGAPS. Par contre, occuper la première ou la dernière position d'un article publié dans une revue à faible impact, classée dans la dernière catégorie dite « NC » (non classée), ne rapporte « que » 4 points (position de l'auteur) multipliés par 1 point (valeur de la revue), soit 4 points.

Les points accumulés par une institution sont alors la somme des points associés aux articles des chercheurs qui lui sont affiliés. Finalement, ces points sont ensuite convertis en valeur monétaire. La valeur d'un point SIGAPS n'est pas fixe et peut changer d'année en année. En 2017, par exemple, un point SIGAPS valait 648 euros. Par conséquent, la publication d'un article dans une revue de catégorie « A » pouvait rapporter jusqu'à près de 21 000 euros en fonction de la position de l'auteur de cette institution sur la liste des co-auteurs, lesquels peuvent appartenir à différentes institutions. Une fois fixée, cette somme est versée, dans le cadre de la dotation MERRI du ministère de la Santé, à l'institution de rattachement du chercheur de façon récurrente durant quatre années consécutives, soit un total de 84 000 euros. L'institution bénéficiaire peut ensuite utiliser cet argent pour mener sa politique de recherche: acheter des équipements ou financer des projets de recherche. Mais elle peut aussi décider, à sa discrétion, de reverser une partie de la somme reçue à ses laboratoires de recherche, au prorata de leurs contributions en points SIGAPS. Comme le rapporte une récente enquête du magazine l'Express (Garcia, 2020), certains établissements hospitaliers privés, qui sont de plus en plus nombreux à pouvoir bénéficier des financements SIGAPS, vont même jusqu'à les reverser directement à leurs médecins sous forme d'honoraires.

Notons enfin que les chercheurs doivent valider la liste de leurs publications, car les repérages automatiques ne suffisent pas, étant donné la complexité des appartenances institutionnelles. De manière assez paradoxale, ils contribuent ainsi à alimenter un système d'évaluation qui peut avoir pour effet de les inciter à produire des publications non pas en fonction de leurs intérêts scientifiques, mais de manière à maximiser les points SIGAPS de leurs laboratoires.

2. Les effets du système de financement SIGAPS

Analysons maintenant les effets prévisibles d'un tel système de financement de la recherche, qui est aussi *de facto* une forme d'évaluation, sur le comportement des chercheurs et des gestionnaires des institutions de recherche.

2.1 La multiplication des publications

Sur le plan des publications d'abord, il est évident que tout établissement de santé éligible au système SIGAPS a désormais intérêt à « booster » son nombre de publications afin d'augmenter sa part de revenus dans l'enveloppe annuelle globale des MERRI actuellement fixée à 1.6 milliard d'euros, dont 60% sont distribués selon les données de SIGAPS. L'effet de l'introduction de systèmes d'évaluation comptables sur l'augmentation de la production des chercheurs est connu et a déjà été documenté. Par exemple, l'Australie, qui avait introduit un tel système au début des années 1990, avait vu son nombre de publications augmenter de 25% en l'espace de quelques années seulement. Par contre, le gouvernement n'avait pas prévu ce qui était pourtant probable, à savoir que l'impact de ces nombreuses publications, mesuré en termes de citations relatives reçues, allait décliner, ce que des études bibliométriques ont d'ailleurs plus tard montré (Butler, 2003).

L'évaluation étant liée aux budgets, on devrait donc observer une croissance du nombre de publications après la mise en application générale du système SIGAPS (Blanc, Bouliou et Bertram, 2019). N'ayant pas accès aux données complètes de SIGAPS, nous allons uniquement évaluer l'évolution de la production des publications des CHU, qui reçoivent 80% de l'enveloppe MERRI et contribuent pour l'essentiel de la recherche médicale française en dehors du CNRS et de l'INSERM. La Figure 1 montre l'évolution du nombre de publications en médecine clinique, sciences biomédicales et biologie des CHU sur la base de données du WoS, entre 2000 et 2018. Nous nous limitons ici à trois catégories de documents: les articles de recherche, les notes de recherche et les synthèses. Contrairement à SIGAPS, nous ne pouvons pas inclure les articles publiés dans les revues de catégories « NC », généralement à faible impact, car elles ne sont pas indexées dans le WoS. De plus, un pourcentage non négligeable de chercheurs de CHU signe leurs articles en indiquant d'autres affiliations institutionnelles, une université ou l'INSERM par exemple. Ces publications ne sont donc pas recensées comme des publications de CHU dans le WoS, alors qu'elles le sont dans SIGAPS, puisque les chercheurs valident personnellement leurs publications dans ce système et peuvent donc inclure même celles qu'ils ont signées avec d'autres adresses institutionnelles. Malgré ces limites, nos données suffisent amplement à observer l'effet de SIGAPS sur les publications.

La Figure 1 montre, selon les données du WoS, que l'on passe, entre 2000 et 2018, de 8116 à 11752 publications, soit une augmentation de 45%. Cette augmentation touche l'ensemble des publications françaises en sciences médicales et biomédicales puisque, par comparaison, la croissance de l'ensemble des publications françaises dans les trois mêmes disciplines et pour la même période est de 47% et suit la même tendance. La courbe montre bien que la croissance des publications des CHU s'accélère à partir de 2006, lorsque le système SIGAPS commence à se diffuser graduellement dans les CHU.³

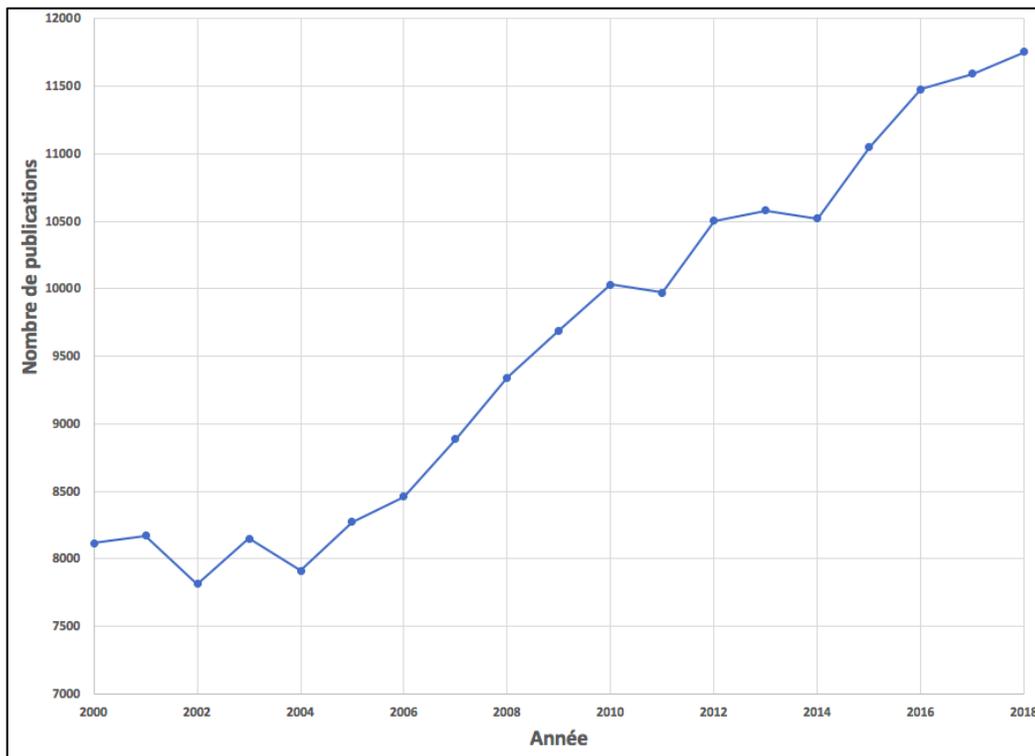


Figure 1 – Évolution du nombre de publications des CHU recensées dans le WoS

³ Comme indiqué plus haut, les nombres de publications que nous obtenons sont inférieurs à ceux établis en 2016 par la cellule bibliométrique du Comité national de coordination de la recherche (CNCR), sur la base de données fournies par SIGAPS. Selon cette dernière, le nombre de publications des CHU, comptabilisant uniquement les articles de recherche hors revues classées « NC », est passé de 10203 à 12874, entre 2005 et 2014, pour une croissance de 26% (Baudoin et Devos, 2016, p. 8). En comparaison, nous obtenons dans la même période, une croissance de 24%, ce qui est très proche comme tendance. La robustesse de notre échantillon est confirmée par la comparaison de notre liste des 20 revues dans lesquelles les chercheurs des CHU publient le plus, avec celles établies dans le rapport du CNCR pour les périodes 2005-2009 et 2010-2014. Dans les deux périodes, nous avons respectivement 18 et 17 revues en commun sur 20. Comme pour le CNCR, les revues du top-20 de notre échantillon incluent environ 25% du total des publications.

Si l'augmentation du nombre total de publications entre 2000 et 2018 est de 45% pour l'ensemble de notre échantillon, la plupart des CHU ont cependant vu leurs publications augmenter de façon beaucoup plus importante. Par exemple, les CHU des grandes villes comme Toulouse, Bordeaux, Montpellier, Lyon, Lille ou Rennes ont au moins doublé leurs nombres de publications entre 2000 et 2018, tandis que l'Assistance Publique - Hôpitaux de Paris (AP-HP) et l'Assistance Publique - Hôpitaux de Marseille (AP-HM), les plus importants producteurs de publications, ont connu des croissances respectives moindres de 23% et 85%. L'écart constaté entre la croissance globale du nombre de publications et la croissance des contributions individuelles de chaque CHU suggère que le nombre de publications des CHU en collaboration avec d'autres institutions, françaises ou étrangères, n'a cessé d'augmenter entre 2000 et 2018. Comme le montre la Figure 2, le nombre moyen d'institutions cosignataires de publications où figure au moins une adresse de CHU est passé de 2.31 en 2000 à 7.45 en 2018. Sur cette période, les CHU signent donc des publications qui comportent en moyenne de plus en plus d'institutions cosignataires. Par comparaison, si l'on considère, sur la même période, l'ensemble des publications en biologie, médecine clinique et recherche biomédicale contenant au moins une adresse institutionnelle française, on voit que le nombre moyen d'institutions cosignataires est passé de 2.17 en 2000 à 5.90 en 2018. La croissance du nombre moyen d'institutions cosignataires d'articles où au moins un CHU est présent est donc plus rapide que celle du nombre moyen d'institutions cosignataires d'une publication médicale ou biomédicale française. On observe d'ailleurs que la pente de la courbe de croissance des publications des CHU augmente à partir de 2006 de telle sorte que l'écart entre ces deux courbes se creuse plus rapidement depuis la mise en place du système SIGAPS. Cela va dans le sens de l'hypothèse d'un effet de ce système sur les pratiques de publication.

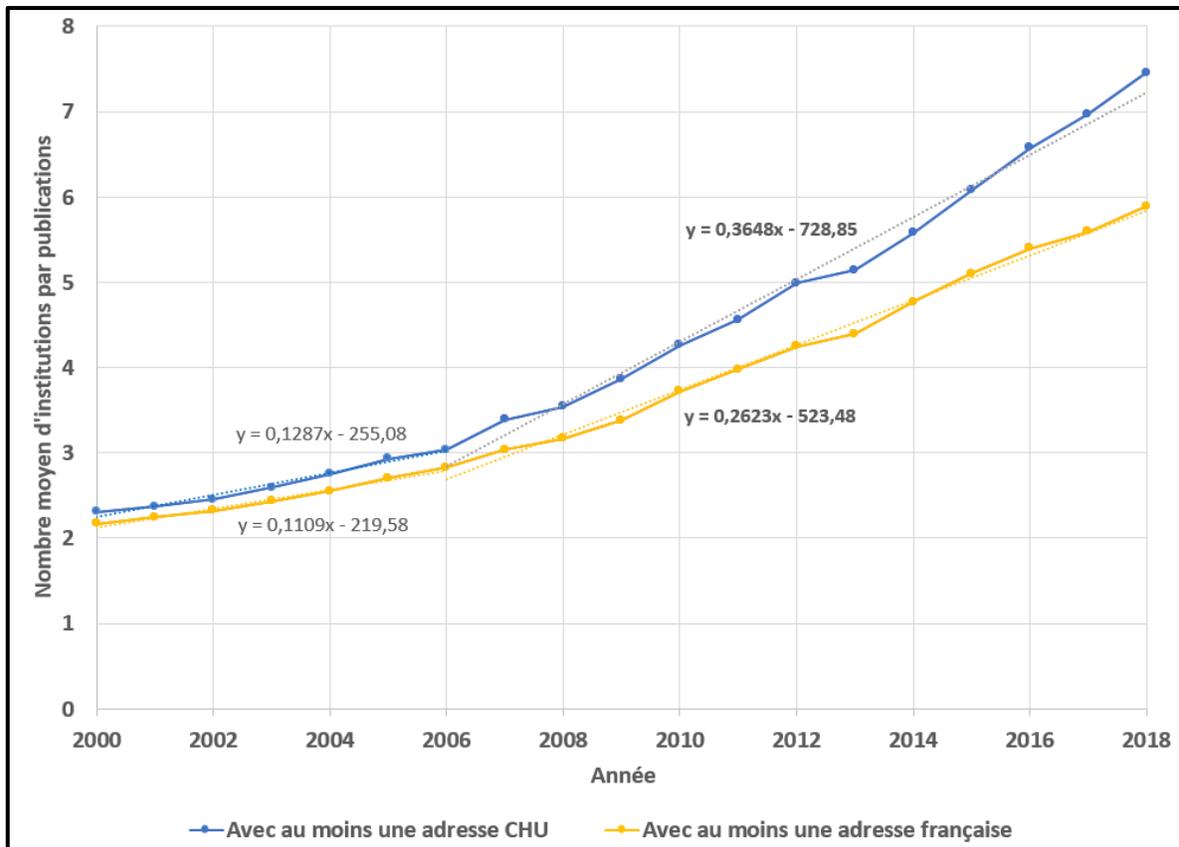


Figure 2 – Évolution du nombre moyen d’institutions cosignataires de publications où figure au moins une adresse de CHU

Les collaborations interinstitutionnelles entre CHU n’ont pas cessé elles-aussi d’augmenter depuis la mise en place du système. Cela est confirmé par l’évolution du nombre moyen de signatures d’établissements par publication, qui est passé de 1.15 à 1.55 entre 2000 et 2018. La Figure 2 montre bien que le taux de collaborations entre CHU s’est accéléré vers 2006, année du début de l’implantation de SIGAPS dans l’ensemble des institutions. On ne peut bien sûr écarter l’idée que d’autres réformes appliquées à cette époque (comme les « Plan hôpital » de 2007 et 2012) aient pu également contribué à ces changements. Les CHU s’orientent donc vers la multiplication des collaborations interinstitutionnelles, que ce soit avec des institutions qui sont éligibles au financement SIGAPS ou non. Ces collaborations sont l’occasion pour chaque participant éligible de produire davantage de publications qui lui permettront d’augmenter ou de maintenir son score SIGAPS. Cela confirme la suggestion de

Fanelli (2020) voulant que la pression à publier soit plus visible dans la croissance des collaborations que dans le nombre d'articles produit par chaque auteur. Cependant, l'enveloppe MERRI étant plafonnée, et les établissements éligibles hors CHU étant de plus en plus nombreux, il s'ensuit nécessairement « une baisse régulière de la part des CHU dans les scores SIGAPS » et ce, même quand leurs scores en valeur absolue continuent d'augmenter ou demeurent stables (Milon, 2018, p. 12). Il faut noter finalement que l'augmentation du nombre d'institutions éligibles à SIGAPS participant à une même publication a pour effet de diminuer les montants octroyés à chaque institution participante. Comme l'enveloppe MERRI est plafonnée, cela signifie que pour maintenir leurs scores, les CHU doivent compenser cette dilution par l'augmentation de leur nombre total de publications.

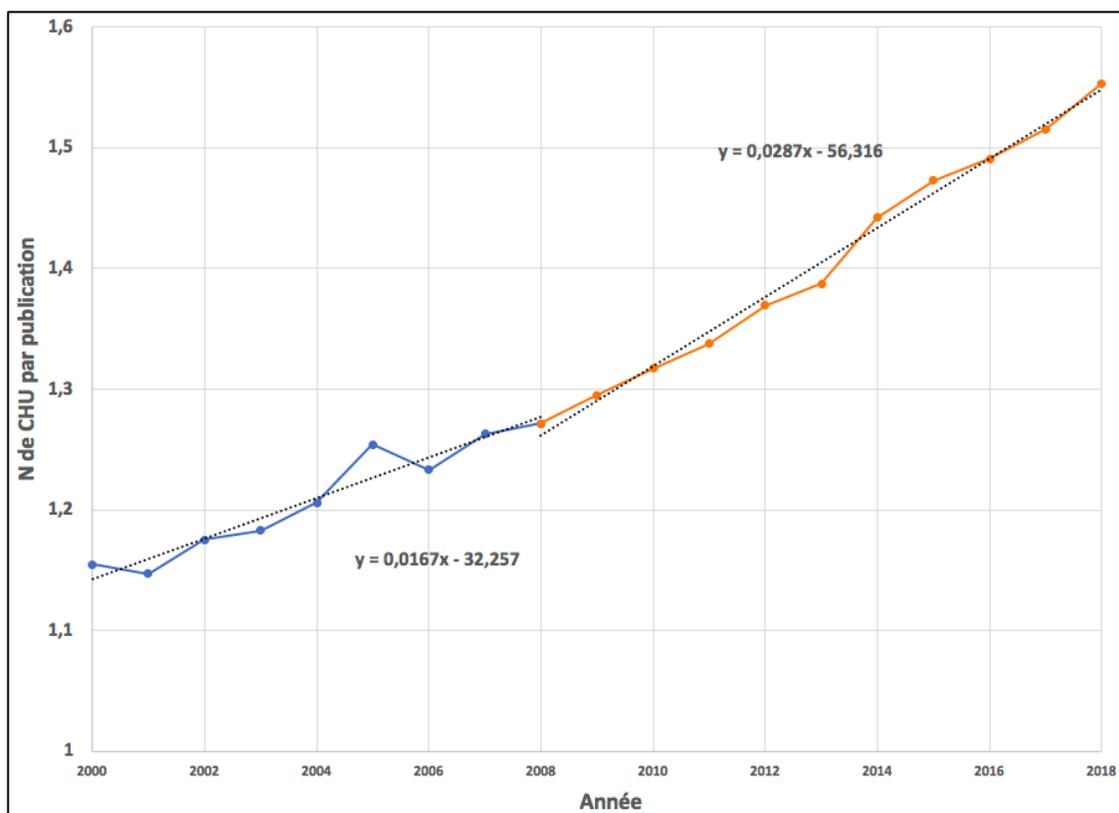


Figure 3 – Évolution du nombre de CHU signataires par publication

Les revues ne rapportant pas toutes le même nombre de points, on devrait s'attendre à des changements de stratégies de publication suite à la mise en place du système SIGAPS. C'est ce que l'on constate justement avec la revue multidisciplinaire en libre accès *PLoS ONE*,

lancée en 2006. On remarque que durant la période 2005-2009 cette revue est peu prisée par les chercheurs et se classe en 440^e place avec seulement 22 publications. Cependant, dès la période suivante, 2010-2014, elle connaît une ascension fulgurante vers la première position avec 1487 publications et maintient ce rang au cours de la période 2015-2018 avec un total de 1222 publications. Or, dans le cadre du système SIGAPS, publier dans *PLoS ONE* comporte un avantage stratégique, car elle est classée au rang « B » et peut donc rapporter un nombre de points élevé (jusqu'à 24 points pour une position de premier ou de dernier auteur). Un autre avantage de cette revue est que son taux d'acceptation d'environ 70% est très élevé. Cela s'explique par la politique de la revue qui privilégie la validité des résultats, plutôt que leur originalité, comme on peut d'ailleurs le lire sur son site internet : « the editors make decisions on submissions based on scientific rigor, regardless of novelty »⁴.

Si l'on raisonne dans les termes purement comptables de SIGAPS, il devient plus intéressant pour un chercheur de publier un article dans *PLoS ONE*, plutôt que dans une revue disciplinaire de même rang, voire même de rang inférieur, mais qui applique des critères de sélectivité plus élevés. Par exemple, publier un article en oncologie dans *PLoS ONE* comme premier ou dernier auteur rapporte 24 points SIGAPS, soit autant que publier le même article dans la revue *European Journal of Cancer*, également classée « B » depuis 2016, mais qui a un taux d'acceptation très faible de 14%. Plus curieux encore, publier ce même article dans la revue *Journal of Oncology*, qui est plus sélective que *PLoS ONE* avec un taux d'acceptation de 31%, mais qui est seulement classée « C », ne rapporte que 18 points. La différence entre les deux stratégies équivaut à 39 000 euros sur quatre ans. Le même phénomène s'observe également pour la revue *Scientific Reports*. Cette autre « méga-revue » en ligne et en accès libre, fondée par l'éditeur privé *Nature Publishing Group* en 2011, est devenue dans la période 2015-2018 la deuxième revue dans laquelle les CHU publient le plus fréquemment. Classée « B » dans le système SIGAPS, cette revue avait un taux d'acceptation de 56% en 2018, soit bien plus que plusieurs revues d'oncologie du même rang, mais qui sont plus sélectives. La montée des « méga-revues », en particulier *PLoS ONE*, comme lieu privilégié de publication des CHU avait déjà été notée par le rapport du CNCR, qui indiquait également une croissance

⁴<https://journals.plos.org/plosone/s/editorial-and-peer-review-process>

de 26% par année du nombre de publications de CHU dans des revues en accès libre, entre 2005 et 2014 (Baudoin et Devos, 2016, p. 15).

Les revues en accès libre nouvellement créées, de qualité très inégale et plus facile d'accès, car payantes, sont classées « NC » dans SIGAPS, en attendant d'être éventuellement indexées dans le WoS et dotées d'un facteur d'impact qui leur permette d'être catégorisées A, B, C, D, ou E. Face à la multiplication et au fort taux de renouvellement des revues en libre accès, le rapport du CNCR note ainsi que « le recours plus fréquent des chercheurs aux revues nouvellement créées, souvent moins sélectives et plus faciles d'accès, entraîne une augmentation de la part NC dans la répartition des revues par catégorie » (Baudoin et Devos, 2016, p. 17). Dans ce rapport, on observe, en effet, une croissance de la part des publications des CHU dans les revues classées « NC » de 7 à 11% entre 2005 et 2014. On note également une croissance de la part des publications dans les revues classées « B », de 18 à 23% dans la même période. Cette augmentation s'inscrit dans la même logique du recours aux revues en libre accès, puisqu'elle s'explique en grande partie par le nombre important de publications parues dans la revue *PLoS ONE*.

Par ailleurs, de 2005 à 2014, la part des publications dans les revues classées « A » est restée stable autour de 19 %, tandis que la part des publications dans les revues classées « C » est passée de 17 à 20%, et dans les revues classées « D » de 11 à 13%. Finalement, la part des publications dans les revues classées « E » a connu une baisse importante de 30 à 14%. Cette baisse s'explique notamment par la chute du nombre de publications en langue française des CHU, question que nous abordons plus loin. En effet, une grande partie des revues médicales françaises sont classées « E », étant donné leur faible visibilité internationale. Le système SIGAPS a donc pour effet de précipiter le déclin des revues médicales françaises à caractère national, alors que le rapport du CNCR note pourtant que celles-ci « servent régulièrement de terrain d'entraînement aux premières publications des jeunes médecins, en même temps qu'elles constituent une source importante de la mise à jour des connaissances des soignants en France » (Baudoin et Devos, 2016, p. 14).

Une autre revue ayant attiré notre attention est *Oncotarget*, qui, avec 274 articles, arrive à la 9^e position des revues dans lesquelles les CHU ont le plus publié dans la période 2015- 2018.

Cette revue, également en libre accès, a été créée en 2010 par un éditeur obscur, Rapamycin Press. Dès 2015, elle a été classée sur la célèbre liste des revues dites « prédatrices » établie par le bibliothécaire américain Jeffrey Beall (2012, 2017). En 2018, Clarivate Analytics, la compagnie qui produit les facteurs d'impacts des revues indexées dans le WoS, a décidé de désindexer la revue de sa base de données. Bien que les raisons justifiant ce retrait n'aient pas été clairement explicitées par Clarivate Analytics, ce genre de mesure est généralement pris à l'encontre des revues qui se rendent coupables de cas de fraudes ou d'autres pratiques contraires à l'éthique scientifique, comme des taux d'autocitations élevées (*citation stacking*) qui visent à « booster » la valeur du facteur d'impact de la revue.

Le médecin français Hervé Maisonneuve a par ailleurs noté l'impressionnant rythme de publication de la revue, qui avoisinait en 2017 les deux numéros par semaine (soit près de 7400 articles selon nos calculs), ce qui n'est jamais un gage de qualité.⁵ Le retrait de *Oncotarget* de la base du WoS signifie que celle-ci n'a désormais plus de facteur d'impact qui permette au système SIGAPS de la classer. Le dernier facteur d'impact calculé par Clarivate Analytics pour cette revue en 2016 était de 5.168, ce qui la classait au rang « B ». Une publication en tant que premier auteur y rapportait donc 24 points, soit 15 500 euros versés à l'institution de rattachement du chercheur durant quatre années consécutives. L'exemple de la revue *Oncotarget* montre qu'une revue bien classée dans SIGAPS n'est pas nécessairement une revue de qualité selon les standards scientifiques. Elle montre également les limites d'un système qui permet d'octroyer plus de 60 000 euros à une institution, pour la « récompenser » d'une publication dans une revue de qualité douteuse.

2.2 Les fluctuations du facteur d'impact

Le rang attribué à chaque revue étant fondé sur le facteur d'impact de la revue, il faut également noter la grande volatilité des rangs SIGAPS des revues au fil des ans.

Tout d'abord, une revue peut conserver le même rang d'une année à l'autre malgré une variation relativement importante de la valeur de son facteur d'impact, puisque son rang dépend non seulement de son propre facteur d'impact, mais également des facteurs d'impact de

⁵ Voir sur le blogue d'Hervé Maisonneuve: <https://www.h2mw.eu/redactionmedicale/2018/03/oncotarget-nest-plus-index%C3%A9-dans-medline-et-web-of-science-sans-raisons-ce-serait-une-revue-pr%C3%A9datrice.html>

toutes les autres revues qui sont associées à la même discipline. De plus, une revue qui conserverait le même facteur d'impact d'une année à la suivante, alors que les autres revues de sa catégorie verraient le leur augmenter, pourrait voir son rang SIGAPS baisser mécaniquement. Or, comme l'a démontré une récente étude bibliométrique (Antonoyiannakis, 2020), les valeurs des facteurs d'impacts des revues sont très volatiles: en 2017, 10% de 11639 revues indexées dans le WoS devaient 50% de la valeur de leur facteur d'impact uniquement à leurs trois publications les plus citées. Ceci prouve que le facteur d'impact d'une revue ne reflète pas la qualité de l'ensemble de ses publications, mais est déterminé en grande partie par un faible nombre de publications qui reçoivent l'essentiel des citations de la revue et qui peuvent grandement faire varier la valeur du facteur d'impact d'une année à l'autre.

Les classements des revues peuvent donc changer annuellement, en fonction de la variation de leurs facteurs d'impacts qui sont tout aussi volatiles. Par exemple, une revue d'oncologie comme *Blood Cancer Journal* passe successivement en l'espace de 5 ans (de 2014 à 2018), du rang « C », au rang « B », au rang « A ». Inversement, durant la même période, la revue *Cancer Journal* passe du rang « B », au rang « C », au rang « D ». Sur les 258 revues qui définissent dans SIGAPS la spécialité « oncologie », 126 revues, soit 53%, ont changé au moins une fois de rang SIGAPS entre 2014 et 2018, parmi lesquelles 10 revues ont changé de deux rangs. Ces changements sont souvent associés à des variations très faibles des facteurs d'impact des revues d'une année à l'autre, variations parfois inférieures à 5%. C'est le cas, par exemple, de *Bone Marrow and Transplantation Research* qui est passée du rang « B » au rang « C » entre 2017 et 2018, pour une variation de son facteur d'impact de 4.497 à 4.674, ou de la revue *Cancer Biology and Medicine* qui est passée du rang « B » au rang « C » pour une variation de son facteur d'impact de 4.467 à 4.607. Or, il devrait être évident que cette variation infime n'est qu'un effet statistique aléatoire et nullement le signe d'une modification tendancielle de la « qualité » de la revue. Ces fluctuations reflètent aussi la fausse précision d'un indicateur curieusement calculé avec trois chiffres significatifs, ce que très peu de mesures en sciences biomédicales peuvent atteindre...

Nonobstant ces problèmes de fluctuation, il faut rappeler que l'utilisation du facteur d'impact des revues pour évaluer les publications des chercheurs constitue une erreur de méthode, car,

comme nous l'avons mentionné plus haut, cet indicateur ne caractérise pas directement l'article, mais bien la revue. Pour illustrer cette erreur d'objet de la mesure dans le système SIGAPS, prenons l'exemple de deux revues d'oncologie similaires: *European Journal of Cancer* et *Clinical Cancer Research*. En 2016, les CHU français ont publié 33 articles dans cette revue, classée « B », qui en 2017 avait un facteur d'impact de 7.191. Cela signifie que les articles parus dans cette revue en 2016 devaient avoir reçu en moyenne 7 citations à la fin de l'année 2017. Or, sur les 33 articles de CHU publiés en 2016 dans cette revue, la moitié (17) ont en fait reçu 9 citations ou plus à la fin de l'année 2017, soit un nombre de citations supérieur à la moyenne de la revue. Considérons à présent la revue *Clinical Cancer Research* qui était classée « A » en 2017, pour un facteur d'impact de 8.911. Les articles publiés dans cette revue en 2016 devaient donc avoir reçu 9 citations en moyenne, à la fin de l'année 2017. Dans les faits, en 2016, les CHU français y ont publié 17 articles, dont 6 articles ont reçu moins que 9 citations, soit un nombre inférieur à la moyenne de la revue.

Cette comparaison montre d'abord que les 17 articles sur 32 publiés dans *European Journal of Cancer*, revue classée « B », ont reçu plus de citations que la moyenne des articles publiés dans *Clinical Cancer Research*, revue pourtant classée « A ». La comparaison montre surtout que ces 17 articles publiés dans une revue classée « B » ont potentiellement reçu moins de points SIGAPS que les 6 articles de CHU publiés *Clinical Cancer Research* classée « A », qui ont pourtant reçu moins de citations. On voit bien par cet exemple que le système SIGAPS n'évalue pas la qualité réelle de chacune des publications des chercheurs, mais s'appuie uniquement sur la visibilité moyenne des revues dans lesquelles elles sont publiées. Or, les distributions des citations aux revues, tout comme d'ailleurs celles des publications des chercheurs et des octrois de recherche, suivent une courbe de type Pareto ou Lotka. Elles sont donc très concentrées, suivant à peu près la règle du 20/80 à l'effet que 80% des citations d'une revue proviennent de seulement 20% des articles qu'elle contient (Larivière et al., 2010). Une conséquence de ce fait général est que le facteur d'impact d'une revue, qui est une valeur moyenne, surestime le nombre réel de citations des articles des chercheurs dans cette revue. Surtout, il pénalise ceux dont les citations réelles reçues, et donc l'impact scientifique, sont largement supérieures à cette moyenne.

La question se pose donc: pourquoi prendre le facteur d'impact de la revue et non pas les citations réelles obtenues par les articles? La réponse est simple: mesurer l'impact *réel* des publications nécessite d'attendre qu'elles soient citées pendant 3 ou 4 ans, alors que le facteur d'impact calculé chaque année permet une évaluation immédiate, au principe qu'une publication dans une revue bien classée a une plus grande *probabilité* d'être citée qu'un article publié dans une revue moins bien classée. On évalue donc indirectement les chercheurs non pas sur l'impact *effectif*, mais sur l'impact *présumé* de leurs publications, ce qui est problématique au plan éthique quand on sait qu'en vertu de la loi de Lotka (un cas particulier d'une loi de puissance de style Pareto), les citations réelles obtenues par une publication dans une revue bien classée peuvent en fait être inférieures à celles reçues par article paru dans une revue moins bien classée dans SIGAPS. On pourrait bien sûr dire qu'en fait SIGAPS n'évalue pas vraiment les chercheurs, mais ce serait oublier que, in fine, si les points obtenus par une institution sont faibles, on se tournera inévitablement vers les chercheurs qui publient dans ces revues mal cotées. Il s'agit donc bien d'une évaluation, même indirecte des chercheurs. D'ailleurs, on verra plus loin une confirmation de cela dans l'usage que certains chercheurs font du classement SIGAPS des revues et que ceux-ci ont bien un effet aussi sur les embauches.

2.3 La langue de publication

Finalement, le système SIGAPS a eu un effet majeur sur l'évolution du nombre de publications en langue française des CHU au cours des 20 dernières années. On sait que, dans toutes les disciplines, la tendance à publier en anglais s'est accentuée depuis les années 1970 (Gingras, 1984, 2008), mais SIGAPS semble avoir accéléré cette tendance dans la recherche médicale française en milieu hospitalier, dont les revues médicales françaises sont pourtant un lieu de publication privilégié. La Figure 4 montre en effet qu'après être demeurée relativement stable entre 25 et 30% de 2000 à 2008, la part des publications de langue française des CHU a rapidement chuté par la suite pour atteindre 9% en 2018. Il est intéressant de comparer la tendance de cette courbe avec celle des publications médicales et biomédicales de l'INSERM et du CNRS. Dans ces deux institutions de recherche fortement insérées dans les réseaux de la recherche internationale, on observe que la part des publications en langue française était déjà très faible en 2000 (à 5%) et qu'elle a légèrement baissé depuis, oscillant

autour de 4% jusqu'en 2018. Cette comparaison fait donc clairement ressortir l'effet de SIGAPS sur la chute du nombre de publications en langue française des CHU. Ce phénomène s'explique par le simple fait que la plupart des revues médicales ou biomédicales françaises ou francophones, bénéficiant de peu de visibilité internationale, ont un facteur d'impact faible et sont donc le plus souvent classées dans les catégories « E » ou « D » par SIGAPS. Cela pousse naturellement les chercheurs des CHU à privilégier les revues de langue anglaise qui monopolisent les catégories supérieures de SIGAPS. Une étude menée sur les données de publications du CHU de Rouen, entre 2007 et 2010, concluait d'ailleurs que « plus les auteurs sont prolifiques (score SIGAPS élevé), moins ils ont tendance à publier en français; publier en français ne semble donc pas être le meilleur moyen d'avoir un score SIGAPS élevé » (Griffon et al., 2012).

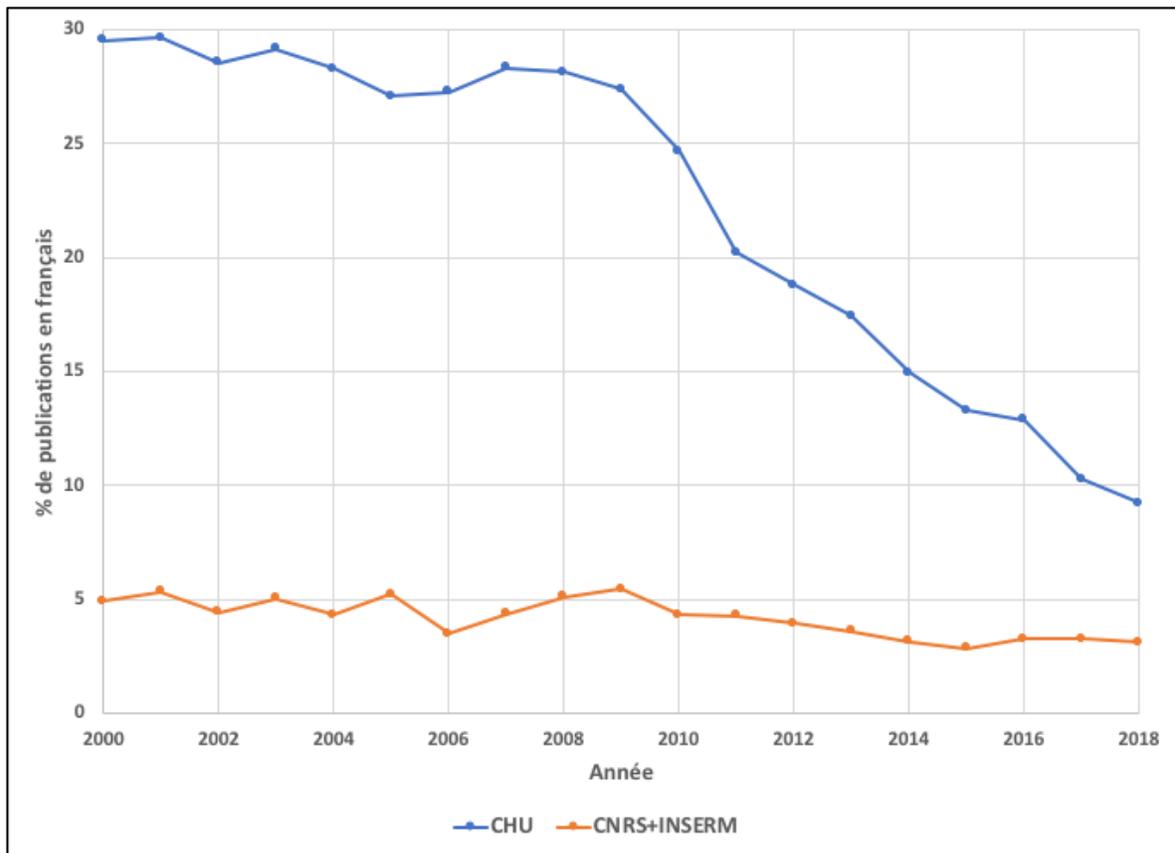


Figure 4 – Évolution de la part des publications de langue française des CHU

Un rapport de la Haute autorité de santé (HAS) relevait aussi cette situation dès 2013, faisant remarquer que « plusieurs revues ont récemment changé de titre et de langue (par exemple

la revue *Gastroentérologie Clinique et Biologique*, devenue en janvier 2011 *Clinics and Research in Hepatology and Gastroenterology*) » (HAS, 2013, p. 27). Cette tendance à l'anglicisation ou au bilinguisme des revues médicales françaises leur permet d'intégrer le *Journal Citations Report* publié annuellement par Clarivate Analytics et d'obtenir ainsi un facteur d'impact censé leur octroyer une plus grande visibilité internationale (Laccourreye, Huy et Maisonneuve, 2020). Or, les données de téléchargements institutionnels d'articles au sein de l'AP-HP, le plus grand CHU de France, indiquaient que sur les 15 revues les plus consultées en 2010 par les chercheurs, les praticiens et les étudiants, 8 étaient de langue française. Le rapport concluait ainsi que: « les auteurs d'articles scientifiques et de formation sont peu nombreux et s'orientent vers des publications en langue anglaise, tandis que la plupart des 200 000 médecins français ont besoin de publications en langue française pour s'informer et se former ». Il ajoutait que « la diminution des publications originales dans des revues françaises est un frein à la diffusion de l'information et de la connaissance médicales parmi les praticiens, car un certain nombre d'entre eux ne maîtrisent pas suffisamment la langue anglaise pour une lecture aisée et fine du contenu des articles » (HAS, 2013, p. 28). Cette chute rapide des publications médicales de langues française pourrait être d'autant plus dommageable qu'une enquête récente menée auprès des chercheurs de l'AP-HP établissait que 40% d'entre eux identifiaient la difficulté à maîtriser la langue anglaise comme un obstacle à la publication de leurs recherches (Duracinsky, Péretz et Chassany, 2017).

2.4 Choix des co-auteurs

Une autre façon d'accroître la « productivité » d'un organisme de recherche est de lui adjoindre des chercheurs extérieurs sous forme de contrat spécifique. Cette stratégie a d'abord été utilisée par l'Université saoudienne King Abdulaziz pour améliorer sa position dans le classement de Shanghai des universités. Pour ce faire, l'université offrait des contrats lucratifs à certaines « stars » mondiales de la recherche, identifiées comme étant parmi les plus cités. En échange, les chercheurs devaient seulement ajouter l'adresse de l'Université King Abdulaziz dans leurs publications. Ce simple geste bibliographique augmentait mécaniquement la position de l'université saoudienne, sans qu'elle n'ait vraiment à créer ses propres laboratoires (Gingras, 2014a).

En France, on constate également que les CHU et autres institutions éligibles ont effectivement largement augmenté leur volume de publications depuis la mise en place du système SIGAPS, en recourant parfois aux mêmes stratégies que l'université d'Arabie saoudite. En effet, un rapport de la Cour des comptes portant sur le rôle des CHU dans l'enseignement supérieur et la recherche médicale notait en 2018 que : « certains établissements de santé contournent le modèle en salariant pour quelques vacations, parfois une seule, des praticiens attachés, par ailleurs auteurs de publications dans le cadre de l'activité de recherche menée auprès de leur employeur principal [...] L'indicateur obtenu de cette manière est déconnecté de l'établissement où se déroule effectivement la recherche en question » (Milon, 2018, p. 88). Moyennant rémunération, un établissement peut ainsi s'attacher contractuellement les services d'un chercheur pour peu qu'il rajoute le nom du CHU aux papiers qu'il signe dans le cadre de recherches qu'il mène dans sa véritable institution de rattachement. Le rapport de la Cour des comptes confirme d'ailleurs cela en soulignant que les CHU peuvent s'attribuer les publications « des chercheurs INSERM et CNRS qui disposent d'un contrat d'interface, à partir du moment où ceux-ci sont salariés de l'établissement de santé » (Milon, 2018, p. 88).

3. Les dérives du système SIGAPS

En plus des effets prévisibles sur le comportement des chercheurs que nous venons d'analyser, un système d'évaluation de la recherche adossé à des indicateurs inadéquats, et associés à des octrois budgétaires, risque fort d'engendrer des comportements déviants, des manipulations et des magouilles. On peut en effet se demander si la recherche scientifique garde son caractère éthique quand des chercheurs ou des laboratoires visent consciemment à publier un grand nombre d'articles dans des revues de faible qualité ou en libre accès, afin d'accumuler rapidement des points, plutôt que de viser une seule publication originale et de grande qualité dans une revue plus sélective, démarche cependant plus longue et au succès plus aléatoire, au vu du fort taux de refus de ces revues. La critique du « publish or perish » est bien sûr ancienne, mais elle est exacerbée par les nouvelles méthodes d'évaluation.

Par exemple, le Dr Didier Raoult a beaucoup publié avec son équipe dans la revue *New Microbes and New Infections* contrôlée par des chercheurs de son institution (Gingras et Khelifaoui, 2020), et classée au plus bas dans SIGAPS (« NC »). On y retrouve 77 articles pour la

seule année 2017, soit un rythme d'une publication tous les 5 jours environ. Ces articles valent ainsi, selon SIGAPS, 308 points, soit près de 200 000 euros qui, versés durant 4 ans à l'AP-HM, rapporteront au total 800 000 euros à l'institution. Même si un tel montant ne représente qu'un pourcentage minime du budget total de l'institution, il demeure élevé en regard des projets de recherche qu'il rend possibles. Par comparaison, cela correspond au nombre de points que rapporterait la publication, en une seule année, de près de 10 publications dans des revues classées « A », comme *Nature*, *Science*, *Journal of the American Medical Association* ou *The Lancet*. Cela étant impossible, même pour les meilleurs chercheurs, on comprend facilement, et sans faire de grands calculs fondés sur la théorie des jeux, que la stratégie optimale est bien de multiplier les publications dans les revues en libre accès qui publient rapidement n'importe quel résultat. Ce n'est d'ailleurs pas un hasard si une revue comme *Oncotarget* s'est faufilée pendant quelques années dans ces classements. Il est d'ailleurs tout à fait plausible de penser que la multiplication de ces revues s'explique en bonne partie par les pressions à publier à tout prix, pour répondre aux exigences de systèmes d'évaluation fondés sur le nombre de publications (Biagioli et Lippman, 2020).

La Cour des comptes ne s'y est pas trompée en notant que : « L'indicateur SIGAPS n'apparaît pas suffisamment discriminant vis-à-vis de la qualité des revues de publication (périmètre des publications retenues, intérêt des revues dites « non classées » dans l'établissement du score, pondération). Cette situation peut conduire à un dévoiement du modèle par des pratiques contestables (course à la publication, y compris d'articles de niveau insuffisant) » (Milon, 2018, p. 89). D'ailleurs, dès 2008, une revue médicale à faible impact avait flairé le bon filon, puisqu'elle encourageait les chercheurs à y publier des articles afin d'accumuler facilement des points SIGAPS. L'un de ses éditoriaux affirmait ainsi qu'« un article publié dans notre revue peut donc rapporter entre cinq et deux points en fonction du rang de signataire de l'auteur : c'est donc bien le moment de publier dans les *Annales de pathologie* ! » (Sabourin et Darmoni, 2008). Mais un laboratoire semi-privé qui commercialise un logiciel de bio-statistiques a su se montrer encore plus opportuniste, en vantant le fait qu'un article publié dans une revue classée « B » par SIGAPS, ayant rapporté 60 000 euros à l'un de ses CHU client, avait été réalisé à l'aide de son programme informatique. La firme ajoutait avec fierté que grâce à son logiciel: « l'analyse statistique a été réalisée en une journée, permettant

d'augmenter la productivité d'une cohorte de quelques publications par an », ce qui justifiait amplement « le prix de 5 000 euros HT par an de sa licence »⁶.

Une autre dérive prévisible générée par SIGAPS est le fait que les CHU commencent à utiliser ce système pour évaluer individuellement leurs chercheurs, en particulier dans les processus d'embauche et de promotion. Ainsi, en 2016, le responsable de la Délégation à la Recherche Clinique et à l'Innovation du CHU de Clermont-Ferrand, Patrick Lacarin, indiquait que pour prétendre à la position de professeur des universités-praticien hospitalier (PU-PH) dans son institution, un chercheur devait avoir accumulé au moins 500 points SIGAPS lors des dix années précédentes, critère validé par une recommandation de la sous-section santé du Conseil national universitaire⁷. Le rapport de la Cour des comptes notait de son côté que : « l'université Paris-Sud procède à des auditions des PU-PH tous les cinq ans en prenant en compte pour la recherche le score SIGAPS » (Milon, 2018, p. 131). En 2016, le Collège national des universitaires de psychiatrie (CNUP) a également recommandé l'intégration de critères liés au score SIGAPS et à l'« indice h » dans les processus de nomination aux postes de professeur des universités (PU) et de maître de conférence des universités-praticien hospitalier (MCU-PH) en psychiatrie adulte. On pouvait lire ainsi dans les critères d'évaluation que: « les candidats peuvent argumenter leur niveau de publications en fournissant au jury leur index H et/ou leur score SIGAPS ». Les critères d'évaluation étaient très précis, puisque le document ajoutait que: « 5 publications dans des revues de bon niveau ($IF \geq 2$) seront nécessaires aux MCU dont au moins 2 en premier ou deuxième auteur. Pour les candidats à un poste de PU, on attend au moins 10 publications en 1^{er}, 2^e ou dernier auteur dans une revue à impact factor ≥ 2 »⁸.

Ces cas sont loin d'être anecdotiques, puisqu'en 2016, un article de la revue *Bulletin du Cancer* présentait les critères minimaux, incluant les scores SIGAPS, requis par la sous-section cancérologie du CNU pour candidater aux postes de PU ou de MCU dans cette spécialité.

⁶<https://www.labcom-risca.com/single-post/Points-SIGAPS-comment-sont-valorises-vos-publications-scientifiques->

⁷<https://alambic.hypotheses.org/tag/sigaps>

⁸ https://cnup.unistra.fr/fileadmin/upload/DUN/cnup/Documents/CRITERES_de_CANDIDATURE_A_LA_NOMINATION.pdf

Pour les postes de PU, il fallait, au cours des 5 dernières années, avoir au moins 10 publications dans une revue à facteur d'impact supérieur à 3, dont 5 articles originaux en tant que premier, deuxième ou dernier auteur dans une revue à facteur d'impact élevé (supérieur à 5 pour l'oncologie médicale, à 4 pour l'oncologie radiothérapie, à 3 pour la chirurgie ou la biologie). En plus de cela, les candidats devaient « démontrer une dynamique croissante de publications » et un score SIGAPS supérieur à 400 (Soria et al. 2016, p. 726). Des critères de nature semblable, mais avec des seuils minimaux inférieurs, étaient exigés pour les postes de MCU.

En 2018, le CHU de Dijon est allé jusqu'à afficher sur son site le classement de ses chercheurs selon leurs scores SIGAPS accumulés durant la période 2013-2016. En plus d'instaurer un climat de compétition malsain entre les chercheurs, le système SIGAPS crée également une fausse compétition au sein et entre les institutions hospitalières. Par exemple, le même CHU de Dijon a instauré depuis 2016, un système d'« intéressement » qui reverse 300 000 euros sur les sommes reçues par l'intermédiaire du financement SIGAPS à ses laboratoires, au prorata de leurs contributions en termes de points.⁹ Autre exemple de compétition induite entre institutions, une revue interne du CHU de Nantes notait en 2014 que : « pour rejoindre le score du CHU de Toulouse, il faudrait augmenter le nombre de publications de plus d'une centaine d'articles par an et augmenter également le taux de publications de rang A » (Brisson, 2014, p. 4). Ainsi, la politique de recherche des établissements est définie, en partie du moins, en fonction d'une course à l'accumulation de « points » qui visent à augmenter des « scores » qui ne disent eux-mêmes pas grand-chose de la qualité et de l'originalité réelle de la production scientifique des CHU.

Les points attribués étant liés à la position des auteurs dans la liste des cosignataires, on peut également prédire que des tensions et conflits seront générés entre collaborateurs par un tel système, surtout pour les jeunes chercheurs qui doivent cumuler des points pour pouvoir candidater sur des postes. Mais seuls des entretiens avec ces chercheurs pourraient confirmer cette hypothèse. La rapide intériorisation de ces normes, imposées sans débat collectif, s'observe déjà dans les CV des chercheurs qui indiquent non seulement le facteur d'impact des

⁹<https://www.chu-dijon.fr/fr/actualites/2018/02/16/publication-points-sigaps-chercheurs-chu>

revues dans lesquelles ils ont publié, mais aussi leur rang SIGAPS qui, contrairement au facteur d'impact, n'a d'existence qu'en France.

Conclusion : Éthique et dérive systémique

Au-delà du cas particulier du système SIGAPS dont nous avons analysé en détail les caractéristiques et les effets mesurables, les partisans de l'application généralisée de l'évaluation bibliométrique aux chercheurs font généralement valoir que celle-ci constitue un critère de jugement impartial, qui élimine les dérives liées au clientélisme et au copinage. Or, force est de constater que l'instauration de ces méthodes d'évaluation basées sur des critères comptables supposément objectifs, loin de régler ces problèmes, peut en fait les amplifier, et même en créer de nouveaux, comme nous l'avons amplement montré dans cette étude.

Par ailleurs, les problèmes générés par le logiciel SIGAPS évoqués récemment dans les médias français (Korda et Simon, 2020; Roy, 2020; Larousserie, 2020) semblent être analysés au seul prisme de l'éthique individuelle des chercheurs qui n'ont pourtant fait que s'adapter aux « règles du jeu ». Un tel raisonnement fait oublier que les éventuels écarts individuels à l'éthique, qu'il ne s'agit pas ici de défendre, ne sont en fait que le symptôme du fonctionnement même d'un système de financement de la recherche qui contient en lui les germes du problème. Pointer vers les dérives individuelles ne mettra malheureusement pas fin aux tentatives de certains chercheurs de détourner le système à leur profit, tant que celui-ci leur permettra de le faire en toute légalité.

On peut également se demander, en terminant, s'il est éthique d'évaluer des chercheurs, même de façon indirecte, sur la base d'un système aussi problématique. En effet, on peut penser qu'il est injuste d'évaluer des personnes sur la base d'indicateurs qui ne mesurent pas directement l'objet à évaluer. Comme on l'a vu, le facteur d'impact ne caractérise pas un article publié, mais la notoriété de la revue dans laquelle il est publié. De plus, la valeur servant au classement fluctue de manière aléatoire d'une année à l'autre. Comme on l'a montré, un chercheur peut voir son article beaucoup plus cité que la moyenne représentée par le facteur d'impact de sa revue, alors qu'un autre classé au même rang sera au contraire beaucoup moins cité. Or, ces deux personnes sont, dans SIGAPS, considérées comme « égales »

pour ce qu'elles « rapportent » à leur institution. On peut aussi se questionner sur le caractère éthique d'un système qui amène, consciemment ou non, les chercheurs et les gestionnaires de la recherche à adapter leurs pratiques pour optimiser les points SIGAPS plutôt que la recherche de qualité. De ce point de vue, on peut trouver curieux que ce soit la Cour des comptes et non pas les chercheurs, leurs sociétés savantes ou les Conseils scientifiques de leurs établissements qui aient le mieux décrit les problèmes systémiques de ce système d'évaluation.

Enfin, il faut souligner que la pression engendrée par ce système comptable s'exerce sur de jeunes chercheurs qui n'ont d'autre choix que de se plier à ses exigences pour pouvoir ne serait-ce qu'espérer obtenir un poste. Ce faisant, ils se voient forcés de troquer l'objectif de produire des connaissances pour celui, bureaucratique, de produire des « unités comptables » (Gingras, 2018) au service des budgets de leurs centres de recherches et peut-être parfois de quelques primes à la « productivité » venant compenser de maigres salaires.

Bibliographie

- Angermuller, J. et T. van Leeuwen. (2019). « On the social uses of scientometrics: The quantification of academic evaluation and the rise of numerocracy in higher education », dans R. Scholz (dir.), *Quantifying Approaches to Discourse for Social Scientists*, London, Palgrave, p. 89-119.
- Antonoyiakinnis, M. (2020). « Impact factor volatility due to a single paper : A comprehensive analysis », *Quantitative Science Studies*, vol. 1, no. 2, p. 639-663.
- Baccini, A., De Nicolao, G. et E. Petrovich. (2019). « Citation gaming induced by bibliometric evaluation: A country-level comparative analysis », *PLoS ONE* 14(9): e0221212. <https://doi.org/10.1371/journal.pone.0221212>
- Barats, C., Bouchard, J. et Haakenstaaad A. (dirs.). (2018). *Faire et dire l'évaluation. L'enseignement supérieur conquis par la performance*, Paris, Presses des mines.
- Baudoin, L. et P. Devos. (2016). *Analyse bibliométrique des publications des CHU, 2005-2014*, Rapport du Comité national de coordination de la recherche, 40 p.
- Beall, J. (2012). « Predatory Publishers are Corrupting Open Access », *Nature*, vol. 489, p. 179.
- Beall, J. (2017). « What I learned from Predatory Publishers », *Biochemia Medica*, vol. 27, no. 2, p. 273-279.
- Biagioli, M. et A. Lippman (dirs.) (2020). *Gaming the Metrics. Misconduct and Manipulation in Academic Research*, Cambridge (Mass.), MIT Press.
- Blanc, E., Bouliou, R., et D. Bertram. (2019). « Bibliometrics and French healthcare institutions from 2004 to 2014 », *Bulletin du Cancer*, vol. 106, no. 3, p. 226-236.
- Butler, L. (2003). « Explaining Australia's increased share of ISI publications—the effects of a funding formula based on publication counts », *Research Policy*, vol. 32, no. 1, p. 143-155.
- Brisson, I. (2014). « Le CHU de Nantes dans le paysage SIGAPS national », *R de recherche*, no. 16, p. 3-5.
- Fanelli, D. (2020). « Pressure to Publish: What Effects Do We See? », dans Biagioli et Lippman (dirs.), *Gaming the Metrics. Misconduct and Manipulation in Academic Research*, Cambridge (Mass.), MIT Press, p. 111-122.
- Garcia, V. (2020). « L'astuce du Pr. Raoult pour publier massivement des études et décrocher des financements », *L'Express*, 30 juin.

Gingras, Y. (1984). « La valeur d'une langue dans un champ scientifique », *Recherches sociographiques*, vol. 25, no. 2, p. 286-296.

Gingras, Y. (2008). « Les langues de la science: Le française et la diffusion de la connaissance », dans Jacques Mauret et al. (dirs), *L'avenir du français*, Paris, Éditions des archives contemporaines, p. 95-97.

Gingras, Y. (2014a). *Les dérives de l'évaluation de la recherche. Du bon usage de la bibliométrie*, Paris, Raisons d'Agir.

Gingras, Y. (2014b). « How to boost your university up the rankings », *University World News*, 18 juillet.

<https://www.universityworldnews.com/post.php?story=20140715142345754>

Gingras, Y. (2018). « Les transformations de la production du savoir: de l'unité de connaissance à l'unité comptable », *Zilsel. Science, technique, société*, vol. 2, no. 4, p. 139-152.

Gingras, Y. et M. Khelifaoui. (2020). « Être juge et partie, ou comment contrôler une revue scientifique », *The Conversation*, 26 juin.

<https://theconversation.com/etre-juge-et-partie-ou-comment-controler-une-revue-scientifique-140595>.

Griffon, N., Devos, P., Gehanno, J.-F., et J.S. Darmoni. (2012). « Existe-t-il une corrélation entre le score SIGAPS et les publications en français ? », *La Presse Médicale*, vol. 41, no. 9, p. e432-e435.

Larivière, V., Macaluso, B., Archambault, E., et Y. Gingras (2010). « Which scientific elites? On the concentration of research funds, publications and citations », *Research Evaluation*, vol. 19, no 1, p. 45-53

Larivière, V. et C. Sugimoto. (2018). *Mesurer la science*, Montréal, Presses de l'Université de Montréal.

Larousserie, D. (2020). « 'Lancetgate' : publier beaucoup dans des revues scientifiques peut rapporter gros », *Le Monde*, 15 juin.

https://www.lemonde.fr/sciences/article/2020/06/15/lancetgate-publier-beaucoup-dans-des-revues-scientifiques-peut-rapporter-gros_6042952_1650684.html

Haute autorité de santé (2013). *Bonnes pratiques et critères de qualité des revues et journaux de la presse médicale française*, 237 p.

Hvistendahl, M. (2013). « China's publication bazaar », *Science*, vol. 342, no. 6162, p. 1035-1039.

Khelifaoui, M., Larrègue, J., Larivière, V. et Y. Gingras. (2020). « Measuring National Self-Referencing Patterns of Major Science Producers », *Scientometrics*, vol. 123, no. 2, p. 979-996.

Korda, R. et C. Simon (2020). « Didier Raoult, une frénésie de publications et des pratiques en question », *Le Parisien*, 12 juin. <https://www.leparisien.fr/societe/didier-raoult-une-frenesie-de-publications-et-des-pratiques-en-question-12-06-2020-8334405.php>

Laccourreye, O., Tran Ba Huy, O. et H. Maisonneuve (2020). « Les journaux médicaux scientifiques en langue française face aux défis du XXI^e siècle », *Bulletin de l'Académie Nationale de Médecine*, vol. 204, p. 410-415.

Mallapty, S. (2020). « China bans cash rewards for publishing papers », *Nature*, vol. 579, no. 7797, p. 18.

Mancini, J., Darmoni, S., Chaudet, H. et M. Fieschi (2009). « Le paradoxe de la 'T2A bibliométrique' SIGAPS : un risque d'effet délétère sur la recherche française ? », *Presse Médicale*, vol. 38, no. 2, p. 174-176.

Milon, A. (2018). « Rapport d'information fait au nom de la commission des affaires sociales (1) sur l'enquête de la Cour des comptes relative au rôle des centres hospitaliers universitaires dans l'enseignement supérieur et la recherche médicale ». <https://www.se-nat.fr/rap/r17-228/r17-2281.pdf>

Nicholas, D., et. al. (2020). « Millennial researchers in a metric-driven scholarly world: An international study », *Research Evaluation*, vol. 29, no. 3, p. 263-274.

Nicholas, D., et. al. (2018). « Early Career Researcher's Quest for reputation in the digital age », *Journal of Scholarly Publishing*, vol. 49, no. 4, p. 375-396.

Pontille, D. (2016). *Signer ensemble. Contribution et évaluation en sciences*, Paris, Economica.

Raoult, D. (2015). « Raoult : publications scientifiques, place à la transparence ! », *Le Point*, 16 mars. https://www.lepoint.fr/invites-du-point/didier_raoult/raoult-publications-scientifiques-place-a-la-transparence-16-03-2015-1913159_445.php

Roy, S. (2020). « Revues scientifiques: quand publier rapporte des budgets de recherche », *Le Figaro*, 16 juin. <https://www.lefigaro.fr/sciences/revues-scientifiques-quand-publier-rapporte-des-budgets-de-recherche-20200616>

Sabourin, J.-C. et S. Darmoni (2008). « Avec le score SIGAPS : c'est le moment de publier dans les *Annales de pathologie* ! », *Annales de Pathologie*, vol. 78, no. 2, p. 77-78.

Soria, J.-C. et al. (2016). « Le Conseil national des universités en cancérologie-radiothérapie : missions et critères de sélection présentés aux lecteurs du *Bulletin du Cancer* », *Bulletin du Cancer*, vol. 103, no. 9, p. 719-729.



Autres titres de cette collection

-
- | | |
|---------|--|
| 2020-04 | Gingras, Yves et Mahdi Khelfaoui
« Academic debates and the complexity of the hydroxychloroquine controversy » |
| 2020-03 | Khelfaoui, Mahdi et Yves Gingras
« Branding Scholarly Journals: Transmuting Symbolic Capital into Economic Capital » |
| 2020-02 | Dandurand, Guillaume , François Claveau , Jean-François Dubé et Florence Millerand
« AI Like Any Other Technology: Social Dynamics of Expectation and Expertise of a Digital Humanitarian Innovation » |
| 2020-01 | Talin, Kristoff et Yves Gingras ,
« + de religion = - de science » |
| 2018-01 | Carlier, Denis
« Numérisation et analyse de documents » |
| 2015-01 | Dias Da Silca, Patricia et Lorna Heaton
« Citizens, amateurs, volunteers: Conceptual struggles in studies of citizen science » |
| 2014-03 | Hanel, Petr
« Is China catching-up human health-related applications of biotechnology ? » |
| 2014-02 | Maroy, C. , P. Doray , M. Kabore
« La politique de financement des universités au Québec à l'épreuve du « Printemps érable » » |
| 2014-01 | Bastien, N. , P. Chenard , P. Doray , B. Laplante
« Économie, société et éducation: l'effet des droits de scolarité sur l'accès aux études universitaires au Québec et en Ontario » |
| 2013-03 | Hanel, Petr , Jie He , Jingyan Fu , Jorge Niosi et Suzan Reid
« A romance of the three kingdoms and the tale of two cities: the role and position of the biotechnology industry cluster in Guangdong province, China » |
| 2013-02 | Gauthier, Elisabeth , Gale E. West et Anne-Marie Handfield
« Why do humans need to do battle? Social representations of alternative pest control approaches » |
| 2013-01 | Bastien, Nicolas , Pierre Chenard , Pierre Doray et Benoit Laplante
« L'accès à l'université: le Québec est-il en retard? » |
| 2012-01 | Prud'homme., Julien , Yves Gingras , Alain Couillard et Daniel Terrasson
« Les mesures de l'interdisciplinarité. Pratiques et attitudes dans un centre de recherche français : l'IRSTEA » |
| 2011-02 | Verdier , Éric , Pierre Doray et Jean-Guy Prévost
« Régionalisation et recomposition du travail statistique : esquisse d'une comparaison France-Québec » |
| 2011-01 | Mayer, Leticia
« PROBABILISM. A Cultural environment that led to the creation of random probability? » |



Centre interuniversitaire
de recherche sur la science
et la technologie

Le CIRST est, au Canada, le principal regroupement interdisciplinaire de chercheurs dont les travaux sont consacrés à l'étude des dimensions historiques, sociales, politiques, philosophiques et économiques de l'activité scientifique et technologique.

Nos travaux visent l'avancement des connaissances et la mise à contribution de celles-ci dans l'élaboration et la mise en œuvre des politiques ainsi que dans la résolution des problèmes de société qui présentent des dimensions scientifiques et technologiques.

Le CIRST rassemble une quarantaine de chercheurs provenant d'une dizaine d'institutions et d'autant de disciplines, telles que l'histoire, la sociologie, la science politique, la philosophie, les sciences économiques, le management et les communications.

Le CIRST fournit un milieu de formation par la recherche à de nombreux étudiants de cycles supérieurs dans les domaines de recherche de ses membres. Créé en 1986, il est reconnu par l'Université du Québec à Montréal, l'Université de Montréal, l'Université Laval, Polytechnique, l'Université de Sherbrooke et l'Université TÉLUQ. Le CIRST est un regroupement stratégique du Fonds de recherche du Québec—Société et culture.