

[affiliations](#) [1], [Atelier bibliométrie](#) [2], [co-signatures](#) [3], [Manuel Durand-Barthez](#) [4], [mentions d'affiliation](#) [5], [publications scientifiques françaises](#) [6], [scientométrie](#) [7], [signatures](#) [8], [INSERM](#) [9], [modèles](#) [10]

Groupe de réflexion sur l'analyse et la caractérisation de la production scientifique – *Outils et Méthodes*

Dixième réunion : 16 février 2012

Compte-rendu élaboré par Manuel DURAND-BARTHEZ

La bibliométrie comme outil d'évaluation et de pilotage à l'INSERM

Conférence de Madame Nicole HAEFFNER-CAVAILLON, DR1 INSERM, responsable de la Cellule Mesures, Indicateurs, Bibliométrie, Département de l'Evaluation et du Suivi des Programmes (DESP)

Biologiste de formation, Docteur ès-Sciences (spécialité : Immunologie), N. Haeffner-Cavaillon a travaillé sur le VIH jusqu'en 2000. Claude Griscelli, alors directeur de l'INSERM, lui proposa de participer à un groupe de travail destiné à réfléchir sur une possible réforme de l'évaluation à l'INSERM. L'approfondissement de cette réflexion l'a conduite à s'investir de manière significative dans l'adaptation des méthodes bibliométriques à l'évaluation de la recherche à l'Institut, jusqu'à assumer la responsabilité de la Cellule chargée de ces études. Ses compétences scientifiques et son statut de chercheur ont grandement et positivement contribué à l'efficacité de la méthodologie qu'elle a mise en œuvre.

La bibliométrie : une aide à l'évaluation et un outil de stratégie

Dans le contexte de l'INSERM, la bibliométrie s'applique uniquement à des chercheurs confirmés : Concours DR (217 ch.), Promo DR (69 ch.) et Prime d'excellence scientifique « PES » (189 candidats acceptés).

Elle constitue également un outil de stratégie ; on examinera des exemples d'application d'indicateurs sur les ITMO (instituts thématiques multiorganismes) et sous forme de *benchmarks* visant des Instituts d'excellence.

Dans le cadre de l'évaluation, cela représente en 2011 : 279 équipes comportant en moyenne 5 chercheurs (donc > 1000 ch.)

Elle constitue également un outil de stratégie ; on examinera des exemples d'application d'indicateurs sur les ITMO (instituts thématiques multiorganismes) et sous forme de *benchmarks* visant des Instituts d'excellence.

Dans le cadre de l'évaluation, cela représente en 2011 : 279 équipes comportant en moyenne 5 chercheurs (donc > 1000 ch.)

Les méthodes appliquées sont particulièrement adaptées au domaine biomédical. Elles ne peuvent s'étendre à d'autres domaines que si d'autres indicateurs sont utilisés ou si l'on applique à ceux-ci des modifications substantielles. S'exprime en l'occurrence le souci de ne pas créer une polémique de type : Bibliométrie vs. Evaluation par les pairs ou bien : Métrie Sciences Sociales vs. Biomédicales.

La mesure de l'activité scientifique s'effectue à partir de notices bibliographiques répertoriées dans les bases de données. Pour l'INSERM, il s'agit du *Web of Science* (ci-après noté WoS) de Thomson-Reuters, dont les indicateurs permettent d'analyser l'information bibliographique.

Cette analyse permet de définir des repères de type socio-économique. Elle met en évidence le critère de visibilité dans le contexte national et international (pays, organismes, chercheurs). Elle est fondamentale vis-à-vis de toute opération d'évaluation et facilite la veille.

Champs des notices : outre les champs classiques, il est particulièrement fait mention des *Acknowledgements* et du support financier (*Grants*) qui autorisent l'identification d'éléments utiles à la connaissance du contexte dans lequel une publication a pu être commise.

Quatre bases sont mentionnées pour mémoire : *Pascal* (INIST-CNRS) qui collationne 18 millions de références et 3260

revues internationales ; *Scopus* d'Elsevier avec 18 000 titres recouvrant 40 millions de références ; le WoS répertoriant un nombre identique de références (issues de 10 000 titres de périodiques + 110 000 actes de colloques) et *PubMed* avec 19 millions de références issues de 5420 titres de périodiques.

Scopus compense sa jeunesse relative par rapport au WoS en introduisant un nombre significatif de journaux issus du domaine des Sciences sociales. La philosophie du WoS consiste à se concentrer sur les journaux offrant le plus de visibilité, qui citent d'autres journaux remarquables, lesquels sont eux-mêmes les plus cités. La focalisation de cette base sur les journaux qui ont un *Impact factor* (ci-après noté I.F.) est une de ses caractéristiques fondamentales. On doit noter qu'en revanche, *Scopus* a ouvert sa base à des journaux qui n'ont pas de système de *reviewing* (environ 1500 sur 18 000). Dans ces deux bases de données, on assiste à l'introduction progressive des ouvrages. Outre les 10 000 journaux qu'inclut le WoS, Thomson y prévoit l'introduction d'environ 30 000 ouvrages.

Dans la suite, on examine les différences de résultats portant sur un mot-clé : « innovation ». La question est successivement posée sur *Scirus* (produit d'appel gratuit d'Elsevier en amont de *Scopus* : 6 927 527 réponses incluant des brevets et des « pages Web » à dominante scientifique) ; *PubMed* (33 393 rép.) *Google Scholar* (2 340 000 rép.) et le WoS (92 701 rép.)

Chacune de ces bases, en fonction des réservoirs qu'elles interrogent, vont présenter des caractéristiques différentes : selon l'analyse que l'on souhaite entreprendre, il conviendra de choisir telle ou telle manière appropriée.

Le WoS est la base de référence à l'INSERM pour l'évaluation.

ur les trente années 1981-2010 consultées sur le WoS, on assiste à un accroissement considérable du nombre de journaux. Fait remarquable : le nombre de publications européennes est resté inférieur à celles issues des Etats-Unis jusqu'en 1995. La Communauté européenne de l'époque dépasse alors les USA jusqu'à ce jour avec respectivement, en 2010, 36 et 28% de la production mondiale (« U.E.-27 » vs. Etats-Unis). Celle de la Chine est en très forte augmentation ces dernières années (13,4% en 2010) tandis que la France plafonne présentement à 5,5%.

En considérant, toujours à partir du WoS, les pourcentages de citations recueillies, on constate que la Chine est en très forte progression, avec des taux devenus à peu près similaires à ceux dont bénéficient les autres pays du peloton de tête, depuis 2009 environ.

Entre 80 et 85% des publications mondiales ont reçu au moins une citation.

La répartition par domaines en 2010 fait apparaître en premier la recherche médicale avec 20% des publications, puis : Chimie (10%), Physique et Ingénierie (à égalité avec 8%, la Physique ayant un peu baissé depuis 2000), Biologie/Biochimie (environ 5%).

Depuis quelques années, on constate dans le WoS une forte augmentation des Sciences sociales, potentiellement due à la stratégie de *Scopus* qui a incité son concurrent Thomson à accroître sa part dans ce domaine. Cela ne signifie pas pour autant que l'on publie plus dans le domaine des Sciences sociales. L'outil auxiliaire du WoS, *InCites*, découpe la recherche en 22 domaines, chaque journal n'étant classé que dans un seul. Plus la communauté est grande et plus l'article a des chances d'avoir été remarqué, lu et cité. La visibilité et le nombre de citations sont proportionnels à la taille de la communauté destinataire.

La répartition *des journaux* par domaines dans le WoS attribue 19% à la médecine, 14% aux Sciences sociales, mais la répartition *des publications* par domaines dans le WoS indique que les Sciences sociales ne représentent que 5%

Les deltas plus ou moins grands observés dans ces résultats sont typiques de différences liées notamment à la taille des communautés, aux usages de ces communautés, et au nombre de journaux. En Sciences sociales, on remarque une très forte spécialisation, et partant un grand nombre de journaux, dans lesquels ne sont publiés qu'un petit nombre d'articles. Ce sont là, par conséquent, des spécificités différentes de celles des domaines de la recherche biomédicale ou en sciences physiques. Il est donc nécessaire de bien connaître tous ces paramètres avant de chercher à utiliser des indicateurs. De fait, indubitablement, trouver une information pertinente dépend de la bonne interprétation de toute une série de paramètres. D'autant plus que, depuis quelques années, le WoS engrange environ un million de références supplémentaires par an.

En poursuivant l'interrogation sur le mot-clé « Innovation » dans le WoS, on voit dans les options consultables en marge gauche des résultats, que le terme est utilisé essentiellement dans le domaine Economie/Business, excessivement peu

dans le domaine biomédical. On voit l'importance des Abstracts (très utiles aux chercheurs parce que l'information arrive plus vite par le biais des abstracts que dans l'article publié proprement dit), les journaux majoritaires, les principaux auteurs, les institutions présentes.

Relative exactitude des références : dans une base de ce type, quelle qu'elle soit, tant le WoS que *Scopus*, une partie des données est scannée, l'autre saisie manuellement. Mais la reprise manuelle est systématique.

Si l'analyse est à l'échelle d'un pays, une ou deux lacunes ne sont pas graves. Si, en revanche, c'est au niveau d'un chercheur de l'institution, une ou deux erreurs peuvent être plus graves. D'où un nombre d'étapes de validation différent en fonction des questions que l'on se pose.

Problème des libellés d'affiliations : multiplicité nocive des graphies, assez particulière à la France. Il faut une règle: *une* affiliation correspond à *un* organisme. Pour le WoS, l'institution qui a le plus haut niveau hiérarchique est l'organisme principal. C'est, aux Etats-Unis, pratiquement toujours l'Université (où s'effectue majoritairement la Recherche). L'Université est donc considérée quasi systématiquement comme l'*organisme principal* et, dans la mesure où ce modèle anglo-saxon a presque valeur de monopole, c'est celui-ci qui a tendance à devoir être appliqué ailleurs. La diversité des graphies françaises apparaît lorsque l'on active l'onglet *Institutions* du WoS. Donc, on ne peut avoir, par le biais de cet onglet, une vue juste de la production d'une institution, mais cela peut donner une indication.

On peut visualiser les résultats avec un tri par langue de publication, sachant que l'anglais est bien évidemment majoritaire afin d'être universellement lu.

Sur un terme interrogé, on peut trier soit en fonction de la date, soit en fonction des citations.

L'activation du menu déroulant de Tri (*Sort by*) permet de visualiser les références dans l'ordre numérique décroissant du nombre de fois citées (*Times cited*).

Il est également possible de visualiser le nombre de fois où, chaque année suivant sa publication, un article a été cité.

L'I.F. porte sur les deux années précédant l'année de référence (i.e. celle où l'on effectue l'analyse). Il s'agit bien de la mesure de la notoriété d'un journal et non de l'impact des articles pris individuellement avec leurs citations.

Dans des journaux d'excellence comme *Nature* et *Science*, la barrière éditoriale est sévère. Moins de 1% des articles soumis sont acceptés. Leur I.F. est très élevé.

Dans certains domaines disciplinaires, les citations n'ont pas coutume d'être effectuées aussi rapidement que dans d'autres. Il faut donc nuancer l'interprétation des I.F.

L'interrogation portant sur le mot « Innovation » a fait apparaître, dans les résultats, une référence portant sur « l'innovation bactérienne ». La progression de cet article de biologie en volume de citations, tant en nombre qu'en rapidité sur le plan temporel, est significative de la différence des usages suivant les disciplines. Dans le domaine des Sciences de la Vie, le nombre de publications est nettement supérieur à celui des Sciences sociales, et de ce fait, l'impact l'est aussi. L'I.F. dans ce domaine sera plus significatif (ou porteur de sens) qu'il ne peut l'être en Sciences sociales.

Les citations vers cet article biomédical proviennent à 30% des USA, 32% de pays de l'UE *non anglophones*. Cette dernière observation est typique de la non exclusivité d'une exploitation états-unienne des articles.

Action du bouton *Create Citation report* : cliquable pour un auteur s'il génère moins de 10 000 publications à l'écran. Un graphique porte sur les 20 dernières années des citations reçues par les articles d'un auteur en particulier. Une attention spéciale doit être accordée à l'élimination des homonymes. Des patronymes associés à des prénoms (ou leur initiale) identiques s'attachent très fréquemment à des personnes différentes, parfois en grand nombre.

L'exemple montré ici porte sur un professeur en Economie. Il a assez peu publié, comme en fait foi le graphique affiché. Mais il est beaucoup cité. Un seul article suffirait à faire monter le visuel graphique des citations. Néanmoins, il apparaît que le nombre moyen de citations par article est relativement élevé. De plus, il a un H index de 16 à ramener à un total de 20 articles (sur les 20 publications, 16 ont été citées au moins 16 fois). Donc ici, le *citation report* lui est très favorable. On peut avantageusement opposer cet exemple à celui d'un auteur qui aurait aussi un H index de 16, mais ramené à 340 papiers. D'où l'impérieuse nécessité d'interpréter les chiffres au plus juste. Un classement de personnes par H index « bruts » ou nombre de citations « brut » est déraisonnable. Message majeur : « aucun indicateur seul ne peut être utilisé pour établir un *ranking* quelconque entre les chercheurs des institutions ou des organismes ; cela n'a aucun sens ». La

variation des paramètres mentionnés dans l'exemple qui précède doit impérativement être analysée avant d'évaluer les données relatives à un pays, une institution ou un chercheur.

Etait issue, de la recherche sur le terme « Innovation », la référence à un article en biologie, bénéficiant d'un nombre de citations extrêmement élevé. La revue dont il est extrait a un I.F. avoisinant 6. Le lien entre l'I.F. d'un journal et le nombre de citations recueilli par un article n'est pas simpliste. Il n'y a pas un lien au niveau individuel des chercheurs. Un nombre significatif d'articles issus de *Nature* et *Science* peuvent collecter un nombre en revanche modeste de citations en dépit du très fort I.F. de ces deux journaux[1]. Un journal comme *Oncogene*, doté d'un I.F. avoisinant 7, inclura en revanche de nombreux articles extrêmement prolifiques en citations. En France, sur les articles les plus cités au niveau mondial, moins de la moitié sont dans des grands journaux dits « généralistes » qui ont des I.F. très élevés (supérieurs à 20). L'I.F. n'est pas l'unique indicateur à mettre en avant lors d'une évaluation. Un article écrit dans une revue française à visée vétérinaire, décrivant pour la première fois une maladie transmissible entre les vaches, pourra être beaucoup cité ; mais ce phénomène sera statistiquement bien moins fréquent que vis-à-vis d'articles publiés dans *Nature* ou *Science*. Les comités de rédaction de ces journaux sélectionnant rigoureusement les articles, les chances dont sont dotés ceux-ci d'être abondamment cités sont plus importantes.

La visualisation des citations d'un article biomédical sur une échelle de temps montre que le plateau du nombre de citations est assez rapidement atteint. Ce phénomène est particulièrement mis en avant lorsqu'on utilise sur le WoS, par le biais de l'onglet *Additional resources*, un autre outil Thomson : *Essential Science Indicators*. On y verra par exemple qu'un article, pour une année donnée, fait partie du *Top 1%* du domaine au niveau mondial.

Un article biomédical portant sur le même sujet, paru la même année dans *Science* et dans une revue à I.F. beaucoup plus faible : *Journal of Biochemical Chemistry* (I.F. 5) affichera un nombre de citations qui, précisément, ne tient pas compte de l'I.F. : soit 477 citations pour l'article paru dans *Science* et 700 pour l'article apparenté paru dans *J Biol Chem*. Si l'on se penche un peu plus sur l'origine de la différence de ces scores, on observe qu'elle est marquée par un phénomène proprement scientifique. Le premier récepteur que cet auteur a décrit est spécifique à une bactérie (*Listeria*) alors que le second ne l'est pas et qu'un grand nombre de bactéries peut s'accrocher à ce récepteur. Le premier article a un nombre de citations (477) qui correspond à son exploitation par une communauté de chercheurs réduite à l'étude du contexte très particulier de ce récepteur (association limitée à *Listeria*). Le second article est lu et cité par une communauté plus large (700 citations) qui profite de la validation de son étude du récepteur vis-à-vis d'un plus grand nombre de bactéries. Cet exemple montre qu'il ne faut pas obligatoirement viser le journal doté d'un I.F. maximum dans la discipline, en tant que cible prioritaire, mais faire en sorte que l'article s'adapte à une communauté aussi large que possible, tout en conservant un niveau de spécialisation appréciable et pertinent, publié dans un journal que fréquentera cette communauté élargie.

Si l'on se concentre maintenant sur le cadre des équipes: il y a 257 équipes à évaluer, ce qui représente à peu près globalement selon les années, 26 000 à 30 000 publications à se partager entre environ 300 membres de commissions si on prend tous les membres sur toutes les disciplines. « Mission impossible ». Alors : on peut certes partir du principe que les articles ont déjà été évalués via la *peer-reviewing* du journal et s'en tenir là sur la validité des travaux ; mais pour diverses raisons tenant – par exemple – à l'identité des *reviewers* et aux tissus relationnels, on est obligé d'exiger des critères de validation complémentaires. Il est clair que le facteur temps est essentiel. Comment prendre le temps de lire la totalité des articles ? Même en Sciences sociales, des volumes de 600 pages ne peuvent être parcourus que brièvement.

Quels critères auxiliaires peut-on envisager ? L'I.F., on l'a vu, peut être associé, suivant son importance numérique, à la difficulté de franchir telle ou telle barrière éditoriale. Est-ce suffisant ? De même, le nombre de citations est-il réellement significatif, pris de façon « brute » ? Cela dépend du nombre d'années, du domaine... Et même, au sein d'un domaine identique, la comparaison du nombre de citations vaut-elle « au premier degré » ?

Seuls 15% des journaux analysés par le WoS détiennent un I.F. > 3.

67% des journaux ont un I.F. < 2.

44% l'ont < 1.

Il serait déplacé de n'examiner que les journaux dont l'I.F. est > 5 voire > 20.

Certes, *La Presse médicale* (I.F. < 1)^[2], qui est le premier journal dans lequel la France publie, constitue un cas exceptionnel.

Pour contrecarrer la critique qui lui était faite de ne calculer l'I.F. que sur deux ans, le WoS présente, depuis 2008, son calcul sur 5 ans.

Un exemple de calcul d'I.F. classique sur deux ans est développé à partir d'une revue de Sciences sociales : *Social Studies of Science*. On voit que le nombre annuel d'articles est peu élevé (une trentaine) et, compte tenu de la relative spécialisation du journal, le nombre de citations est proportionnellement aussi faible. Donc l'I.F. ne peut logiquement pas devenir significativement élevé. L'I.F. sur deux ans de l'exemple est à 1,7. Rapporté à cinq ans, il ne passe qu'à 2,3. Ce sont les papiers publiés en 2005 qui ont attiré, en 2010, le plus de citations. Par conséquent, en Sciences sociales, si l'on voulait calculer des I.F. « corrects », ils devraient être calculés sur dix ans.

Un deuxième exemple concerne la revue *Immunity* : le nombre d'articles publiés dans l'année est triple de celui de la précédente revue de Sciences sociales, ce qui génère un I.F. en conséquence. Autre corollaire, en contraste par rapport au premier exemple de *Soc. stud. sci.* : le nombre important de citations rapidement générées par les articles issus d'*Immunity* baisse dans le temps, de façon concomitante en regard de l'évolution de la science et de l'innovation. En Sciences sociales au contraire, il monte.

Un troisième exemple reflète les caractéristiques du deuxième, cette fois en médecine clinique (*Pediatric Infectious Disease Journal*) : publication abondante, à un rythme très rapide. Donc une exploitation proportionnellement plus rapide, tant en lecture qu'en citation. Les éléments de comparaison, en fondamental (*Immunity*) et en clinique (*Pediatr. infect. dis. j.*) induisent les mêmes observations. Le plateau de deux ans est pertinent et significatif. L'extension de ces modalités d'évaluation, très opportunes dans le domaine biomédical, ne sont objectivement pas adaptables telles quelles à la généralité des domaines.

Comparons maintenant l'évolution des I.F. de deux journaux biomédicaux : *The Lancet*, généraliste, qui passe en 5 ans (2000-2004) de 10,3 à 21,7 ; ce qui pourrait logiquement laisser supposer qu'en peu de temps, des innovations exceptionnelles auront pu marquer l'ascension de cette revue. Or, par contraste, *Immunity* voit son I.F. chuter sur la même période de 21 à 15,4.

En réalité, la rédaction de *The Lancet*, prenant conscience de l'importance stratégique croissante de l'I.F., a accru le nombre, dans ses sommaires, des articles de type *review* qui bénéficient usuellement d'une capitalisation importante de citations, parfois le décuple d'un *research article* proprement dit. *The Lancet* avait publié 37 *reviews* en 2000 contre 110 en 2004 et 780 *research articles* en 2000 contre 367 en 2004.

Immunity a rencontré un concurrent : *Nature Immunology*^[3]. D'où une baisse relative de l'I.F. du premier journal. Afin de redresser la barre, *Immunity* a accru le nombre de *reviews*. Son I.F. est alors passé de 15 à 24 entre 2004 et 2010 et son coefficient de *reviews*, de 10 à 18% de 2006 à 2010.

La nationalité des membres de l'*editorial board* d'un journal est importante : s'il est composé quasi exclusivement d'Américains, la difficulté pour un chercheur français de franchir la barrière sera supérieure à celle qu'il rencontrera si ce même *board* est composé pour moitié d'Américains et d'Européens.

Dans cet ordre d'idées, des analyses ont été effectuées sur le JCR. On a pris les 20 premiers journaux – en fonction de leur I.F. – de chaque *Subject category*, pour déterminer en pourcentage la nationalité des membres du *board*. Dans ces différents « top 20 », on a constaté la présence d'une majorité « américaine » de membres des *boards* à hauteur de 53%, contre 4% de Français. Par comparaison, les publications « américaines » représentent 28% des publications mondiales, les françaises 5,5.

Ainsi, lorsqu'un chercheur français déclare faire partie d'un *editorial board*, cet élément doit être sérieusement pris en compte dans l'évaluation. Suivant le contexte, cette position peut être très valorisante.

Dans le domaine biomédical pris en particulier, la proportion américaine représente 60,42% et la France n'occupe que 2,19% du terrain dans les *editorial boards*.

Une analyse parallèle est effectuée en se référant cette fois à la nationalité du journal.

Dans *Science* (USA), trois quarts des affiliations sont américaines. Un peu moins du quart, européennes, essentiellement représentées par le Royaume-Uni, l'Allemagne, la France et l'Italie. Les autres pays ne constituent qu'une quantité négligeable.

Dans *Nature* (UK), un peu moins de trois quarts des affiliations sont américaines ; environ un quart, européennes (avec le même groupe de 4 : Royaume-Uni, Allemagne, France, Italie) et les « autres pays du monde » commencent à être représentés.

Si l'on examine maintenant deux grands journaux généralistes biomédicaux : le *JAMA* américain et *The Lancet* anglais, on voit que les proportions du *JAMA* sont quasiment identiques à celles de *Science* alors que dans *The Lancet*, la part européenne constitue la moitié.

Il est clair qu'il est plus difficile pour un chercheur français de publier dans *Science* que pour un chercheur Américain.

Si l'on regarde maintenant sur un histogramme (portant les citations en ordonnée et les articles en abscisse) les citations reçues en 2004 et 2005 par les articles publiés en 2004 dans *Nature*, typique des grandes revues généralistes, on voit que certains articles atteignent des pics exceptionnels en nombre de citations ; en revanche beaucoup sont « moyens », tandis qu'un nombre significatif d'entre eux ne récoltent aucune citation. Si l'on trace une ligne transversale sur les colonnes de l'histogramme, matérialisant l'I.F. 26,681 (en 2006), on voit que certains articles « phares » sont certes en dessus, mais nombreux sont ceux qui restent en-deçà. Donc, la puissance numérique exemplaire de son I.F. ne signifie pas que tout article paru dans ce journal sera obligatoirement cité de façon notable, ou même cité tout court. Un article dont la teneur est de grande qualité mais n'intéresse personne, ne sera pas cité ou fort peu, même s'il est publié dans un tel journal.

Toujours dans *Nature* (année 2001), si l'on fait des calculs séparés sur les articles (*research articles*) et les *reviews*, à raison de 1332 articles et de 60 *reviews*, on voit que la moyenne de citation (calculée jusqu'en 2005) des *reviews* est de 212 et celle des articles n'est que de 74.

Dans la partie *reviews*, la moyenne de 212 a été « tirée » par une seule d'entre elles qui a été citée 4000 fois.

Dans l'autre partie, celle des articles, 70% se situent en-dessous de la moyenne de 74.

Soulignons que ces chiffres sont calculés sur une période de 5 ans et non de 2.

Dans *Science* et *Nature*, de très nombreux domaines sont présents. L'analyse de la représentativité respective de ceux-ci est essentielle pour interpréter l'évaluation de ces journaux. Cette représentativité, dans ces deux journaux, montre évidemment des distorsions.

Une analyse effectuée dans *Nature* sur la période 2001-2011 met en avant deux phénomènes. D'une part, le taux de citations par article le plus important se situe en *Clinical medicine* (254,50 pour 886 publications), mais cette discipline n'arrive qu'en 5ème position pour ce qui est du nombre de publications. Les deux domaines qui viennent en tête du nombre de publications sont *Molecular Biology & Genetics* et *Biology & Biochemistry*. Ce sont deux domaines transversaux dans les sciences biomédicales. Les communautés touchées par ces deux disciplines sont vastes et diverses : neurologie, immunologie, microbiologie etc., ce qui accroît en conséquence le nombre de citations. Il y en aura plus par le biais de ces thèmes transversaux qu'à travers des disciplines circonscrites et très spécialisées, comme les neurosciences prises isolément, voire la chimie. Les communautés lectrices et, pourtant, citantes, risquent alors d'être plus restreintes.

Sur cet échantillon de dix années de *Nature* (2001-2011), en Sciences sociales, se dégagent 59 articles, soit 6 par an. En fait, avoir publié dans ce domaine dans *Nature* n'implique pas, loin s'en faut, que l'on soit lu largement. Il apparaît qu'un publiant dans cette configuration, est généralement mal représenté et, en outre, mal cité, parce que ce domaine n'est pas typiquement représentatif du profil du journal.

Prenons maintenant l'exemple de *Science*, toujours sur la période 2001-2011. Physique et Chimie sont bien représentées dans ce journal comme dans *Nature*. Dans *Science*, la Médecine clinique est légèrement moins bien placée que dans *Nature* au niveau du nombre de publications (7ème rang contre 5ème) mais reste première au plan des moyennes de citations par article (259,18 pour 556 publications). Il y a certes vingt publications de plus en Sciences sociales (79), dans le journal *Science* que dans *Nature*, mais le niveau des citations reste toujours faible.

Toujours vis-à-vis des Sciences sociales, le comparatif en nombre de papiers et de citations est éloquent, en regard de la Médecine clinique.

Dans *Nature* (2001-2011), il y a quinze fois plus de papiers en Médecine clinique qu'en Sciences sociales, et cinq fois plus de citations dans le même rapport.

Dans *Science*, il y a sept fois plus de papiers en Médecine clinique qu'en Sciences sociales, et 7,6 fois plus de citations dans le même rapport. *Science* est donc un peu plus « large » en faveur des Sciences sociales, mais l'observation relative à la pertinence qu'il y aurait de publier en Sciences sociales dans *Science* est aussi valable que vis-à-vis de *Nature*.

Lors d'une évaluation des équipes en 2007 sur 18 773 publications, sur la période 2002-2006, on a tracé la corrélation entre l'I.F. des différents journaux et le nombre de citations. Si l'on reste en I.F. < 10, on aura une quantité très significative d'articles très bien cités. Lorsque l'on fait glisser le curseur vers les journaux dont l'I.F. est très élevé (avoisinant 50), on aura, certes, des articles très bien cités mais la proportion d'articles faiblement ou pas cités n'aura pas diminué de manière significative. La corrélation I.F. / Nombre de citations n'est pas bonne. : $r = ,45626$

Si l'on regarde cette corrélation, non pas dans l'ensemble des articles, mais dans les articles les plus cités, sachant que sur 18 773 articles, 723 font partie du Top 1% des articles les plus cités, elle est encore plus mauvaise que dans le cas précédent, soit $r = ,38397$. On voit bien qu'il y a beaucoup d'articles très cités dans les journaux de faible impact et toujours des articles peu cités dans les journaux à I.F. élevé.

Est évoqué un rapport émis en 2009 par le Centre for Science and Technology Studies de l'Université de Leiden (Pays-Bas), très importante structure d'études bibliométriques au plan mondial, dont l'une des figures marquantes est, depuis une trentaine d'années, Ton Van Raan, physicien de formation. Ce rapport fait état d'une évaluation menée sur des chercheurs britanniques. Des calculs ont été effectués pour connaître la part de citations qu'on obtient, du fait qu'on est publié dans un journal de grande notoriété. Il conclut que la notoriété du journal contribue pour moins de 10% au nombre de citations générées.

Le facteur H : il est assez souvent très élevé chez les chercheurs du domaine biomédical. En France, le plus élevé est détenu par Pierre Chambon (142), en biologie moléculaire. C'est assez logique puisque la génétique est globalement le domaine le plus cité. Le H est à manipuler avec précaution. Si l'on veut faire des comparaisons, qu'au moins elles se fassent au sein d'une même discipline.

Examinons le cas de deux chercheurs ayant l'un et l'autre un H index de 9. L'article le plus cité du premier recueille 20 citations, celui de l'autre 249. Ce dernier détient deux articles phares, le second totalisant 200 citations, et un troisième, plus loin derrière, avec environ 70 citations. On voit donc que le H seul ne signifie rien. Il est regrettable que certaines universités en viennent à utiliser le H dans les concours de recrutement.

Le WoS a découpé sa base en 22 « grands domaines ». Dans le module dénommé *Essential Science Indicators*, tous les deux mois est éditée une table permettant de déterminer combien de citations une publication de telle ou telle année (précision *chronologique* importante) doit recueillir pour être classée dans le *Top 1* ou le *Top 10 %* d'une discipline donnée (là aussi, précision *disciplinaire* importante). Le classement de Leiden tient compte de la part qu'a une institution, par le biais de ses articles, dans le *Top 10 %*.

On peut améliorer l'utilisation de l'I.F., on peut le normer sur le domaine : on fait la moyenne des I.F. d'un domaine.

Quand un chercheur appartient à la Génétique et qu'il publie dans le 25ème journal de sa discipline, dont le facteur est égal à 7, on divise ce chiffre par le facteur d'impact moyen de la discipline, i.e. 4,1. Si l'on exécute la même opération sur les Sciences sociales, dotées d'un I.F. moyen de 0,964, elle ne va pratiquement rien diviser. Cela permet de rétablir les équilibres entre les domaines.

On peut également mesurer le seuil d'I.F. exigible dans une discipline pour qu'un journal puisse figurer dans le Top 10%. A supposer qu'il y ait 181 journaux en Sciences de l'Agriculture, on voit que le 18ème journal (donc au 10ème percentile) a un I.F. de 2,51 qui va par conséquent correspondre à ce seuil.

Les calculs comparatifs doivent idéalement s'affranchir des biais générés par l'année et par le domaine de publication. Affranchissement vis-à-vis de l'année (pour comparer des chercheurs qui n'ont pas le même âge dans la recherche) mais aussi des différences de spécialités à l'intérieur d'une même discipline (ici le Biomédical). Ainsi la Dermatologie n'engendre-t-elle habituellement pas une multitude de publications, et partant, des citations en grand nombre ; ce n'est pas pour autant qu'il faut marginaliser cette spécialité. En Dermatologie le meilleur journal aura un I.F. proche de 6 tandis qu'en Cardiologie, il avoisinera 15. D'une façon générale, dans le domaine biomédical, on aura de très grands écarts : au lieu de se situer entre 0 et 5, ils s'étaleront entre 5 et 20.

Evaluation des Unité ou Centres ; Indicateurs bibliométriques équipes

Certifiée ISO 9001, la méthodologie bibliométrique en usage à l'INSERM se décompose en une dizaine d'étapes avant la production d'un document final. Elle est agréée aussi bien par le Conseil scientifique que par les Commissions scientifiques spécialisées de l'institution, qui en sont destinataires.

La quotité de travail ou d'emploi des auteurs : un chercheur EPST a un ETP égal à 1, un chercheur universitaire a un ETP égal à 0,5 puisqu'il est censé faire la moitié du temps de la recherche et l'autre moitié de l'enseignement. Un hospitalo-universitaire aura 0,33 (hôpital + recherche + enseignement). On peut avoir aussi affaire à des chercheurs n'émargeant pas directement à telle ou telle de ces catégories prédéfinies : un médecin ou un universitaire peuvent par exemple être détachés pendant un an à 100% sur de la recherche.

De nombreuses corrections doivent être apportées aux données, notamment sur les noms d'auteurs, parfois libellés de façon différente, avec des variantes, par les intéressés eux-mêmes (problèmes du nom marital, des noms doubles, des prénoms multiples...). Les données fournies par les chercheurs sont gérées dans *End Note* et corrigées le cas échéant afin que la base soit propre, en vue de l'exploiter sous forme de requêtes dans d'autres outils.

Chaque journal sera affilié à un domaine. La question se pose pour les grands journaux généralistes. Si l'on est spécialiste en neurosciences et si l'on publie dans *Nature*, on restera dans le multidisciplinaire, ce qui n'est pas significatif. Pour ce périodique, comme pour *Science* et *PNAS*, on retypera l'article en consultant au moins son titre et son abstract.

Lorsque le WoS donne les indicateurs de seuil, ce retypage est effectué automatiquement en fonction des références présentes dans le journal. Si une dominante de domaine y apparaît, il le classera dans celui-ci. S'il en ressort plusieurs, il le basculera en multidisciplinaire.

Utilisation de *ReseauLu* : la société Aguidel[4] a mis au point ce logiciel de traitement de données hétérogènes permettant de transformer les données bibliographiques en tables de données relationnelles. Il est utilisé en combinaison avec des bases Access en vue de produire des représentations graphiques significatives.

La bibliométrie est donc utilisée afin d'apporter une vision « objective » sur la visibilité des recherches. Celle-ci exclut les interférences « humaines » associées par exemple à des réseaux d'influence susceptibles de biaiser le signalement de telle ou telle publication. De ce fait, si un chercheur est classé 25ème en dépit d'une excellente bibliométrie, la Commission sera tenue de justifier ce classement paradoxal. Deux *J Exp Med* et deux *J Biol Chem* pouvaient usuellement, à une époque antérieure, permettre d'attribuer d'office l'appréciation « A+ » à un chercheur. Aujourd'hui, ce ne sera plus suffisant, on sera obligé d'avoir d'autres arguments.

Il est clair que l'évaluation par les pairs n'est pas évincée par la bibliométrie. En outre, celle-ci s'appuie sur un grand nombre d'indicateurs.

Le premier d'entre eux est certes purement quantitatif : le nombre d'articles reflète la production en tant que telle.

Ensuite, on a vu que la visibilité internationale pouvait se traduire en plusieurs « tranches », identifiées comme les *Top 1*, *10*, *20* et *50* % (médiane mondiale). Elles tiennent compte aussi bien de l'année de publication que du domaine (cf. ces deux nuances déjà exprimées plus haut). Le *Top 10%* fait l'objet d'une attention particulière, comparé au nombre total de publications, afin de définir un taux de réussite, autrement dit un pourcentage d'excellence.

Il faut tenir compte de la taille des équipes. Certaines sont grosses, avec 25 chercheurs revendiquant par exemple 142 publications, d'autres beaucoup plus petites avec 2 chercheurs et 5 publications. Leur comparaison ne peut se faire en s'appuyant seulement sur des nombres. Donc, à l'instar de Leiden et d'autres cellules bibliométriques internationales, on évite de manipuler les « tranches de Top » sans discernement. Ainsi le *Top 1%* convient-il assez bien au niveau d'un pays, mais moins franchement à celui d'une équipe, car à cette échelle, des résultats extrêmement faibles risquent d'être particulièrement fréquents et donc pas assez discriminants.

La position des auteurs : on sait que celle-ci, dans l'énoncé des co-auteurs en tête d'un article, est plutôt déterminante dans le secteur biomédical. Le jeune chercheur sera souvent mentionné en tête. Le dernier mentionné est également doté d'une importance spécifique.

Dans une expérience, le *senior author*, qui a dirigé en général les travaux, figure plutôt en dernière position dans l'énoncé des co-auteurs. Lorsqu'il appartient à une équipe dans laquelle un chercheur notable estime avoir vocation à signer tous les papiers en dernier, il alternera avec son binôme la dernière et l'avant-dernière positions. Il convient donc de connaître les coutumes des chercheurs qu'on évalue pour pouvoir utiliser ce genre d'indicateur.

Ces positions se voient attribuer un indice spécifique : un indice de 1 est décerné à la première et dernière position ; 0,5 à la 2ème ou l'avant-dernière ; 0,25 à toutes les autres positions. Des tableaux permettent de voir très vite la différence de score avant et après « correction » ou réajustement en fonction des positions.

Ensuite, on prend en compte l'I.F. brut du journal pour chaque article, qui reflète les difficultés éprouvées par l'auteur pour publier, et l'I.F. normé par domaine. Le H index est aussi évoqué en dépit des réticences signalées, avec une référence à 10 ans et sur la carrière.

Un graphique permet de visualiser le nombre de publications dans le *Top 1%* par équipes (253 équipes pour 2007-2008). La médiane se trouve à 1 tandis que 70% des données sont inférieures à 3. De ce fait, comme dit plus haut, un indicateur d'excellence trop élevé n'est pas franchement significatif pour apprécier valablement des équipes. A l'inverse, des indicateurs trop larges, agrégeant une trop grande quantité de chercheurs voire leur globalité, ne sont pas non plus significatifs.

Un schéma identique portant cette fois sur le *Top 10%* permet d'avoir des chiffres plus « raisonnables » avec une médiane à 11 et une moyenne à 14. Aucune équipe n'avait pas d'article dans le *Top 10%*.

Des tableaux synthétiques permettent de visualiser de façon synoptique les éléments utiles aux décisions des Commissions en matière d'évolution de carrière des chercheurs à partir de cinq ans de travaux.

On voit apparaître, entre autres, sur le tableau : le nombre de publications ; ce nombre corrigé en fonction de la position d'auteur, du nombre de citations, du *Top 10%*, du nombre de publications situées au-dessus de la médiane mondiale ; l'I.F. moyen ; le nombre de journaux dont l'impact relève du *Top 10%*.

L'ensemble de ces données est compilé, dédoublonné compte tenu des chercheurs qui ont collaboré entre eux, et cela donne une idée des travaux de l'équipe.

Un tableau est agencé avec des onglets permettant d'accéder aux données de chaque chercheur afin de visualiser ses différentes caractéristiques en regard de ses publications. Les données sont manipulables et combinables de sorte que les membres de la Commission puissent se faire une idée équilibrée, à partir d'indicateurs multiples et nuancés, de la situation d'un chercheur. Les caractéristiques d'un lot d'articles remarquables étant signalées, les membres de la Commission peuvent se rendre sur les bases de données adéquates, accéder au *full text* pour se faire une idée plus précise, etc.

Un outil graphique permet de visualiser, pour chaque équipe, les publications appartenant au *Top 10%* (en ordonnée) en regard de l'I.F. moyen (en abscisse). Ces données peuvent apparaître de manière concomitante sur le même tableau pour plusieurs équipes. Ainsi, telle équipe aura publié plusieurs fois dans des journaux à très fort I.F. mais il se trouve que leurs articles sont très peu visibles. A l'inverse, telle autre équipe aura de nombreux articles dans la zone du *Top 10%*, mais les journaux correspondants auront un I.F. modéré.

Ces équipes appartiennent généralement à un Centre. On établit donc les indicateurs déjà détaillés précédemment pour l'ensemble du Centre. Ensuite, on tente de dégager les spécialités dominantes pour chaque Centre. Il s'en suivra un indice de citation moyen normé par rapport à une référence mondiale dans une spécialité donnée. Tel Centre spécialisé en immunologie verra ainsi ses papiers dotés d'un indice spécifiant que leur taux de citation est trois fois supérieur à la moyenne mondiale correspondante dans cette spécialité, sur une période identique.

Au niveau de la production du Centre, on voit que, par exemple, une unité récoltera 71 publications en *Top 10%* parmi lesquelles (précision importante) 30 seront attribuables à un chercheur en première ou en dernière position. Ce peut être n'importe quel chercheur : un post-doc, un doctorant (en première position), associés à une personne appartenant au Centre, figurant par exemple en dernière position.

Ces détails et nuances font qu'un évaluateur ne peut pas se permettre de lire les tableaux en diagonale. On est vraiment obligé d'exploiter au maximum les précisions incluses dans le tableau pour en tirer le maximum d'informations. Il convient donc de former les membres des Commissions à tous ces indicateurs afin qu'ils les exploitent au mieux.

Au sein d'une équipe, on analysera la productivité par chercheur. Telle équipe aura par exemple un taux de cinq articles par an et par chercheur, telle autre 0,17.

Au niveau d'un Centre, on peut aussi analyser les indices de collaboration. On précisera le nombre d'équipes avec lesquelles chaque équipe a collaboré au sein du Centre. Il importe de savoir s'il y a synergie entre équipes différentes ou simple juxtaposition de deux chercheurs.

Si l'on fait la somme des publications, on peut faire en sorte que le résultat final corresponde à un nombre sans doublons. Mais si tous les chercheurs ont contribué à toutes les publications, on aura affaire à un indice « dilué ».

Une cartographie appropriée permettra de définir les collaborations au sein d'un Centre, sur le plan disciplinaire. Les réseaux peuvent être générés par des mots-clés issus des notices, les pays figurant dans les zones d'adresse, les institutions. Des couleurs permettront de repérer les chercheurs qui appartiennent à la même équipe.

On peut repérer aussi sur l'image les chercheurs qui travaillent de façon relativement isolée par rapport au réseau. A l'opposé de ces équipes « éclatées » apparaissent clairement des clusters d'équipes travaillant en synergie. Ces images sont importantes pour le Comité scientifique afin de se faire une idée de la vie des Centres.

D'autres représentations graphiques permettent, en 3D, d'avoir en ordonnée l'I.F., en abscisse l'excellence, i.e. la part en *Top 10%*. En faisant varier les paramètres sur deux axes médians, l'un indiquant le score zéro de l'I.F. moyen pour la France, l'autre le score zéro de la France dans un domaine donné, on voit où se situent des équipes ou des entités.

Ce schéma peut être appliqué à un Centre qui relèvera (du fait qu'il est monothématique) d'une Commission spécifique. Les lots de publications rattachés à des équipes sont représentés par des cercles de diamètre plus ou moins grands en fonction de leur importance, qui sont assortis d'une note de type B, A ou A+ correspondant à l'évaluation AERES, avant que la Commission ne se prononce elle-même. Des cercles figureront de part et d'autre des deux axes médians mentionnés précédemment, donc au-dessous ou au-dessus de zéro, en termes d'I.F. moyen ou de positionnement par rapport au *Top 10%*.

Le modèle s'applique également, naturellement, aux Centres pluri-thématiques. Là, chaque équipe est normée sur sa propre Commission.

L'hétérogénéité des résultats, i.e. une correspondance toute relative entre les notations AERES et la situation des cercles par rapport aux deux axes, suscite une réflexion sur la complémentarité des indicateurs bibliométriques. De fait, les évaluateurs ont du mal à juger la qualité des publications.

Pendant deux ans, à l'époque où Christian Bréchet était directeur de l'INSERM, a été mise en place une méthodologie qui, en plus des évaluations effectuées à l'initiative des membres des Commissions, prévoyait l'envoi de tous les dossiers à deux experts anonymes étrangers extérieurs qui notaient également les équipes, de leur côté. La confrontation des jugements permettait d'affiner ceux-ci.

Seules 38 % des équipes ont reçu la même note si l'on compare les jugements émis par les deux parties (corrélation $r=0,34$) sur 272 équipes analysées en 2007. Ces scores se décomposent ainsi, pour les notes attribuées de façon identique par les membres des Comités et par les experts étrangers : note B 17% de convergence, A 50%, A+ 35%.

Evaluation des chercheurs ; Indicateurs bibliométriques

L'ancienneté de carrière des chercheurs est, on l'a vu, importante, notamment pour le calcul du facteur H en regard de l'ensemble de la période d'activité analysée. La production d'un chercheur ayant débuté en 1982 n'est évidemment pas identique à celle d'un autre qui aura débuté en 1991. On prend en considération le facteur H depuis 10 ans. Ce genre de donnée est sujet à modulation. Un graphe de production individuelle montrera que celle-ci n'est pas régulière tous les ans. Il y aura des pics et des « trous ».

Si l'on observe un graphique de production individuelle (aujourd'hui accessible à partir de 1975) on ne peut définir une généralité portant sur un pic en milieu de carrière. Les cas de figure avec une remontée en fin de carrière ne sont pas rares, il faut examiner le parcours de chaque chercheur dans le détail. Sur un graphique, on aura en ordonnée une échelle de temps (années) avec en regard les titres des périodiques dans lesquels des articles seront parus, assortis de curseurs horizontaux se déplaçant en abscisse sur une échelle de nombres de citations. Un histogramme mentionne également la synthèse avec le nombre de citations en ordonnée et l'échelle des années de publication en abscisse. Les colonnes

permettent de visualiser la variation des taux de citations ainsi que d'auto-citations. Les citations peuvent en outre, à partir d'un « camembert », être détaillées en pays d'origine. Une variété internationale d'origines peut éventuellement amoindrir le risque de citations dues à un cercle rapproché de relations personnelles.

Lorsqu'un pic paraît exceptionnel en regard du schéma général, il peut être intéressant de voir son contexte. Sans doute s'agit-il d'une découverte marquante qui générera de nombreuses citations pendant un temps, avant que le cours de la science n'érode son étude et donc la citation de l'article correspondant.

De même, une erreur scientifique notable dans un article de recherche peut générer des citations sans que l'on puisse en tirer des conclusions hâtives.

Le problème des auto-citations est important. Théoriquement, on peut imaginer qu'un chercheur réutilise ses données antérieures lorsqu'il progresse dans son cheminement vers une découverte. ; réutilisation qui se traduirait logiquement par des auto-citations. Or, ce qui est étonnant, c'est que 30% des publications de l'INSERM ne sont pas citées, ce qui signifie que les chercheurs les ont écrites et ne les ont jamais réutilisées. C'est dans ce sens qu'il faut se poser la question, en particulier à propos des auteurs d'une publication qui sont très cités : par exemple, si l'on est cité 215 fois, il faudrait avoir écrit, après cette publication, au moins 215 papiers pour pouvoir se citer soi-même. Cette configuration est donc singulièrement complexe et, de ce fait, les auto-citations pour les papiers très cités ne compteront que de manière très faible. L'histogramme susmentionné en rend compte, notamment : la bande colorée illustrant le volume des auto-citations, à la base de chaque colonne, est soit d'une élévation minime, soit carrément absente.

L'exemple pris sur un chercheur en immunologie de l'INSERM montre des pics de citations à un niveau identique, tant dans une revue généraliste comme *Nature* que dans une autre, très spécialisée, comme le *Journal of Immunology*. Le plafond de citations qu'elles atteignent ne peut qu'être difficilement franchi, car le sujet abordé majoritairement dans les articles de ce chercheur est extrêmement spécialisé et ne peut donc concerner qu'une population relativement homogène et globalement peu mouvante.

Le fait que, pour une année donnée, un chercheur fasse partie du *Top 1%* des scientifiques mondiaux, est un critère utilisé dans le calcul de la Prime d'Excellence scientifique. A l'INSERM, certains chercheurs se la voient attribuer d'office compte tenu d'un dossier dont la qualité ne soulève aucune contestation notoire. Le positionnement d'un auteur en numéro 1, 2 ou ultime, dans des articles le propulsant en *Top 1%*, renforce les arguments en faveur de cette attribution d'office. D'autres chercheurs font en revanche l'objet d'une analyse plus approfondie. Par exemple, l'analyse de leur évolution sur les 5 dernières années entre en ligne de compte.

En 2011 il y avait 200 candidats à cette prime. En 2010 : 350. Cela entraîne des travaux impliquant une réflexion importante de la part des Commissions.

Le WoS va, tout comme les publications en 22 grandes disciplines, classer les scientifiques eux-mêmes, et ce sur 10 ans, dans *Essential Science Indicators*. Ils sont toujours classés au nombre de citations, mais par domaine. Il faut exister dans *une* discipline pour exister dans le *Top 1%*. Si c'est la somme de toutes les disciplines qui fait le nombre de citations, cela ne vaut pas. Il faut que dans *une* discipline, le *Top 1%* soit atteint. On peut situer un chercheur dans le lot du *Top 1%*, au sein d'une liste de chercheurs de même rang au niveau international. Figurer dans le peloton de tête ou de queue de ce premier percentile n'a, à la limite, pas d'importance. Simplement, il faut en être.

Le clic sur un icône en marge de la mention d'un chercheur fait apparaître les titres figurant en *Top 1%* pour ce chercheur.

La France, dans toutes les analyses bibliométriques, figure en général approximativement au cinquième rang mondial en nombre de publications et en nombre de citations, mais dès que l'on commence à définir des indicateurs qui rapportent le nombre de citations par publication, on descend à la dix-huitième place. Ce dernier élément étant notamment dû au fait que l'on publie un très grand nombre d'articles dans des revues qui ne sont que peu, ou pas citées du tout, ou citées par exemple deux fois dans toute la carrière du chercheur.

Donc, beaucoup d'énergie dans la rédaction de ces papiers, mais pas rentabilisée rationnellement au service d'une consolidation la recherche. Ils ne sont certes pas inutiles : lorsqu'on publie dans un journal de gastro-entérologie en langue française, c'est en partie pédagogique et lisible par un public qui ne fait pas de recherche ; mais dans ce cas, cela n'a rien à faire dans le WoS. Néanmoins, paradoxalement, des chercheurs continuent de publier dans des journaux à très faible I.F., inférieur à zéro, et en tirent malgré tout une certaine fierté du fait qu'ils figurent dans le WoS. Cela n'a guère de sens. Que leur papier soit autrement visible dans *PubMed*, cela se conçoit et peut être louable, mais il serait préférable que le journal ne soit pas indexé dans le WoS

De nombreux chercheurs persistent à penser qu'il faut beaucoup publier. En réalité, il est préférable de publier moins et

d'avoir plus d'impact sur la communauté scientifique.

Cela engendre une réflexion complexe vis-à-vis de l'évaluation. Doit-on prendre en compte toutes les publications, quelles qu'elles soient, ou ne doit-on agréer, par exemple, que les papiers en langue anglaise ? Celle-ci, bien évidemment, accroît le facteur de visibilité internationale. C'est important si l'on souhaite attirer des post-docs ou des chercheurs en général, et obtenir des financements, notamment d'industriels étrangers etc. Il apparaît que, de plus en plus, les laboratoires pharmaceutiques étrangers désertent la France. De même, les laboratoires industriels français se délocalisent assez souvent en Angleterre, aux Etats-Unis...

Ces entreprises sont abonnées au WoS et le surveillent.

Un usage sain de la bibliométrie devrait permettre d'éviter l'emballement des publications, au profit par exemple de doctorants qui doivent impérativement publier plusieurs des papiers, mais dont la qualité et l'originalité ne sont pas forcément convaincantes, du moins pour l'extérieur.

La recherche est, d'un autre côté, une mission essentielle à laquelle les membres d'une institution telle que l'INSERM ont vocation. Or, là comme ailleurs, il peut exister des chercheurs non-publiants ; il paraît naturellement souhaitable d'en diminuer sensiblement le nombre. Il peut arriver que des non-publiants soient employés à des tâches à caractère plutôt « technique ». Non point qu'il faille que tout chercheur, indistinctement, publie systématiquement un ou deux articles par an. Mais il est indispensable que s'établissent des synergies dans les équipes et cela peut en partie résulter d'une pression supplémentaire exercée sur les chercheurs. En Angleterre elle a porté ses fruits. Ce pays nous précède, de même que l'Allemagne. Depuis environ cinq ans, l'Italie nous a aussi dépassés. L'importance des publications est aussi liée – c'est loin d'être négligeable – à l'industrie, aux brevets, aux biotechnologies. A cet égard la bibliométrie est, d'une manière ou d'une autre, le reflet de la façon dont l'extérieur nous voit. Elle n'est pas uniquement un outil d'incitation ferme ou « d'émulation administrative » vis-à-vis du chercheur.

Sciences de la vie et de la Santé – 2006-2007 ; Indicateurs bibliométriques

La mutualisation institutionnelle peut être une solution en vue d'optimiser les synergies et de mieux structurer la recherche française. C'est l'un des objectifs affichés par les ITMO (Instituts thématiques multiorganismes) qui doivent, en sus des avantages inhérents à leur structure même, s'appuyer sur une logistique financière *ad hoc*.

L'INSERM a contribué à définir la caractérisation des ITMO en travaillant sur les données bibliométriques applicables à chacun des domaines rattachés à ces Instituts.

L'utilisation de mots-clés appropriés, en lien avec les thématiques des journaux de spécialité, permet de cerner les domaines d'une entité donnée en récupérant ses publications. On utilise en parallèle les sous-catégorisations de l'ISI.

Ainsi voit-on, dans ce contexte, qu'environ 40 à 44% des publications contenues dans le WoS sont, en France, du domaine Sciences du Vivant et de la Santé (repérage à hauteur de 43,9% par le biais des mots-clés associés aux journaux de spécialité ; à hauteur de 39,8% par le biais des sous-catégories ISI). Cette opération a été appliquée sur les deux années 2006-2007. En regard des ITMO, les corpus considérés ont dû faire l'objet de rectifications manuelles nombreuses et « chronophages » dans les libellés d'affiliations (universités, CHU, INSERM, CNRS, multi-partenariats...). Il a souvent fallu rajouter une ligne pour les Universités, parfois pour le CNRS (cas des UMR).

Dans ce genre d'analyse, l'Université apparaît comme le premier producteur de la France puisque 80% des publications des Sciences du Vivant et de la Santé sont co-signées par au moins une affiliation universitaire. CNRS et INSERM sont à 25%.

Derrière les indicateurs se profilent des *benchmarks*. Ainsi peut-on en établir pour les universités dans les différentes disciplines (Cancer ; Circulation-Métabolisme-Nutrition ; Neurosciences-Sciencescognitives-Neurologie-Psychiatrie ; Technologies de la Santé), d'abord en fonction du nombre de publications (i.e. en termes de production), ensuite en termes d'excellence (i.e. *Top 10%* et I.F. > 9). On établira aussi des comparaisons avec d'autres universités internationales et l'on verra que telle ou telle institution française peut fort bien se détacher en excellence mais en même temps produire beaucoup d'articles peu visibles.

D'autres analyses permettent de différencier les institutions par le nombre d'articles publiés dans *Science* et *Nature*.

Une analyse effectuée sur des institutions internationales de haut rang dans le domaine de la Médecine clinique permet

d'afficher des différences sensibles sur l'échelle de la « réputation » (I.F. > 20) mais pratiquement pas sur celle de la visibilité (*Top 1%*) : étalement d'un côté, concentration de l'autre.

Outre la co-occurrence des auteurs sur une publication, on peut aussi, en regard des articles publiés par les auteurs d'une institution, voir dans quels journaux sont parus les articles qu'ils citent dans la bibliographie terminale de leurs articles. On peut dresser une typologie de ces titres cités. Ainsi *Blood*, par exemple, aura-t-il un caractère mixte entre le fondamental et le clinique. De même, on différenciera les journaux SHS ou à dominante technologique. Cette typologie permet de mieux caractériser les spécificités d'un domaine et leur interface (par exemple le Cancer, avec un pourcentage de répartition entre le fondamental et le clinique, en localisant sur un schéma le croisement de ces deux caractéristiques). Ceci peut être aussi, pour un domaine, rapporté à un pays.

Dans un tel schéma, on verra que certains chercheurs, par exemple en recherche clinique, seront relativement marginalisés, inversement en fondamentale – et réciproquement. L'intérêt de leur croisement, s'il apparaît dans une telle configuration, consiste à repérer les chercheurs qui suscitent cette interface (une compétence transversale en recherche clinique et fondamentale) et sont donc les mieux à même de créer des synergies propices à l'innovation. On a pu utiliser de telles méthodes pour optimiser la structuration d'équipes dans les cancéropôles.

Conclusion

Les outils bibliométriques peuvent être utilisés dans différentes situations. Il faut bien analyser à chaque fois ce qu'il faut choisir comme indicateurs et ne pas forcer le système à dire des choses qu'il ne peut pas dire. Ils sont complémentaires et ne peuvent être considérés comme des indices absolus. L'analyse est rendue pertinente par une juste interprétation de leur convergence. Par ailleurs, il importe de souligner qu'une compétence, une expertise dans le domaine analysé, sont plus qu'opportunes. La bibliométrie ne peut remplacer les experts, mais la combinaison des deux est utile. Un bibliomètre familier des notions scientifiques du domaine étudié, titulaire de diplômes dans la discipline, ayant exercé en laboratoire, et publié de façon remarquable, sera bien mieux à même d'interpréter les résultats des analyses bibliométriques.

La cellule bibliométrique de l'INSERM^[5], créée en 2002, est aujourd'hui constituée de 10 personnes.

[1] I.F. respectif de ces deux journaux dans l'édition 2010 du *Journal Citation Report* : 36 et 31. Ce cours varie, certes, mais relativement peu.

[2] I.F. 0,48 en 2010/p>

[3] démarrage de la publication en 2000

[4] <http://www.aguidel.com/page23.html> ^[11]

[5] On consultera avec profit son site sous : <https://www.eva2.inserm.fr/EVA/jsp/Bibliometrie/index.jsp> ^[12]

→ Mots-clefs

URL source: <http://urfist.chartes.psl.eu/atelier-evaluation-caracterisation-bibliometrie-seance-du-16-fevrier-2012>

Liens

[1] <http://urfist.chartes.psl.eu/tags/affiliations>

[2] <http://urfist.chartes.psl.eu/tags/atelier-bibliometrie>

[3] <http://urfist.chartes.psl.eu/tags/co-signatures>

[4] <http://urfist.chartes.psl.eu/tags/manuel-durand-barthez>

[5] <http://urfist.chartes.psl.eu/tags/mentions-daffiliation>

[6] <http://urfist.chartes.psl.eu/tags/publications-scientifiques-francaises>

[7] <http://urfist.chartes.psl.eu/tags/scientometrie>

[8] <http://urfist.chartes.psl.eu/tags/signatures>

[9] <http://urfist.chartes.psl.eu/tags/inserm>

[10] <http://urfist.chartes.psl.eu/tags/mod%C3%A8les>

[11] <http://www.aguidel.com/page23.html>

[12] <https://www.eva2.inserm.fr/EVA/jsp/Bibliometrie/index.jsp>

[13] <http://urfist.chartes.psl.eu/types-de-public/doctortant>

- [14] <http://urfist.chartes.psl.eu/types-de-public/enseignant-du-sup%C3%A9rieur-chercheur>
- [15] <http://urfist.chartes.psl.eu/types-de-public/professionnels-de-l%E2%80%99information>