

L'ANALYSE BIBLIOMETRIQUE COMME OUTIL D'AIDE A LA DECISION STRATEGIQUE

PRESENTATION D'UN CAS : LES ETATS DES LIEUX DES ITMO

Nicole Haeffner-Cavaillon, Marie-Anne Jan

nicole.haeffner@inserm.fr, marie-anne.jan@inserm.fr

INSERM, Département de l'Evaluation et du Suivi des Programmes
Cellule Mesures, Indicateurs, Bibliométrie
101 rue de Tolbiac - 75013 Paris
Tél : 01 44 23 61 02 – 01 44 23 60 70

Mots clefs :

Bibliométrie, Analyse bibliométrique, Evaluation de la recherche, Données statistiques, Cartographies, Réseau de collaboration

Keywords:

Bibliometrics, bibliometric analysis, research evaluation, statistics, mapping, collaborative network

Palabras clave :

Bibliometría, análisis bibliométrico, evaluación de la investigación, estadísticas, mapas, redes de colaboración

Résumé

Créée en avril 2009, l'Alliance nationale pour les sciences de la vie et de la santé AVIESAN rassemble les grands acteurs en sciences de la vie et de la santé en France. La mise en œuvre des objectifs d'Aviesan se fait au sein de 10 Instituts Thématiques Multi-Organismes (ITMO) dont le rôle principal est d'animer la réflexion stratégique au sein de leur propre domaine scientifique. La cellule Mesure, Indicateurs et Bibliométrie de l'Inserm a eu pour mission d'établir des bilan-diagnostic de la production scientifique par Institut pour le pilotage de cette nouvelle organisation. La présentation montre, pour chaque Institut, les indicateurs bibliométriques permettant de positionner la recherche en France. Différents logiciels ont permis de mettre en place cette méthode : recherche de mots clés avec un logiciel de Textmining, utilisation du logiciel de gestion bibliographique Endnote, Access et Excel pour établir des requêtes et les données statistiques à partir des imports de notices bibliographiques du Web of Science (Thomson-Reuters), analyses des réseaux de collaboration à partir du logiciel Reseau-Lu, représentations 3D à l'aide d'une macro Excel, et répartition géographique à l'aide du logiciel Arctique.

1 Introduction

L'analyse bibliométrique permet de mesurer la production scientifique d'un pays, d'un institut, d'un chercheur, elle fournit un outil de comparaison globale ou dans différentes disciplines, elle permet d'identifier les réseaux de collaboration, d'estimer l'impact des travaux, elle apporte une appréciation complémentaire à l'expertise scientifique. La bibliométrie est considérée actuellement comme une mesure "objective" de l'activité de la recherche. Devenue accessible à tous, elle est de plus en plus utilisée dans le monde académique et politique comme un outil d'aide à la décision.

Les missions du groupe de bibliométrie de l'Inserm sont de mettre à la disposition de la direction générale, des ITMO, des commissions scientifiques et des chercheurs, des indicateurs permettant d'apprécier la position de notre pays, de notre organisme, des unités ou équipes de recherche, dans le contexte européen et mondial, des études thématiques ciblées sur les domaines de recherche de l'Inserm et des points de repères individuels et collectifs de productivité dans le domaine biomédical :

- Etablir des points de repères
 - Ces techniques se basent sur des acquis méthodologiques forgés au cours des 30 dernières années par des organismes de renom internationaux et nationaux comme: "The Institute for Scientific Information" (ISI), les "Observatoires des Sciences et de la Technologie" (OST) français, canadien et britannique, "The Center for Sciences and Technologie Studies of Leiden University (CWTS), etc. Ces points de repères ne peuvent se comparer qu'au sein d'une même discipline et sur des périodes similaires:
Repères pour 22 grands domaines par année, du nombre de publications (articles, lettres, revues), du total de citations de ces articles, de l'appartenance au Top 10% et 1% mondial des publications les plus citées, du facteur d'impact des journaux, du H-index etc.
 - Repères de productivité (nombre de publications par an par chercheur), d'index de citation moyen, de facteur d'impact moyen, etc.
- Produire des outils d'aide à l'évaluation scientifique

Etudes bibliométriques concernant les travaux des chercheurs et les unités de recherche de l'Inserm. Ces études sont mises à disposition des instances scientifiques de l'Inserm et serviront de base à la réflexion des commissions scientifiques en charge de l'évaluation et à la direction.
- Constituer des indicateurs thématiques : la bibliométrie prospective pour les instituts et les programmes nationaux de recherche

Etudes concernant des grandes spécialités médicales de l'Inserm telles que la nutrition, le diabète, les maladies cardiovasculaires, la rhumatologie et les pathologies osseuses, la psychiatrie et la santé mentale, etc.

Ces travaux permettent de mettre en exergue le rayonnement, l'ouverture et l'intégration de la recherche française dans la communauté scientifique européenne et internationale.

- Identifier des faits marquants de la France

Chaque année, la cellule identifie les faits marquants sur la base des citations, des grandes découvertes et des "domaines de pointe" de la recherche biomédicale en France.

Elle produit également des cartographies de pôles d'excellence.

Suite à la création des ITMO en 2009, la cellule de Bibliométrie a eu pour tâche d'établir des Etats des lieux de la production scientifique de chaque Institut qui vont servir pour le pilotage de la nouvelle organisation AVIESAN.

C'est la méthode d'élaboration des indicateurs bibliométriques qui est décrite ci-dessous.

2 Méthodologie de l'étude

2.1 Indicateurs bibliométriques

La recherche des publications (articles, lettres, revues) a été effectuée à partir de la base de l'ISI. Cette base est constituée de l'ensemble des articles provenant de 5800 journaux scientifiques (Science citation index ; SCI) indexés depuis 1945, de 1700 journaux de sciences sociales (Social sciences CI) indexés depuis 1956 et de plus de 8000 revues (Arts and Humanities CI) indexées depuis 1975. Cette base constitue aujourd'hui la base internationale la plus complète au niveau des notices bibliographiques et sert de référence mondiale à la majorité des bibliomètres de renom (CTSW de Leiden, Université de Shanghai, etc....).

L'évaluation des performances dans le domaine de la recherche biomédicale est souvent basée sur le nombre de publications et sur l'impact des journaux (Impact Facteur : IF) qui les éditent. Toutefois, l'accès récent aux bases de données qui référencent les citations des articles permet d'apprécier directement la visibilité internationale des travaux et de s'affranchir de l'utilisation de la seule notoriété des revues (IF). Les découvertes importantes qui jalonnent les progrès de la science font l'objet de publications dans de grandes revues dites généralistes mais aussi dans les revues de spécialité. Leur grand nombre de citations atteste de l'intérêt de la communauté scientifique.

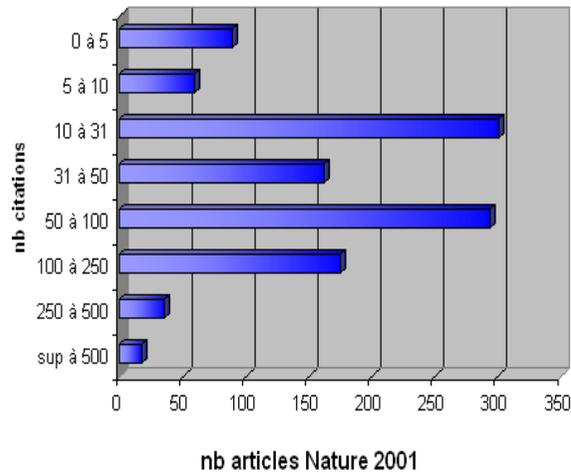


Figure 1 – Nombre de citations des articles publiés dans Nature

Le facteur d'impact d'un journal peut contribuer à la notoriété des articles qu'il édite de par sa grande visibilité. Ce facteur ne représente toutefois qu'une moyenne de citations des articles dont seuls quelques uns peuvent être très cités. A titre d'exemple pour les articles (hors lettres et revues) publiés dans Nature en 2001 la moyenne des citations obtenue en 2005 est de 74. Plus de 70% des articles sont cités moins de 74 fois. Seul 1,5% sont cités plus de 500 fois (figure 1). Toutefois, publier dans des journaux de prestige de fort impact représente une difficulté particulière qui doit être reconnue dans les analyses bibliométriques.

Une autre façon d'évaluer la répercussion ou la visibilité internationale d'une publication est de la situer dans le contexte international au sein de son domaine. Thomson ISI classe l'ensemble des publications mondiales dans 22 disciplines pour chaque année de publication et détermine les Top 1%, Top 10%, Top 20% etc... Chaque publication est assignée à une seule discipline. Une autre catégorisation est effectuée par l'ISI en 105 sous-disciplines en fonction du journal mais une publication peut appartenir à plusieurs sous-disciplines. La catégorisation en 22 disciplines est utilisée dans cette étude avec les seuils établis par l'ISI pour déterminer automatiquement pour chaque publication son « Top statut ». Ils se répartissent en Top 1% 10%, 20%, 50% (médiane mondiale). Pour chaque Institut, Unité, groupe ou chercheur, les publications peuvent être assignées à l'un de ces groupes. Cet indicateur est le seul qui s'affranchit du domaine et du temps.

A titre d'information, pour la recherche biomédicale en France entre 1996 et 2006 (publications avec au moins une affiliation française), 1,07% des publications appartiennent au Top 1%. Seuls 38% de ces articles sont publiés dans les grandes revues généralistes.

Différentes études démontrent que la reconnaissance d'un article n'est pas liée à l'IF du journal mais à sa qualité propre ; il peut être repéré par le nombre de ses citations, reflet de son utilisation par les pairs. Les articles très cités décrivent comme illustré précédemment de vraies découvertes. Si quelques « erreurs ou supercheries » ont pu faire couler beaucoup d'encre, la polémique soulevée par l'utilisation de cet indicateur est probablement dépassée.

Toutefois le nombre d'articles présents dans le Top 1% ou 10% ne doit pas être utilisé comme indicateur à lui seul car il dépend du potentiel humain et financier du groupe considéré.

La performance peut plutôt être exprimée par le pourcentage d'articles que le groupe peut positionner dans le Top 1% ou le Top 10% : indicateur que nous avons désigné à l'Inserm sous le terme de « % d'excellence ou de succès ».

Ces différents indicateurs d'excellence sont utilisés par les commissions européennes et dans de nombreuses études de « ranking » comme par exemple le « ranking de Shanghai » par l'université de Jiao Tong et le « Mapping of Excellence » par l'université de Leiden.

Les études bibliométriques ou de « ranking » sont de plus en plus nombreuses et doivent être menées avec une grande rigueur. Nous tenons à rappeler qu'il est possible de comparer des organismes, des Instituts de recherches, des départements, équipes ou même des chercheurs à condition de prendre en compte différents indicateurs qui se complètent. Il n'existe cependant pas de « formule miracle » capable de classer directement des équipes, des centres ou des Instituts.

Les indicateurs bibliométriques sont aussi « taille-dépendant » : ils sont statistiquement influencés par le niveau d'agrégation des données. Les scores d'un institut de taille modeste contribuant pour quelques dizaines de publications dans un domaine ne peuvent pas être comparés à ceux d'un Institut où sont agrégées les données de centaines de publications issues du travail d'un grand nombre de groupes multidisciplinaires.

De plus, il faut garder en mémoire que les « coutumes » bibliométriques peuvent être différentes dans diverses disciplines et que les thèmes émergents ont besoin de temps pour être cités. Le temps est aussi un élément à prendre en compte dans la propagation de l'information scientifique qui n'est pas identique d'un pays à l'autre. Dans la mesure où les analyses bibliométriques sont en vigueur depuis longtemps aux USA, que ce pays est un gros producteur et qu'il abrite un très grand nombre d'éditeurs, les articles des chercheurs américains sont plus rapidement et plus fréquemment cités que les autres. Ce phénomène se traduit par exemple par un plus fort pourcentage de leurs publications au Top 1% (plus de 70% alors qu'ils contribuent pour 30% de la littérature scientifique mondiale) et probablement une plus forte rapidité de citations. Ce dernier phénomène est à prendre en compte dans les analyses bibliométriques effectuées à 2 ans à un moment où les différences entre les pays ne sont pas stabilisées.

Vient à l'appui de nos conclusions une étude récente (US scientists dominate as journal Gatekeeper ; the Scientist march 2005¹) qui indique l'influence décisive des USA de par ses éditeurs sur « quand, où et quoi » est publié dans le monde. En effet, 53,8% de Gatekeepers sont originaires des USA versus 3,9% de France et 6,38 de l'Allemagne, 10,1 du Royaume Uni.

Ainsi à l'Inserm, nous préconisons pour les analyses bibliométriques l'utilisation d'un grand nombre d'indicateurs (voir encadré). Il est important d'indiquer que la bibliométrie ne remplace pas l'avis des pairs mais apporte une vision « objective » sur la visibilité des recherches.

- Nombre de publications (articles, lettres, revues) traduisant la production
- Citations de chaque publication, total des citations et indice de citation moyen traduisant la « visibilité » internationale
- Impact facteur moyen (IF moyen) traduisant la difficulté qu'a eue l'auteur à publier
- Nombre de publications au Top 1% et 10% : traduit la visibilité internationale des travaux adaptée à chaque domaine et tenant compte de l'année de la publication (seuils de citations ESI)
- Pourcentage de Top 1% et Top 10% traduisant le pourcentage d'excellence
- IF agrégé par sous domaine : IF moyen des journaux de chaque domaine : permet de comparer à la valeur moyenne du domaine
- Publications d'excellence: nb de publications au Top 1+10% + nombre de publications dont l'IF est supérieur ou égal à 2xIF agrégé
- H index : maximum de publications n qui sont cités n fois : traduit la production et l'impact de la personne ou du groupe en fonction du niveau de citation de ses publications

¹ Etude détaillée : *Scientometrics* 2005 62 :3

2.2 Constitution des corpus de publications

La recherche a été effectuée sur la France pour les années 2006-2007. Les citations des publications ont été recueillies jusqu'en mars-juin 2009 selon la date de téléchargement de chaque institut.

Procédure

- A. Pour chaque sous domaines des Instituts, le corpus de publications a été établi à partir de mots-clés établis en 2 temps :
- 1) Mots clés fournis par les experts de l'Institut,
 - 2) Recherche affinée et complétée par une recherche de mots clés soit avec le logiciel de Textmining Clementine (analyse textuelle) sur les résumés d'articles dans les journaux de spécialité, soit par une analyse manuelle sur les articles des meilleurs des journaux de spécialité. Le corpus obtenu est analysé pour éliminer les publications hors domaines soit sur le titre du journal soit sur le titre de la publication.
- B. Des Tests de validité du corpus obtenu ont été effectués : à partir des meilleurs journaux de la spécialité (classés par IF dans la base du JCR). Nous testons la requête obtenue avec les mots clés majeurs de la discipline sur le corpus des journaux. Si nous recueillons au moins 80% des publications, alors la requête est validée et le corpus est établi sur l'ensemble des journaux de la base. Une vérification est aussi effectuée avec des noms de chercheurs du domaine pour lequel 80% des publications doivent être retrouvées par la requête.

Lorsque le corpus est établi définitivement, les articles sont importés vers un logiciel de gestion bibliographique (Endnote) selon un filtre établi qui permet de n'exporter que les champs utiles pour le traitement des données (références de l'article, abstract, mots clés, nombre de citations et adresses). Les données sont ensuite exportées à l'aide d'une application spécifique développée par la société AGUIDEL vers Access puis Excel ou Statistica pour établir des données statistiques (ex : nombre d'articles, nombre de citations, index de citation) et des graphiques (ex : répartition des articles français par région).

3 Mise en œuvre des indicateurs

Pour chaque Instituts, nous avons établis différents tableaux de données à l'aide des corpus de publications : données bibliométriques de la France par sous-domaines, comparaisons internationales, productivité et cartographies (répartition géographiques des sous domaines en nombre de publications, réseaux de collaborations des organismes d'un Institut).

3.1 Publications 2006-2007 de la France (Articles, lettres, revues)

Institut	Sous-Domains	Nb Pub	Total Citations	Indice de Citation moy	H-Index	IFm	Nb Top1% WOS*	% Top1% WOS	Nb Top10% WOS**	% Top 10% WOS
Institut	Sous-Domaine A	Nombre total articles, lettres, revues Total des citations des publications 2006-2007 reçues à la date de chargement	Total des citations/nb pub	Max de pubs n citées n fois	Moyenne des Impact Facteur 2008	Utilisation des champs disciplinaires de l'ISI affectés en fonction des journaux (22 disciplines) et application des seuils ESI (voir page 13) . Pour les journaux généralistes, les publications sont affectées manuellement à un domaine.				
Institut	Sous-Domaine B									
Institut	Sous-Domaine C									
Institut	Sous-Domaine D									
Institut	Sous-Domaine E									
Institut	Sous-Domaine F									
Institut	Sous-Domaine G									
Institut	Sous-Domaine H									
Institut	Ces données concernent les publications de tous les sous-domaines de l'Institut réunis sans doublons.		Attention, le H-Index est sensible aux domaines de recherche comme les citations.	ⓐ L'IF utilisé pour ce calcul est celui de 2008 quelle que soit l'année de publication de l'article. Les résultats de cette colonne doivent donc être pris à titre indicatif. Les IF des journaux évoluent dans le temps: par exemple Lancet a gagné 10 points entre 2003 (IF:18,31) et 2007 (IF:28,64).						

* Les Top 1% WOS représentent les publications de la France dans le premier percentile du corpus mondial des 22 domaines de l'ISI Web of Knowledge

** Les Top 10% WOS représentent les publications de la France dans le premier décile (Top1%+10%) à partir des seuils de citations établis pour chaque domaine et chaque année

Tableau 1

3.1.1 Répartition des publications 2006-2007 par sous-domaines

Les sous-domaines ont été définis à l'aide de mots clés et des journaux de spécialités (*Tableau 2*)

Sous domaines	Sous-Domaine A	Sous-Domaine B	Sous-Domaine C	Sous-Domaine D
Sous-Domaine A	aaaa	0	0	0
Sous-Domaine B	0	bbbb	0	0
Sous-Domaine C	0	0	cccc	0
Sous-Domaine D	0	0	0	dddd

Ce tableau indique le nombre de publications total d'un sous-domaine ainsi que le nombre de publications communes entre 2 sous-domaines

Tableau 2

3.2 Comparaison des données internationales des publications 2006-2007

Nous avons utilisé les requêtes déjà établies pour les indicateurs de la France pour chaque sous-domaine des Instituts mais sans chargement de corpus. Les données du tableau ont été obtenues directement de la base de l'ISI qui donne pour chaque requête effectuée le nombre de publications, la position du pays en nombre de publications, le nombre de citations total et moyen, le H index et le nombre d'auteurs.

La comparaison s'effectue avec les pays représentatifs en nombre de publications (*Tableau 3*).

Institut	Sous-domaines	Nb pub	Total citations ¹	Indice de citations moyen	
					h-index ²
Institut	Sous-Domaine A	0	0	0	0
	Sous-Domaine B	0	0	0	0
	Sous-Domaine C	0	0	0	0
	Sous-Domaine D	0	0	0	0

Tableau 3

3.3 Productivité

Comparaison populations / potentiel recherche en 2007

A l'aide des données obtenues par l'OCDE², nous avons établi un ratio du potentiel humain de la recherche par pays pour l'année 2007 pour positionner la France dans un contexte international (Tableau 4). Chaque ITMO a également recensé l'ensemble des équipes de recherche et le potentiel humain qui les constitue. Pour comparer la productivité par sous-domaines des différents pays avec la France, nous utilisons le ratio par rapport aux nombres de chercheurs (ETP) en France (Ratio A Tableau 2), puis par sous domaines des Instituts avec le nombre de publications. Le rapport de ces 2 ratios permet de situer pour chaque sous-domaine la productivité du pays par rapport à celui de la France (Tableau 5).

	Population mondiale OCDE	ETP Chercheurs OCDE	Ratio (potentiel humain de la recherche)
France	61707000	211129	0,34
Royaume Uni	60975000	183535	0,30
Allemagne	82257000	286000	0,35
Italie	58880000	88430	0,15
USA	301621200	1387882	0,46

Ratio = ETP Chercheurs / Population mondiale

Tableau 4

		Pays 1	Pays 2	Pays 3	Pays 4
	Ratio A : ETP pays/ETP France	0,00	0,00	0,00	0,00
Sous-Domaine A	Ratio B : Nbre pubs pays/Nbre pubs France	0,00	0,00	0,00	0,00
	Ratio B/Ratio A	0,00	0,00	0,00	0,00
Sous-Domaine B	Ratio B : Nbre pubs pays/Nbre pubs France	0,00	0,00	0,00	0,00
	Ratio B/Ratio A	0,00	0,00	0,00	0,00
Sous-Domaine C	Ratio B : Nbre pubs pays/Nbre pubs France	0,00	0,00	0,00	0,00
	Ratio B/Ratio A	0,00	0,00	0,00	0,00

Si le Ratio est > 1, alors la productivité est supérieure à celle de la France
Si le Ratio est < 1, alors la productivité est inférieure à celle de la France

Tableau 5

² Source OCDE, Population mondiale <http://stats.oecd.org/index.aspx?queryid=254>

Source OCDE, Principaux indicateurs de la science et de la technologie, Volume 2008/2 <http://www.oecd.org/dataoecd/9/44/41850733.pdf>

3.4 Universités et organismes

Les affiliations présentes dans les notices bibliographiques ont été corrigées manuellement, les CHU et hôpitaux universitaires ont été rattachés à l'Université correspondante. Pour toutes les unités Inserm mixtes avec l'Université, l'affiliation indique les deux partenaires. Pour les unités CNRS ayant un label UMR, l'affiliation indique le partenaire de mixité. Les résultats sont exprimés en comptes entiers (1 pour chaque partenaire). Pour chaque institut, les universités ont été classées en nombre de publications, en nombre de publications d'excellence, en pourcentage de publications d'excellence.

- **Nombre de publications total par organisme** (Tableau 6)

Organismes	Nb de Publications d'excellence	%
Universités		
INSERM		
CNRS		
INRA		
Institut Pasteur		
CEA		
...		

Tableau 6

- **Publications des 10 universités publiant le plus** (Tableau 7)

Le tableau indique le nombre de publications total de chaque Université (incluant celles en collaboration avec un partenaire EPST) ainsi que la part de l'Inserm et du CNRS dans ces publications.

Rang	Universités	Nb de Publications total	Inserm		CNRS	
			Nb pub en partenariat avec Univ	%	Nb pub en partenariat avec Univ	%
1	Univ Paris 05					
2	Univ Paris 06					
3	Univ Paris 07					
4	Univ Lyon					
5	Univ Lille					
6	Univ Paris 11					
7	Univ Aix Marseille					
8	Univ Bordeaux					
9	Univ Toulouse					
10	Univ Montpellier					

Tableau 7

- **Top 10 des affiliations les plus fréquentes dans les publications** (Tableau 8)

Pour les Universités, les résultats sont sans les organismes (les affiliations de type Organisme-Univ ne sont pas prises en compte), ce qui représente leur propre production.

Rang	Affiliations	Nb de Publications
1	Univ Paris 05	
2	Univ Paris 06	
3	Univ Paris 07	
4	Univ Lyon	
5	Univ Lille	
6	Univ Paris 11	
7	Univ Aix Marseille	
8	Univ Bordeaux	
9	INSERM-Univ Paris 06	
10	Univ Toulouse	

Tableau 8

- **Classement par Institut de la production des Universités**

Pour chaque institut, les universités ont été classées en nombre de publications décroissant (incluant celles en collaboration avec un partenaire EPST). L'université de rang 1 est donc celle qui publie le plus dans le domaine (Tableau 9).

Institut 1		Institut 2		Institut 3		Institut 4	
Université	Rang	Université	Rang	Université	Rang	Université	Rang
Univ Paris 05		Univ Paris 05		Univ Paris 05		Univ Paris 05	
Univ Paris 06		Univ Paris 06		Univ Paris 06		Univ Paris 06	
Univ Paris 07		Univ Paris 07		Univ Paris 07		Univ Paris 07	
Univ Paris 11		Univ Paris 11		Univ Paris 11		Univ Paris 11	
Univ Aix Marseille		Univ Aix Marseille		Univ Aix Marseille		Univ Aix Marseille	
Univ Bordeaux		Univ Bordeaux		Univ Lille		Univ Bordeaux	
Univ Lille		Univ Lille		Univ Lyon		Univ Lille	
Univ Lyon		Univ Lyon		Univ Montpellier		Univ Lyon	
Univ Strasbourg		Univ Montpellier		Univ Strasbourg		Univ Montpellier	
Univ Toulouse		Univ Toulouse		Univ Toulouse		Univ Toulouse	

Tableau 9

- **Nombre de publications d'excellence des universités par Institut**

Pour chaque institut, les universités ont été classées en nombre décroissant de publications d'excellence. L'université de rang 1 est donc celle qui a le plus de publications d'excellence dans le domaine (Tableau 10).

Institut 1		Institut 2		Institut 3		Institut 4	
Université	Nb Pub Excellence	Université	Nb Pub Excellence	Université	Nb Pub Excellence	Université	Nb Pub Excellence
Univ Paris 06		Univ Paris 05		Univ Paris 05		Univ Paris 05	
Univ Paris 11		Univ Paris 06		Univ Paris 06		Univ Paris 07	
Univ Paris 05		Univ Paris 07		Univ Paris 11		Univ Paris 06	
Univ Paris 07		Univ Paris 11		Univ Paris 07		Univ Paris 11	
Univ Lyon		Univ Lille		Univ Montpellier		Univ Lyon	
Univ Aix Marseille		Univ Lyon		Univ Aix Marseille		Univ Aix Marseille	
Univ Toulouse		Univ Toulouse		Univ Strasbourg		Univ Toulouse	
Univ Bordeaux		Univ Strasbourg		Univ Lyon		Univ Lille	
Univ Lille		Univ Bordeaux		Univ Toulouse		Univ Montpellier	
Univ Montpellier		Univ Montpellier		Univ Bordeaux		Univ Strasbourg	

Tableau 10

- **Classement par Institut des Universités en termes d'Excellence**

Pour chaque institut, le pourcentage d'excellence des universités a été normé par rapport au pourcentage de la France dans le domaine. L'université de rang 1 est donc celle qui a le plus fort taux d'excellence dans le domaine (Tableau 11).

Institut 1		Institut 2		Institut 3		Institut 4	
Université	Rang	Université	Rang	Université	Rang	Université	Rang
Univ Nice		Univ Paris 11		Univ Evry		Univ Reims	
Univ Paris 11		Univ Lille		Univ Nice		Univ Angers	
Univ Paris 12		Univ Strasbourg		Univ Angers		Univ Bordeaux	
Univ Nantes		Univ Toulouse		Univ Nantes		Univ Dijon	
Univ Paris 13		Univ Nice		Univ Strasbourg		Univ Poitiers	
Univ Paris 07		Univ Nantes		Univ Bordeaux		Univ Paris 05	
Univ Toulouse		Univ Caen		Univ Paris 11		Univ St Etienne	
Univ Angers		Univ Montpellier		Univ Grenoble		Univ Lille	
Univ Dijon		Univ Paris 06		Univ Paris 05		Univ Paris 12	
Univ Paris 06		Univ Paris 07		Univ Paris 07		Univ Grenoble	

Tableau 11

3.5 Réseau de collaboration des organismes d'un Institut

L'analyse des réseaux de collaborations identifie les organismes ayant collaboré de manière plus ou moins importante en cosignant les publications (la surface des points est proportionnelle au nombre de liens présents dans les réseaux). Les analyses de réseaux sont effectuées à l'aide du logiciel Reseau-Lu qui utilise différents algorithmes construisant des matrices d'associations et qui positionnent chaque point en fonction des relations spécifiques avec les autres. Les données (auteurs, ville, pays, mots clés, ...) sont transformées en 3 étapes : 1) matrice de poids, 2) matrice d'associations escomptées, 3) matrice de proximité. Ci-dessous un exemple de réseau (Figure 2) : les partenariats de la société Aguidel³ (développeur du logiciel Reseau-Lu).

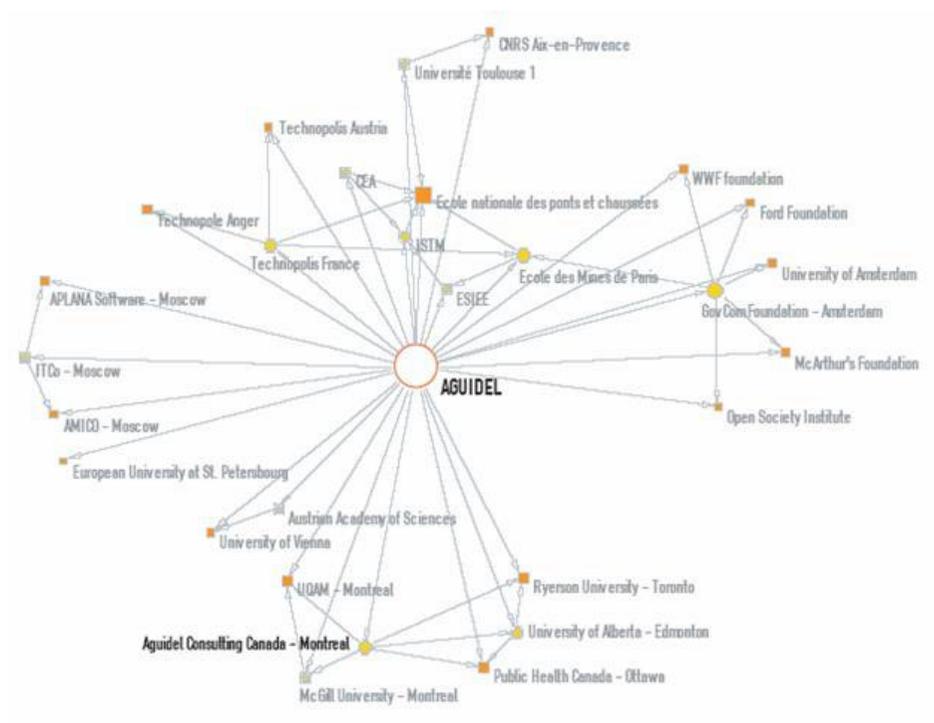


Figure 2

³ <http://www.aguidel.com/fr>

3.6 Représentation graphique 3D

La représentation 3D est fonction du nombre total de publication, du facteur d'impact moyen et du pourcentage de publications d'«excellence».

Les publications d'excellence sont les publications appartenant au Top10% (1%+10%) ainsi que les publications parues dans des journaux dont le facteur d'impact est supérieur à 2,29 fois le facteur d'impact agrégé de la catégorie.

Le facteur d'impact agrégé d'une catégorie est calculé par le Journal Citation Reports (JCR) de la même manière que le facteur d'impact d'un journal.

Ex : Facteur d'impact agrégé 2008 de la catégorie Maladies infectieuses directement accessible dans le Web of Science :

Category (linked to category information)	Total Cites	Median Impact Factor	Aggregate Impact Factor	Aggregate Immediacy Index	Aggregate Cited Half-Life	# Journals	Articles
INFECTIOUS DISEASES	290946	2.516	3.644	0.794	5.7	51	8640

Ainsi si on prend l'exemple d'un institut pour lequel l'IF agrégé de la catégorie (ou des catégories) correspondante(s) est de 3,8 ; les publications parues dans un journal dont le facteur d'impact est supérieur ou égal à 9 ($3,8 \times 2,3$) sont considérées comme faisant partie de l'excellence.

Dans l'exemple des Maladies Infectieuses, seuls 2 journaux du domaine sont inclus outre les grands généralistes : Lancet Infectious Diseases (IF=13,1) et Plos Pathogen (IF=9,1).

Pour la représentation graphique en 3D (présentée page suivante Figure 3), le facteur d'impact moyen et le pourcentage de publications d'excellence sont normés par rapport à ceux de la France dans le domaine. Ainsi, les zéro des axes correspondent à l'IF moyen et au pourcentage d'excellence moyen de la France dans le domaine. Les ronds représentent la taille de l'élément observé en nombre de publications. Il y a 2 sortes de représentations :

- Par Institut, le positionnement des régions
- Par région, le positionnement des Instituts

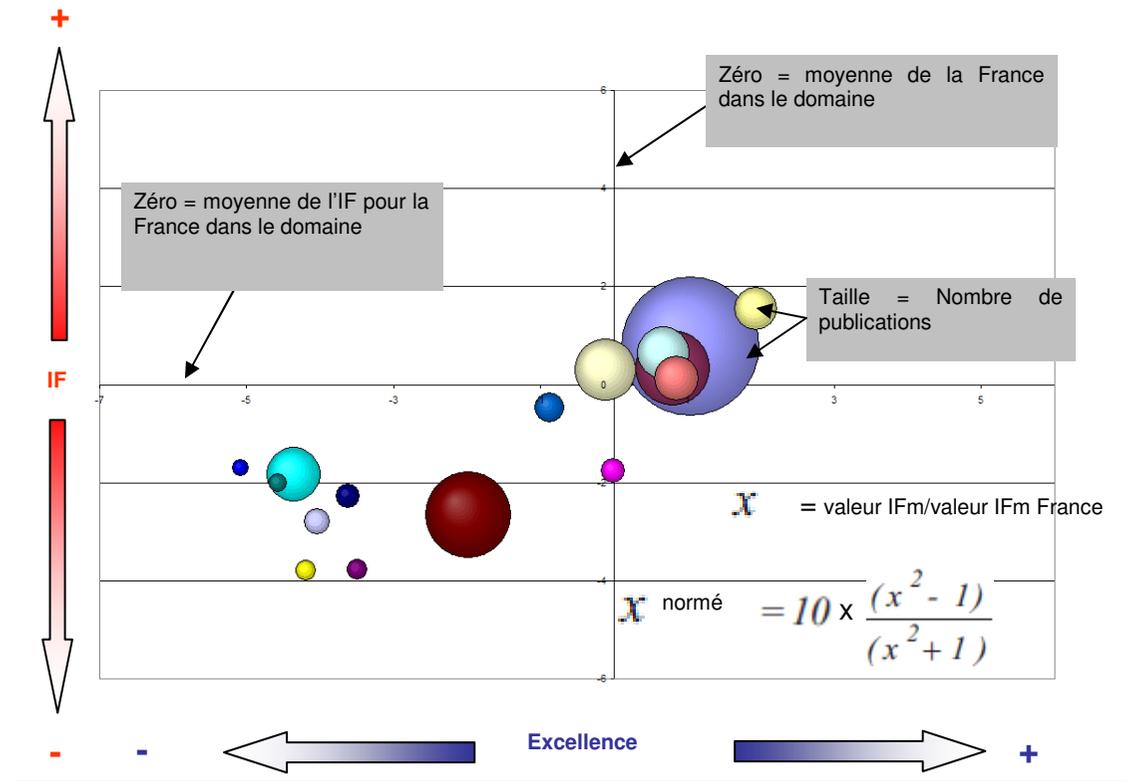


Figure 3

4 Cartographies

A l'aide des adresses des publications du corpus initial, il est possible de faire une répartition géographique des sous-domaines d'un Institut en nombre de publications, d'établir un réseau de collaborations entre les différents organismes de recherche et une représentation des régions 3D par Institut.

Répartition géographique des publications par sous domaine

Pour cette cartographie, le logiciel utilisé est Arctique⁴ : les cartes sont constituées sur le nombre de publications (Figure 4). Les villes des adresses ont été extraites et une table des correspondances villes/régions a été utilisée pour associer chaque publication à 1 ou plusieurs régions (si plusieurs régions, la publication compte pour 1 pour chaque région).

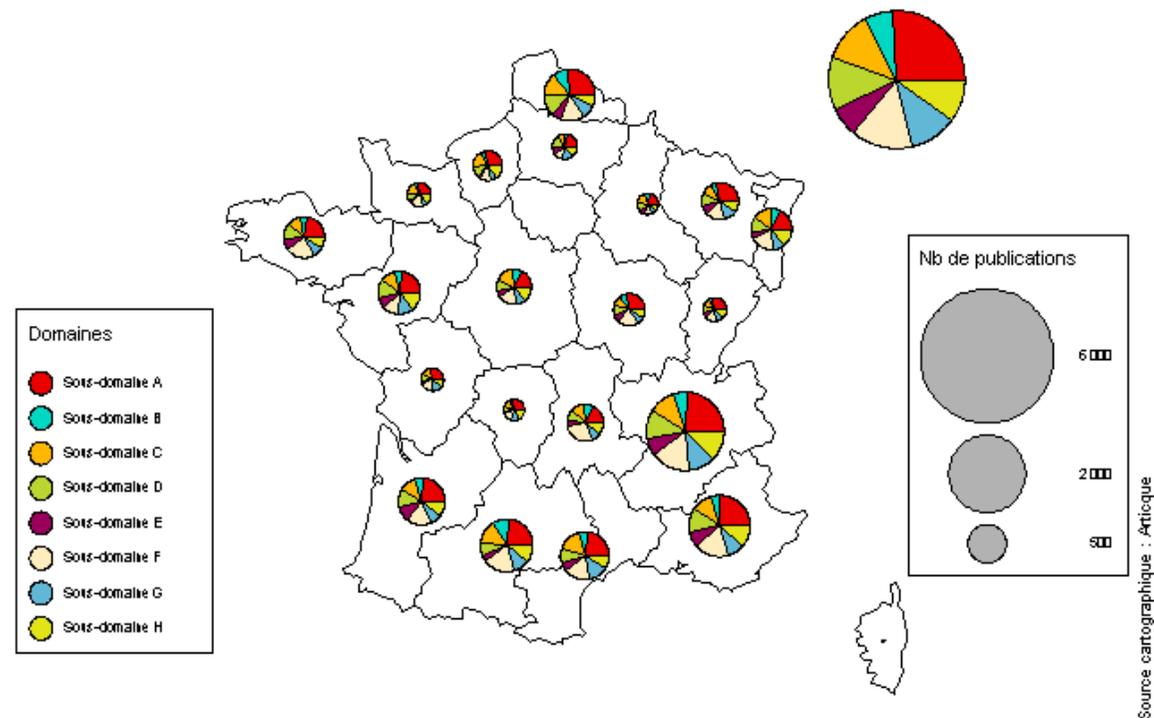


Figure 4

⁴ www.arctique.com

5 Conclusion

La France se situe actuellement au **cinquième rang mondial de la production scientifique dans le domaine des sciences de la vie et de la santé**. Les publications de ses grands organismes de recherche ont un indice d'impact supérieur à 1, c'est-à-dire que leur influence mesurée par citations se situe au-dessus de la moyenne internationale. Cette recherche de qualité est assurée par une pluralité d'acteurs : huit grands établissements publics auxquels s'associent les universités et les CHU. Aviesan est née de la **volonté d'accroître encore ces performances de la recherche française**, en favorisant sa cohérence, sa créativité et son excellence. Cette mission appelle une coordination scientifique des grandes thématiques de recherche, transversales à tous les organismes, et une coordination opérationnelle des projets, des ressources et des moyens, d'où la création des ITMOs.

L'ensemble des indicateurs bibliométriques décrits ont permis de dresser un état des lieux de la recherche biomédicale en France.

Ces données vont alimenter la réflexion stratégique propre à chaque ITMO et servent à en illustrer le document stratégique que l'on retrouve sur le site AVIESAN www.aviesan.fr.

6 Bibliographie

- [1] HAEFFNER-CAVAILLON N, GRAILLOT-GAK C (2009) *The use of bibliometrics indicators to help peer-review assessment*. Arch. Immunol. Ther. Exp., 2009, 57, 33-38
- [2] AKSNES DW, TAXT RE (2004) *Peer reviews and bibliometric indicators: a comparative study at a Norwegian university*. Res. Evaluation 13:33-41
- [3] CHERUBINI P (2008) *Impact factor fever*. Science 322:191
- [4] GARFIELD E (2006) *The history and meaning of the journal impact factor*. JAMA 295:90-93
- [5] HIRSCH JE (2007) *Does the H index have a predictive power?* Proc. Natl. Acad. Sci. USA 104 :10193-19198
- [6] JUNQUERA B, MITRE M (2007) *Value of bibliometric analysis for research policy : A case study of Spanish research into innovation and technology management*. Scientometrics 71:443-454
- [7] NEDHERHOF AJ (2008) *Policy impact of bibliometrics rankings of research performance of departments and individuals in economics*. Scientometrics 74:163-174

7 Annexe

Liste des Instituts Thématiques Multi-Organismes ITMO

- **Bases moléculaires et structurales du vivant**
- **Biologie cellulaire, développement et évolution**
- **Cancer**
- **Circulation, Métabolisme, Nutrition**
- **Génétique, génomique et bioinformatique**
- **Immunologie, Hématologie, Pneumologie**
- **Microbiologie, Maladies Infectieuses**
- **Neurosciences, Neurologie, Psychiatrie**
- **Santé Publique**
- **Technologies pour la Santé**