

**Production coopérative d'indicateurs inter-institutionnels
de politique scientifique**

Indicateurs bibliométriques des institutions publiques de recherche
Année 1997 – hors sciences humaines et sociales

Essai méthodologique

Groupe "IPSIR" : indicateurs de publications scientifiques
des institutions de recherche

Travail réalisé à l'OST

- ◆ coordonné par : Michel Zitt (OST/Inra)
Serge Bauin (Unips-CNRS)
Ghislaine Filliatreau (DR-MR/OST)
- ◆ avec les représentants des institutions concernées (1) : CEA (F. Rousseau-Hans, D. de Prunelé), Cemagref (N. Delherbe, O. Hologne), CNRS (M. Crance), Inra (E. Bassecoulard, A. Chartier, H. Pampouille), Inserm (N. Barberousse, N. Pinhas), IRD (M. Stjepanovic), Universités (J-P Finance, M. Laurent).
- ◆ et la collaboration de S. Ramanana-Rahary et de N. Teixeira (OST)

Septembre 2002

(1) D'autres institutions ont été associées à certaines des étapes de travail du groupe (Cnes, Ined, Inrets, Inria). Pour diverses raisons (place de la production académique, mode de diffusion principal, etc.), le repérage de leurs publications n'a pas été entrepris pour le présent exercice.

INTRODUCTION

Dans le cadre de la "production coopérative d'indicateurs de politique scientifique", le groupe "IPSIR" (publications scientifiques) a pour mission d'élaborer des indicateurs bibliométriques communs aux organismes publics de recherche et universités.

Les indicateurs bibliométriques sont calculés à partir des articles publiés par les laboratoires de recherche dans les journaux scientifiques. Ces articles, qui contiennent des connaissances certifiées sont l'une des productions principales des laboratoires de recherche académique, sachant qu'il y a, bien entendu, d'autres dimensions dans l'activité des laboratoires.

Les résultats ici présentés sont exploratoires : ils sont limités à une seule année de publication un peu ancienne (1997)¹, et à une partie des institutions de recherche du secteur public. De plus, ils reposent sur un repérage des articles assez simple. Quelques-uns seulement des indicateurs possibles sont donc présentés ici, pour illustrer la faisabilité et les potentialités d'une exploration bibliométrique institutionnelle de la recherche française. L'étape suivante nous permettra, avec des indicateurs plus précis et des données plus récentes, d'explorer la dynamique à l'œuvre au cours des dernières années.

Il convient enfin de souligner que la situation de la recherche a largement évolué depuis 1997², en particulier du fait de la généralisation rapide de la mixité des laboratoires. Cette mixité est devenue quasi systématique entre les universités et le CNRS, mais elle se généralise également entre les universités et d'autres EPST comme l'Inserm et, de plus en plus souvent, entre plusieurs de ces partenaires. Cette évolution aura d'ailleurs des conséquences importantes sur la répartition institutionnelle des publications scientifiques dans les prochaines années.

DONNÉES, MÉTHODES et LIMITES

Source

La base de données utilisée est le seul *Science Citation Index* (SCI) de l'Institute for Scientific Information, outil habituel des bibliométriciens dans les domaines des sciences de la matière et de la vie. C'est une base très sélective, plutôt orientée vers l'activité scientifique académique, et de bonne qualité pour les disciplines fondamentales bien internationalisées. Sa représentativité est moins bonne dans les disciplines appliquées, de terrain, ou dans les domaines à forte tradition nationale, ou encore dans les disciplines (notamment l'informatique) où les informations passent aussi par d'autres canaux de diffusion. Rappelons aussi que SCI n'est pas construit comme un "échantillon représentatif" de la production scientifique mondiale, mais plutôt comme un recensement des meilleures sources académiques.

Enfin, les bases homologues pour les sciences humaines et sociales (SSCI ou AHCI), trop insatisfaisantes quant à leur couverture pour la plupart des disciplines et sans alternative crédible à ce jour, ne peuvent être utilisées en l'état. Cela ne nous a pas permis d'aborder, pour ce travail, les productions en sciences humaines et sociales, ce qui introduit un biais évident dès lors que l'on parle de "la" production scientifique de l'ensemble du monde académique.

¹Au départ du chantier, le fait de travailler sur l'année 1997 avait plusieurs avantages dans le cadre d'une opération pilote : pour cette année là, toutes les institutions disposaient des renseignements nécessaires à l'identification des articles et, surtout, le recul était suffisant pour les études d'impact.

²Il faut également mentionner que les articles sont normalement publiés plusieurs mois (de l'ordre de 3 à 24 mois selon les disciplines et les journaux) après la fin des travaux qui ont permis de dégager les résultats présentés.

Types de documents

Les indicateurs bibliométriques portent en général sur quatre types de documents : articles, articles de synthèse, notes, lettres, en excluant les comptes-rendus de congrès ("proceedings" ou CRC). Dans cette étude, nous avons choisi de calculer certains indicateurs avec et sans cette extension CRC, ce qui modifie le volume de publications d'environ 15 % pour la France, mais avec de fortes différences selon les institutions. Ainsi, pour le CEA, l'accroissement dû aux CRC est très important (environ 34 %). Dans le chapitre sur les nombres de publications, des résultats avec et sans CRC sont présentés. Dans les chapitres co-publications et visibilité (citations) il a été convenu entre institutions participantes de se limiter aux publications hors CRC.

Classification disciplinaire

La classification disciplinaire utilisée est la classification standard de l'OST, et résulte d'une agrégation des spécialités proposées par le SCI ("subject category codes"), classifications qui sont définies au niveau des journaux (voir Annexe). L'informatique n'y est pas classée dans les mathématiques mais dans les sciences pour l'ingénieur ; les sciences de l'univers, outre l'astronomie et les géosciences, comprennent les sciences et technologies de l'environnement et, au titre de la "terre fluide", l'océanographie, la biologie marine et l'hydrobiologie.

Repérage des institutions

Chacune des institutions participantes (CEA, Cemagref, CNRS, Inra, Inserm, IRD et universités) a repéré les publications de ses laboratoires dans le SCI, pour l'année de publication 1997.

Les principes de repérage ont été les suivants :

CEA	Sont pris en compte les laboratoires du CEA civil et de défense, ainsi que de l'IPSN et les unités mixtes.
Cemagref	Les affiliations Cemagref, à quelques unités près, sont aisément identifiées dans le SCI.
CNRS	Sont prises en compte les unités évaluées et financées par le CNRS, propres, mixtes et associées, avec une adresse française.
Inra	Sont prises en compte les unités de recherche propres ou associées répertoriées dans l'annuaire des structures de l'Inra. Pour une part de l'activité des structures d'enseignement supérieur agronomique associé à l'Inra, il subsiste une marge d'erreur (de l'ordre de 5 %) due à l'imprécision des affiliations au niveau laboratoire.
Inserm	Sont prises en compte les publications des "formations de recherche" de l'Inserm. Les publications des chercheurs Inserm hors formations n'ont pas été comptabilisées à ce stade.
IRD	Ne sont prises en compte pour l'instant que les notices avec adresse IRD (ou adresse d'un co-publiant avec IRD) en France métropolitaine. Elles représentent environ les trois quarts du total général IRD. Les lignes IRD sont donc notées IRD-France.
Universités	Ont été considérées toutes les structures évaluées et/ ou co-financées par la Mission Scientifique Universitaire, y compris les unités mixtes avec d'autres institutions - essentiellement le CNRS s'agissant de l'année 1997. Cette large mixité explique d'ailleurs les difficultés de repérage rencontrées pour des structures comme les instituts de recherche ou les observatoires, dans lesquels cohabitent des unités mixtes et des unités non mixtes. Pour l'année 1997, ces problèmes n'ont pas toujours pu être résolus de manière incontestable, et le choix a été fait d'approcher les résultats par défaut pour les universités. Ainsi beaucoup d'observatoires astronomiques ou des structures comme le LURE n'ont pas été pris en compte pour ce premier repérage.

Dans tous les tableaux, la ligne "Autres FRA" identifie les publications avec adresse française sans revendication d'une des institutions mentionnées ci-dessus. Les publications des "autres acteurs français" (institutionnels ou privés), en collaboration ou association avec l'une des sept institutions prises en compte, ne peuvent donc être identifiées compte tenu du mode de repérage. Il y a, par conséquent, une forte minoration de la place de ces "autres institutions".

Principes de comptage

Après le repérage de leurs publications par chacune des sept institutions participantes, les calculs bibliométriques ont été réalisés par l'OST à partir de sa base de données.

Le principe de comptage utilisé pour chaque institution est celui du *compte de présence*. Dans cette logique, le nombre de publications attribuées à une institution correspond aux publications où cette institution apparaît au moins une fois. On ne fait pas la différence entre la présence d'un seul ou de plusieurs laboratoires de cette institution.

Cependant, le recouvrement institutionnel est souvent très important, comme dans le cas des universités et du CNRS. En conséquence, la production des structures mixtes entre ces deux institutions est attribuée à chacune d'entre elles, ce qui veut dire par exemple que les données relatives aux universités incluent celles de la plus grande partie des laboratoires du CNRS, qui étaient déjà mixtes avec l'université en 1997 ; réciproquement, les données relatives au CNRS recouvrent largement celles des laboratoires universitaires³.

Les nombres ou pourcentages correspondants ne peuvent donc être consolidés entre institutions, et la somme des parts françaises des différentes institutions est bien supérieure à 100 % en raison du grand nombre d'articles produits par des laboratoires mixtes, auxquels s'ajoutent bien entendu les articles produits en collaboration - et donc co-signés - par des laboratoires d'institutions différentes.

Le compte de présence permet à chaque institution d'estimer la part de la science mondiale dans laquelle elle est active, part plus grande que sa "contribution" relative telle qu'elle apparaît lors d'un décompte fractionnaire (compte adopté par exemple dans le rapport biennal de l'OST). Outre la simplicité de l'interprétation, l'avantage du compte de présence est la stabilité en cas d'information incomplète sur les institutions participantes. C'était la seule méthode de comptage ici praticable compte tenu du mode de repérage adopté par les organismes. Les inconvénients sont l'absence d'additivité, d'où l'instabilité lorsque l'on change le niveau d'observation (du laboratoire à l'institution, de la région au pays, etc.). Enfin, ce mode de comptage favorise les institutions qui collaborent le plus.

Nous avons conservé une approche fractionnaire pour les disciplines (un article publié dans une revue assignée à deux disciplines est compté pour ½ dans chacune, etc.).

Les totaux France obtenus ici ne sont pas directement comparables avec les indicateurs standards publiés par l'OST, d'abord en raison du mode de comptage, ensuite en raison de la différence de couverture des bases utilisées. La différence du type de compte adopté pour les citations entraîne des différences importantes pour les indicateurs de visibilité (impact relatif).

Contexte et limites des résultats

Le tableau 1 présente le nombre de publications décomptées pour chacune des institutions participantes, avec ou sans prise en compte des résumés des présentations lors de congrès (CRC).

³ Fin 1996, environ 900 unités mixtes et associées du CNRS l'étaient avec l'enseignement supérieur (sur 1 100 unités mixtes et associées, et 1 300 unités de recherche CNRS). Les établissements de l'enseignement supérieur, c'est-à-dire relevant du ministère de l'éducation nationale, incluaient en 1996 la majorité des écoles d'ingénieurs.

Tableau 1 : Nombre de publications où apparaît l'institution

	Nb. Docs. CRC* exclus	Nb. Docs. CRC inclus	Soit en plus (en %)
CEA	1 803	2 414	34
Cemagref	45	52	16
CNRS	17 838	20 698	16
Inra	2 268	2 409	6
Inserm	4 179	4 582	10
IRD-France	332	354	7
Universités	24 365	27 660	14
Autres FRA**	3 072	3 751	22
France	34 197	39 329	15
Monde	511 835	577 102	13

*comptes-rendus de congrès

année 1997

**chiffres par défaut, voir texte

Attention : pas d'additivité en colonnes (comptages multiples)

L'interprétation doit naturellement tenir compte des tailles des institutions. On pourra par exemple approcher ces proportions en se reportant au rapport "Démographie de la recherche scientifique française", publié par l'OST en janvier 2000⁴ et portant sur des données 1999. En ne s'intéressant qu'aux enseignants-chercheurs et chercheurs hors SHS, on y voit par exemple que le personnel chercheur des universités représente environ 14 700 "équivalents temps plein recherche", contre environ 9 400 pour le CNRS, près de 2 000 pour l'Inserm, environ 1 500 pour l'Inra, 600 pour l'IRD et moins de 100 pour le Cemagref. En 1997 cependant, les effectifs universitaires étaient sensiblement plus faibles en proportion, car ce sont les universités qui ont bénéficié des plus forts recrutements au cours des cinq dernières années.

Ces données de personnels ne sont pas toujours simples à utiliser : pour reprendre les cas du CNRS et des universités, l'annexe chiffrée du rapport d'activité 2000 du CNRS⁵ montre que, hors SHS, 8 300 chercheurs du CNRS seulement sur les 9 400 que compte l'organisme travaillent dans des laboratoires repérables comme appartenant au CNRS cependant que travaillent également, dans ces mêmes laboratoires, de très nombreux enseignants-chercheurs représentant 6 200 équivalents temps plein recherche. Par ailleurs, ces chiffres ne tiennent pas compte des nombreuses autres catégories de personnels présents dans les laboratoires et susceptibles de publier (ingénieurs de recherche, ingénieurs agronomes, visiteurs étrangers, étudiants en thèse ou en stage, etc.).

Un autre élément très important pour la lecture des résultats concerne les universités, qui constituent un ensemble d'établissements hétérogène au regard de l'activité de recherche. À titre d'illustration nous avons inclus, dans le chapitre consacré aux comptages de citations, un exemple éloquent sur la comparaison d'une université de grande taille, active en recherche académique de niveau international, et une université de taille modeste axée principalement sur la fonction d'enseignement et - comme souvent en ce cas - pratiquant une recherche plutôt plus appliquée. Cet exemple montre bien que les chiffres globaux donnés pour les universités recouvrent en fait de notables disparités.

Indicateurs

Nous présentons successivement les trois familles classiques d'indicateurs bibliométriques :

- la production scientifique : le décompte des publications et la spécialisation (section I),
- les collaborations scientifiques : les co-publications (section II),
- la visibilité : le comptage des citations et les différentes mesures d'impact (section III).

⁴<http://www.obs-ost.fr>

⁵<http://www.cnrs.fr/cw/fr/band/cnrs/rapports/rap00/chiffres.pdf>

Les précautions d'usage dans l'interprétation des données statistiques s'appliquent aux données bibliométriques (par exemple, s'il s'agit d'établir un profil disciplinaire pour un institut de taille modeste), et tout particulièrement pour certains indicateurs dont les distributions statistiques sont atypiques, comme les impacts.

I. PRODUCTION SCIENTIFIQUE

Le tableau 2 permet une première approche des positionnements des institutions, en nombre et en pourcentage (non additifs) par rapport à la France

Tableau 2 : Nombre de publications où apparaît l'institution et poids dans la France (1997)

		Biologie Fondamentale		Recherche Médicale		Biologie Appliquée - Ecologie		Chimie		Toutes disc. confondues	
CEA	CRC exclus	293	4,3 %	103	1,2 %	37	1,9 %	259	5,1 %	1 803	5,3 %
	CRC inclus	311	4,3 %	109	1,1 %	39	1,8 %	376	6,1 %	2 414	6,1 %
Cemagref	CRC exclus	3	0,0 %	1	0,0 %	14	0,7 %	1	0,0 %	45	0,1 %
	CRC inclus	3	0,0 %	1	0,0 %	16	0,8 %	3	0,1 %	52	0,1 %
CNRS	CRC exclus	3 486	51,0 %	1 236	13,9 %	800	41,0 %	4 038	79,0 %	17 838	52,2 %
	CRC inclus	3 647	50,7 %	1 359	13,6 %	867	41,1 %	4 839	78,6 %	20 698	52,6 %
Inra	CRC exclus	867	12,7 %	325	3,7 %	796	40,8 %	117	2,3 %	2 268	6,6 %
	CRC inclus	907	12,6 %	348	3,5 %	844	40,0 %	135	2,2 %	2 409	6,1 %
Inserm	CRC exclus	1 796	26,3 %	2 155	24,3 %	48	2,4 %	62	1,2 %	4 179	12,2 %
	CRC inclus	1 885	26,2 %	2 438	24,4 %	63	3,0 %	72	1,2 %	4 582	11,7 %
IRD	CRC exclus	55	0,8 %	35	0,4 %	82	4,2 %	4	0,1 %	332	1,0 %
	CRC inclus	57	0,8 %	37	0,4 %	94	4,4 %	4	0,1 %	354	0,9 %
Université	CRC exclus	4 507	65,9 %	7 123	80,2 %	1 035	53,0 %	3 947	77,3 %	24 365	71,2 %
	CRC inclus	4 740	65,9 %	8 017	80,4 %	1 127	53,4 %	4 627	75,2 %	27 660	70,3 %
Autres FRA	CRC exclus	534	7,8 %	875	9,9 %	151	7,7 %	297	5,8 %	3 072	9,0 %
	CRC inclus	567	7,9 %	1 003	10,1 %	163	7,7 %	389	6,3 %	3 751	9,5 %
FRA	CRC exclus	6 839	100,0 %	8 880	100,0 %	1 951	100,0 %	5 108	100,0 %	34 197	100,0 %
	CRC inclus	7 190	100,0 %	9 974	100,0 %	2 109	100,0 %	6 154	100,0 %	39 329	100,0 %

Attention aux arrondis, les totaux peuvent différer légèrement de la somme en ligne

Attention : pas d'additivité en colonnes (comptage fractionnaire)

Tableau 2 (suite)

		Physique		Sciences de l'Univers		Sciences pour l'Ingénieur		Autres*		Toutes disc. confondues	
CEA	CRC exclus	746	14,8 %	202	7,2 %	107	6,1 %	56	3,1 %	1 803	5,3 %
	CRC inclus	1 122	16,7 %	215	7,0 %	185	8,2 %	59	3,1 %	2 414	6,1 %
Cemagref	CRC exclus	0	0,0 %	15	0,5 %	11	0,6 %	1	0,1 %	45	0,1 %
	CRC inclus	0	0,0 %	17	0,6 %	11	0,5 %	1	0,1 %	52	0,1 %
CNRS	CRC exclus	3 915	77,9 %	2 204	78,2 %	1 018	58,4 %	1 141	62,3 %	17 838	52,2 %
	CRC inclus	5 140	76,7 %	2 378	77,7 %	1 295	57,5 %	1 174	62,2 %	20 698	52,6 %
Inra	CRC exclus	18	0,3 %	83	3,0 %	31	1,8 %	32	1,7 %	2 268	6,6 %
	CRC inclus	21	0,3 %	88	2,9 %	33	1,5 %	33	1,7 %	2 409	6,1 %
Inserm	CRC exclus	12	0,2 %	4	0,1 %	9	0,5 %	93	5,1 %	4 179	12,2 %
	CRC inclus	14	0,2 %	5	0,2 %	11	0,5 %	94	5,0 %	4 582	11,7 %
IRD	CRC exclus	1	0,0 %	134	4,8 %	14	0,8 %	8	0,4 %	332	1,0 %
	CRC inclus	2	0,0 %	139	4,6 %	14	0,6 %	8	0,4 %	354	0,9 %
Université	CRC exclus	3 544	70,5 %	1 739	61,7 %	1 029	59,1 %	1 442	78,7 %	24 365	71,2 %
	CRC inclus	4 485	66,9 %	1 901	62,1 %	1 278	56,7 %	1 485	78,7 %	27 660	70,3 %
Autres FRA	CRC exclus	431	8,6 %	257	9,1 %	372	21,3 %	156	8,5 %	3 072	9,0 %
	CRC inclus	664	9,9 %	299	9,8 %	502	22,3 %	164	8,7 %	3 751	9,5 %
FRA	CRC exclus	5 026	100,0 %	2 818	100,0 %	1 742	100,0 %	1 832	100,0 %	34 197	100,0 %
	CRC inclus	6 700	100,0 %	3 061	100,0 %	2 254	100,0 %	1 887	100,0 %	39 329	100,0 %

*Mathématiques et journaux multidisciplinaires

Attention aux arrondis, les totaux peuvent différer légèrement de la somme en ligne

Attention : pas d'additivité en colonnes (comptage fractionnaire)

Hors CRC, les universités signent ou co-signent plus de 70 % des articles français, le CNRS plus de 50 %. Les universités et le CNRS, tous deux à vocation multidisciplinaire et largement associés institutionnellement, sont naturellement présents partout, avec quelques nuances : le CNRS est par exemple, moins actif en recherche médicale, les universités en biologie appliquée.

Le CEA, autour de sa discipline dominante, la physique, a une activité importante en sciences de l'univers, et aussi en chimie, sciences de l'ingénieur, biologie fondamentale.

Le Cemagref apparaît en biologie appliquée, sciences de l'univers, sciences pour l'ingénieur.

Les autres instituts spécialisés ont chacun deux disciplines majeures : biologie appliquée et sciences de l'univers pour l'IRD-France, biologie appliquée et biologie fondamentale pour l'Inra, biologie fondamentale et recherche médicale pour l'Inserm.

L'inclusion des CRC modifie principalement la position du CEA en physique, chimie et sciences pour l'ingénieur. Il passe globalement de 5,3 à 6,1 % des publications françaises.

Tableau 3 : Profils disciplinaires des institutions (%)

		Biologie Fondamentale	Recherche Médicale	Biologie Appliquée - Ecologie	Chimie	Physique	Sciences de l'Univers	Sciences pour l'Ingénieur	Autres*	Toutes disc. confondues
CEA	CRC exclus	16,2	5,7	2,1	14,4	41,4	11,2	5,9	3,1	100,0
	CRC inclus	12,9	4,5	1,6	15,6	46,5	8,9	7,7	2,4	100,0
Cemagref	CRC exclus	6,0	1,3	30,4	3,0	0,0	33,7	23,3	2,2	100,0
	CRC inclus	5,2	1,2	31,2	6,4	0,0	33,3	20,8	1,9	100,0
CNRS	CRC exclus	19,5	6,9	4,5	22,6	21,9	12,4	5,7	6,4	100,0
	CRC inclus	17,6	6,6	4,2	23,4	24,8	11,5	6,3	5,7	100,0
Inra	CRC exclus	38,2	14,3	35,1	5,2	0,8	3,7	1,4	1,4	100,0
	CRC inclus	37,7	14,4	35,0	5,6	0,9	3,7	1,4	1,4	100,0
Inserm	CRC exclus	43,0	51,6	1,1	1,5	0,3	0,1	0,2	2,2	100,0
	CRC inclus	41,1	53,2	1,4	1,6	0,3	0,1	0,2	2,1	100,0
IRD-France	CRC exclus	16,7	10,5	24,7	1,1	0,3	40,4	4,1	2,3	100,0
	CRC inclus	16,1	10,5	26,4	1,0	0,6	39,4	3,9	2,1	100,0
Universités	CRC exclus	18,5	29,2	4,2	16,2	14,5	7,1	4,2	5,9	100,0
	CRC inclus	17,1	29,0	4,1	16,7	16,2	6,9	4,6	5,4	100,0
Autres FRA	CRC exclus	17,4	28,5	4,9	9,7	14,0	8,4	12,1	5,1	100,0
	CRC inclus	15,1	26,7	4,4	10,4	17,7	8,0	13,4	4,4	100,0
France	CRC exclus	20,0	26,0	5,7	14,9	14,7	8,2	5,1	5,4	100,0
	CRC inclus	18,3	25,4	5,4	15,6	17,0	7,8	5,7	4,8	100,0
Monde	CRC exclus	18,8	30,5	7,3	14,6	12,2	6,2	6,5	4,0	100,0
	CRC inclus	17,7	30,2	6,9	14,9	13,5	6,1	7,1	3,7	100,0

* Mathématiques et journaux multidisciplinaires

année 1997

Le profil disciplinaire des institutions présenté dans le tableau 3 permet de retrouver les points forts attendus (biologie et médecine pour l'Inserm, physique pour le CEA, etc.). Cependant, les disciplines ayant elles-mêmes des poids très inégaux à l'échelle française, une vue complémentaire est obtenue en comparant pour chaque institution son poids dans la France par discipline à son poids toutes disciplines confondues. C'est aussi le ratio de la part d'activité que l'institution consacre à une discipline à la part d'activité de la France dans cette discipline. Ce ratio est "l'indice d'activité" ou de spécialisation, qui exprime l'engagement relatif dans les divers secteurs de connaissance. Il est ici normalisé entre -1 et +1 (tableau 4) :

Tableau 4 : Indices de spécialisation par rapport à la France (CRC exclus)

		Biologie Fondamentale	Recherche Médicale	Biologie Appliquée - Ecologie	Chimie	Physique	Sciences de l'Univers	Sciences pour l'Ingénieur	Autres*
CEA		-0,21	-0,91	-0,77	-0,04	0,78	0,30	0,15	-0,50
Cemagref		-0,83	-0,99	0,93	-0,92	-1,00	0,89	0,91	-0,71
CNRS		-0,02	-0,87	-0,24	0,39	0,38	0,38	0,11	0,18
Inra		0,57	-0,53	0,95	-0,79	-0,99	-0,67	-0,87	-0,87
Inserm		0,64	0,60	-0,92	-0,98	-1,00	-1,00	-1,00	-0,70
IRD-France		-0,18	-0,72	0,90	-0,99	-1,00	0,92	-0,22	-0,70
Universités		-0,08	0,12	-0,29	0,08	-0,01	-0,14	-0,19	0,10
Autres FRA		-0,14	0,09	-0,15	-0,41	-0,05	0,02	0,70	-0,05

* Mathématiques et journaux multidisciplinaires

année 1997

en italique : chiffres correspondant à moins de 50 publications

Si 'a' désigne le niveau de spécialisation, la normalisation entre -1 et 1 consiste à faire la transformation $(a^2-1)/(a^2+1)$

Dans ce tableau, les spécialisations oscillent entre -1 (sous-spécialisation maximale) et +1 (spécialisation maximale). Les données "toutes disciplines confondues" et "France entière" sont la référence, et seraient donc égales à zéro. Les chiffres en italiques correspondent à moins de 50 publications. On sait qu'en ce cas les chiffres obtenus sont très sensibles à l'activité de quelques laboratoires (c'est le cas, par exemple, pour le Cemagref).

A nouveau, les engagements forts de chaque institution sont bien visibles : la physique, les sciences pour l'ingénieur et les sciences de l'univers pour le CEA, la biologie fondamentale et appliquée pour l'Inra, la biologie fondamentale et la recherche médicale pour l'Inserm, la biologie appliquée et les sciences de l'univers pour l'IRD-France.

Les universités et le CNRS, de profils logiquement assez proches, sont relativement moins engagés dans ces disciplines couvertes par les instituts spécialisés. Leurs domaines privilégiés restent les sciences physiques (physique, chimie, sciences de l'univers).

II - COOPERATIONS SCIENTIFIQUES

L'activité scientifique est par nature coopérative, que ce soit à l'intérieur des équipes de recherche ou entre équipes de recherche. On réserve normalement le terme de "co-publications" aux publications résultant de collaborations entre laboratoires différents ; on a vu que d'autres formes de coopération plus institutionnalisées, comme les unités mixtes, existent : on parlerait plutôt, en ce cas, de "co-présence institutionnelle".

En cohérence avec le compte de présence, on définit ici le lien entre deux acteurs par le nombre de leurs "co-présences". Du fait de la méthode employée pour le repérage, ces co-présences incluent, de manière indifférenciée, les co-présences institutionnelles et les collaborations.

Co-présences entre institutions

Tableau 5 : Co-présences entre les institutions participantes (nombres de publications, CRC exclus)

co-présences	CEA	Cemagref	CNRS	Inra	Inserm	IRD-France	Universités	Total coop- 7***	Rappel total France
CEA		-	1 063	34	98	10	777	1 203	1 803
Cemagref	-		12	9	-	4	10	25	45
CNRS	1 063	12		575	1 107	152	13 919	14 692	17 838
Inra	34	9	575		145	16	908	1 115	2 268
Inserm	98	-	1 107	145		13	3 058	3 332	4 179
IRD-France	10	4	152	16	13		147	205	332
Universités	777	10	13 919	908	3 058	147		16 584	24 365

Co-présences non additives (compte de présence)

année 1997

*(***) : le total des partenariats français ne couvre que les partenaires de la coopérative actuellement renseignés*

Dans ce tableau, et pour chaque institution (lecture en ligne), le nombre total d'articles signés en partenariat avec l'un des autres partenaires de l'étude et le nombre total d'articles de l'institution sont rappelés.

Quantitativement, et sans revenir sur le recouvrement institutionnel massif du CNRS et des universités, les partenariats les plus importants - qu'il s'agisse d'unités mixtes ou de co-publications - sont entre les universités et l'Inserm, puis entre le CNRS et l'Inserm, le CNRS et le CEA, l'Inra et les universités (tableau 5).

Ces chiffres bruts peuvent être ramenés, pour chacune des sept institutions, à l'ensemble des partenariats qu'elle entretient avec les six autres :

Tableau 6 : Co-présences entre institutions, en % des partenariats de chaque institution (CRC exclus)

co-présences (%)	CEA	Cemagref	CNRS	Inra	Inserm	IRD-France	Universités	Rappel du total coop-7***
CEA		-	88,4	2,8	8,1	0,8	64,6	1 203
Cemagref	-		48,0	36,0	-	16,0	40,0	25
CNRS	7,2	0,1		3,9	7,5	1,0	94,7	14 692
Inra	3,0	0,8	51,6		13,0	1,4	81,4	1 115
Inserm	2,9	-	33,2	4,4		0,4	91,8	3 332
IRD-France	4,9	2,0	74,1	7,8	6,3		71,7	205
Universités	4,7	0,1	83,9	5,5	18,4	0,9		16 584

Pourcentages non additifs (compte de présence)

année 1997

****La référence (total des partenariats français) ne couvre que les partenaires de la coopérative actuellement renseignés.*

Lecture : dans la ligne CEA, le chiffre de la colonne CNRS indique que pour 100 articles en partenariat français avec adresse CEA, 88,4 % ont une co-signature CNRS

Le tableau 6 montre par exemple que les universités sont les partenaires d'élection pour l'Inserm et l'Inra, que le CNRS collabore largement avec le CEA, l'IRD et le Cemagref.

Collaborations européennes et internationales

Tableau 7 : Taux de collaboration (%) européenne et internationale (CRC exclus)

	Collaborations européennes	Collaborations internationales	Rappel total France
CEA	29,7	49,3	1 803
Cemagref	20,0	37,8	45
CNRS	20,5	41,6	17 838
Inra	15,8	31,3	2 268
Inserm	17,2	33,1	4 179
IRD-France	15,7	61,7	332
Universités	16,7	34,1	24 365

Pourcentages non additifs (compte de présence) année 1997

Le tableau 7, dont la dernière colonne est un simple rappel du nombre total de publications enregistrées pour l'institution considérée, montre que le taux de collaboration internationale est très élevé structurellement pour l'IRD-France (62 %), assez élevé pour le CEA (49 %), et le CNRS (42 %), et moins élevé pour les universités, l'Inserm et l'Inra (moins de 35 %). Pour mémoire, le taux de collaboration pour la France toute entière est de 38,4 %. Par ailleurs l'IRD-France et l'Inra enregistrent fréquemment des partenariats étrangers sans autre présence française, ce qui reflète leur faible recouvrement avec les autres institutions.

Le taux de collaboration européenne, en regard de la collaboration internationale totale, est particulièrement élevé pour le CEA et faible pour l'IRD. Par ailleurs, le taux de coopérations multipartenaires (nombre moyen de pays partenaires sur les articles en co-publication internationale) est élevé pour le CEA et à un moindre degré l'IRD, et faible pour l'Inra.

Globalement, les principaux partenaires de la France sont dans l'ordre les USA, l'Allemagne, le Royaume-Uni, l'Italie, l'Espagne, la Suisse, le Canada, la Belgique, les Pays-Bas et la Russie (voir tableau 8), les institutions ayant des profils assez contrastés :

Le CEA privilégie logiquement les grands pays physiques : USA, Allemagne, Royaume-Uni, qui sont aussi les premiers partenaires scientifiques de la France en termes quantitatifs, ainsi que la Suisse (notamment le Cern), l'Italie et le Japon.

Le CNRS, proche de la moyenne générale française - qu'il conditionne en partie -, ne fait apparaître que de légères préférences relatives, au profit de l'Allemagne et de l'Espagne.

L'université dans son ensemble a naturellement un profil assez proche du CNRS, avec peu d'écarts à la moyenne, et une légère préférence pour l'Italie, l'Espagne, la Suisse, la Belgique et le Canada et parmi les petits pays, pour le Maroc.

L'INRA montre une préférence pour le Royaume-Uni, pays actif en sciences de la vie, pour l'Espagne, le Canada, la Belgique, et l'Australie, puis pour le Danemark, Israël et la Chine.

L'Inserm a comme partenaires privilégiés les pays de pointe en recherche médicale et biologique, Etats-Unis et le Royaume-Uni, et également l'Italie, l'Allemagne, la Suisse, le Canada, la Belgique, puis avec des partenariats plus modestes la Suède, le Danemark et Israël.

L'IRD-France a, comme on peut s'y attendre, une configuration atypique et la liste de ses premiers partenaires, en dehors des Etats-Unis et du Royaume-Uni, est très différente de la moyenne française. Les principaux partenaires sont en Afrique de l'Ouest, puis en Amérique du Sud. On note aussi une coopération proportionnellement importante avec l'Australie.

Tableau 8 : Co-publications internationales (%) par ordre des partenaires de la France (CRC exclus)

	CEA	Cemagref	CNRS	Inra	Inserm	IRD-France	Universités	*France
USA	36,8	23,5	24,5	21,2	32,7	20,5	25,3	26,0
DEU	24,5	11,8	14,7	10,4	11,3	1,5	13,1	14,1
GBR	15,6	17,6	11,8	14,0	17,5	13,7	12,6	13,4
ITA	15,9	0,0	9,7	5,4	12,7	4,4	10,3	10,0
ESP	9,8	5,9	7,8	9,2	4,1	2,0	7,6	7,4
CHE	10,8	0,0	6,3	3,5	7,5	1,0	6,6	6,5
CAN	4,0	5,9	4,9	6,9	6,9	2,0	5,9	5,8
BEL	7,0	0,0	4,8	8,5	7,4	3,4	6,2	5,6
NLD	6,2	23,5	4,3	4,2	5,1	5,4	4,9	4,9
CEI	8,2	0,0	6,3	1,4	2,1	1,0	5,0	4,8
JPN	4,4	5,9	3,8	1,8	4,5	2,4	3,4	3,5
SWE	3,7	5,9	2,5	2,4	3,7	0,5	2,8	2,9
POL	5,8	0,0	3,4	1,3	0,7	0,0	2,8	2,5
AUS	1,6	0,0	2,2	4,5	1,6	6,8	2,1	2,3
BRA	4,0	0,0	2,7	2,1	1,7	6,8	2,3	2,3
MAR	0,7	0,0	2,7	1,3	0,8	0,5	3,0	2,2
DNK	6,1	0,0	1,8	2,0	2,5	2,9	1,8	2,1
ISR	1,6	0,0	1,7	2,0	2,1	0,0	1,9	1,7
GRC	7,0	17,6	2,0	1,4	1,2	0,0	1,8	1,7
CHN	3,4	0,0	1,8	1,6	1,2	0,0	1,8	1,6
...								
Total International*	889	17	7 427	709	1 384	205	8 305	13 124

% des publications étrangères de chaque institution

année 1997

*nombre d'articles en partenariat international

Les chiffres en gras sont supérieurs à la moyenne française

En italique : chiffres peu significatifs, correspondant à moins de 10 articles

Si, au lieu de considérer les profils de chaque institution par rapport à la moyenne française, on normalisait par la capacité d'échange de chaque partenaire, on constaterait la relativement faible affinité française pour les plus grandes puissances scientifiques (USA, Royaume-Uni, Japon, et même l'Allemagne). Les relations les plus étroites, lorsque l'on corrige l'effet de taille, sont celles avec les voisins francophones ou méditerranéens : la Belgique, l'Espagne, l'Italie, la Suisse, la Grèce, et naturellement avec les partenaires traditionnels d'Afrique, mais cela ne correspond qu'à des coopérations très faibles dans l'absolu.

III. INDICATEURS DE VISIBILITE : CITATIONS ET IMPACTS

Les articles scientifiques comportent presque toujours une bibliographie, les "*articles cités*". Le nombre de références faites à un article ("*citations reçues*" par cet article) est, ainsi, une mesure de sa visibilité - on parle également de son "*impact*".

Les citations dépendent de nombreux facteurs, parmi lesquels le journal et la langue de publication, les habitudes de citations de la discipline, l'orientation de l'article : fondamentale, appliquée, méthodologique, etc. L'interprétation des citations en simples termes de "qualité" est incorrecte (par exemple, la version anglaise aura au moins cinq fois plus de chances d'être citée que la version française du même article), mais il est incontestable qu'elles mesurent la visibilité et l'influence d'un article.

La complexité des mécanismes à l'œuvre rend parfois l'interprétation délicate. Les citations sont par exemple sujettes à d'importants effets d'amplification en faveur des articles les plus visibles. Le nombre moyen de citations recueillies par un article étant très différent en biologie fondamentale et en mathématiques par exemple, il est important de normaliser les citations, au moins au niveau des disciplines, mais plusieurs niveaux de normalisation sont possibles, comme on le verra plus loin avec le "*ratio de citation relative*".

Par souci de cohérence entre indicateurs de publication et de citation, nous appliquons le même principe de comptage : les citations reçues par chaque article sont intégralement attribuées à chaque institution signataire. Dans cette étude, on décompte les citations reçues sur une fenêtre de trois ans, année de publication comprise - c'est-à-dire les citations émises dans les années 1997, 1998, 1999.

Le groupe de travail a décidé de ne présenter que la visibilité hors CRC. Les CRC sont en général moins cités que les articles ordinaires, mais là encore la situation varie selon les disciplines.

"Parts de marché" des citations

Dans le tableau 9, la part de citations des différentes institutions est calculée par le rapport entre le nombre des citations reçues par les articles où une institution apparaît, et le nombre de citations reçues par tous les articles avec adresse française (ensemble des citations de toutes les institutions françaises). L'indicateur n'est pas additif entre institutions en raison des doubles comptes (unités mixtes et co-publications).

Tableau 9 : Parts de citations, en % de la France (CRC exclus)

	Biologie Fondamentale	Recherche Médicale	Biologie Appliquée - Ecologie	Chimie	Physique	Sciences de l'Univers	Sciences pour l'Ingénieur	Autres*	Toutes disc. confondues
CEA	4,4	1,4	2,6	6,6	20,1	12,2	7,5	8,6	6,9
Cemagref	0,0	0,0	0,5	0,0	0,0	0,2	0,3	0,0	0,0
CNRS	58,9	21,2	52,2	83,3	79,8	84,2	63,8	69,3	57,2
Inra	9,2	2,5	41,7	2,0	0,3	1,9	1,8	7,5	5,8
Inserm	32,5	40,0	4,0	1,4	0,1	0,0	0,4	25,5	21,9
IRD-France	0,6	0,4	3,1	0,1	0,0	3,5	0,5	0,6	0,7
Universités	60,6	72,6	52,9	76,7	69,8	59,0	61,7	60,7	66,4
Autres FRA	7,5	11,1	6,0	5,6	8,8	7,3	18,8	10,6	8,7

* *Mathématiques et journaux multidisciplinaires*

année 1997

En italique : chiffres correspondant à moins de 50 publications

Attention : pas d'additivité en colonnes (comptages multiples)

Impact

En bibliométrie, "l'impact" désigne le nombre moyen de citations par article. On appellera impact d'une institution le nombre moyen de citations par article pour un acteur donné (ici une des sept institutions), sur un mode de comptage et des fenêtres temporelles définies⁶.

Le tableau 10 présente les impacts relatifs (par rapport au monde et hors CRC) des différentes institutions :

Tableau 10 : Impact relatif par rapport au monde (CRC exclus)

	Biologie Fondamentale	Recherche Médicale	Biologie Appliquée - Ecologie	Chimie	Physique	Sciences de l'Univers	Sciences pour l'Ingénieur	Autres*	Toutes disc. confondues
CEA	1,03	1,33	1,68	1,36	1,59	1,97	1,48	ns	1,42
Cemagref	<i>ns</i>	<i>ns</i>	0,85	<i>ns</i>	-	0,43	0,69	<i>ns</i>	0,36
CNRS	1,16	1,73	1,57	1,11	1,20	1,25	1,32	<i>ns</i>	1,19
Inra	0,73	0,77	1,26	0,91	0,96	0,75	1,25	<i>ns</i>	0,95
Inserm	1,25	1,88	2,05	1,23	0,61	<i>ns</i>	1,01	<i>ns</i>	1,94
IRD-France	0,74	1,19	0,91	<i>ns</i>	<i>ns</i>	0,85	0,82	<i>ns</i>	0,80
Universités	0,92	1,03	1,23	1,05	1,16	1,11	1,26	<i>ns</i>	1,01
Autres FRA	0,96	1,28	0,96	1,01	1,20	0,93	1,06	<i>ns</i>	1,05
France	1,01	1,14	1,24	1,06	1,17	1,16	1,21	<i>ns</i>	1,08
Monde	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

(*) *Mathématiques et journaux multidisciplinaires, l'impact composite n'a pas de signification.*

année 1997

En italique : chiffres peu significatifs, correspondant à moins de 50 publications

ns : pour la colonne "Autres", ou lorsqu'il y a moins de 8 publications

On voit dans ce tableau que la visibilité des universités est peu différente de la moyenne française dans la plupart des disciplines, un peu au-dessus pour les sciences pour l'ingénieur et la biologie appliquée - écologie, un peu en dessous pour la biologie fondamentale et la recherche médicale.

La visibilité du CEA est particulièrement forte en sciences de l'univers, et très bonne dans ses domaines de spécialisation : physique, sciences pour l'ingénieur, chimie et également, en dehors de ses spécialisations, en recherche médicale et biologie appliquée.

De même, la visibilité de l'Inserm est très forte dans ses spécialités, recherche médicale et biologie fondamentale, un domaine où la concurrence - américaine notamment - est particulièrement vive.

Le CNRS bénéficie d'une visibilité forte dans ses disciplines de spécialisation (physique, chimie, sciences de l'univers), et plus forte encore sur des domaines de moindre spécialisation (recherche médicale, biologie appliquée).

La visibilité de l'Inra est élevée en biologie appliquée-écologie, et plus modeste en biologie fondamentale.

L'IRD-France, dont la mission est plus centrée sur le développement, est en retrait de la moyenne pour ce critère d'impact, qui est typique de la recherche académique.

Le faible nombre de publications du Cemagref ne permet pas d'interpréter les résultats d'impact.

⁶Notons que le fait de publier des articles peu visibles, par exemple en raison du choix du journal ou du thème de recherche, aura évidemment des conséquences très différentes sur l'indicateur de "part de marché des citations" et sur l'indicateur d'impact d'une institution donnée : par exemple, le fait d'ajouter un article de faible visibilité à un portefeuille de publications laisse inchangée la "part de marché de citations", mais tend à abaisser l'indicateur d'impact.

Il faut noter que le mode de comptage utilisé affecte fortement le niveau des indicateurs d'impact. Ainsi, l'impact relatif de la France par rapport au monde apparaît comme nettement plus élevé en compte de présence qu'en compte fractionnaire⁷.

Enfin, comme on l'a déjà mentionné, l'interprétation des chiffres des universités est délicate, car les laboratoires y sont nettement plus hétérogènes que dans les autres institutions. Pour s'en convaincre, il suffit d'observer deux exemples contrastés, l'un correspondant à une université de grande taille que nous appellerons A, bien insérée dans le courant international de la recherche, et l'autre une université de taille modeste B, avec un poids important des activités d'enseignement, dont les publications sont peu nombreuses et plus orientées vers la recherche appliquée, c'est-à-dire se faisant dans des journaux de moindre visibilité internationale ou des journaux de transfert. Les différences d'impact sont très fortes : le chiffre global des universités est une moyenne cachant des variations considérables d'impact.

Tableau 11 : Parts de publications et impact relatif pour deux universités à vocations différentes

Universités A et B	Biologie		Biologie Appliquée			Sciences de l'Univers	Sciences pour l'Ingénieur		Toutes disc. confondues
	Fondamentale	Recherche Médicale	- Ecologie	Chimie	Physique		Autres*		
A parts de publications**	11,98	11,88	4,56	5,88	13,89	27,87	8,97	20,07	11,86
A impact relatif	1,20	1,47	1,03	0,75	1,09	1,20	2,03	0,39	1,22
B parts de publications**	<i>0,29</i>	<i>0,35</i>	-	1,90	1,12	<i>0,79</i>	-	-	0,66
B impact relatif	<i>0,64</i>	<i>0,43</i>	-	0,95	0,32	<i>0,62</i>	-	-	0,56

* *Mathématiques et journaux multidisciplinaires*

année 1997

** *parts mondiales en dix millièmes*

En italique : chiffres peu significatifs, correspondant à moins de 7 publications ; tirets : moins de 2

Ratio de citation relative

Pour des raisons tenant aux différences de comportement de citation entre disciplines, le chiffre agrégé des "impacts relatifs" pour un acteur à un niveau donné (par exemple "toutes disciplines"), est sensible à la distribution des articles entre disciplines, par exemple à une forte présence dans les disciplines à fort taux de citation comme la biologie. Des normalisations à divers niveaux (journaux, spécialités, sous-disciplines, disciplines,...) sont donc utiles. Nous nous limiterons ici à l'indicateur le plus finement normalisé, le "ratio de citation relative", en considérant chaque journal comme une micro-spécialité scientifique.

Dans cette optique, les citations reçues par un acteur pour un article sont considérées comme résultant de deux composantes : d'une part, le nombre de citations attendu du journal où l'article est publié ("l'impact" du journal)⁸, d'autre part le nombre relatif de citations reçues effectivement par l'article par

⁷Science et Technologie, Indicateurs Edition 2002, Rapport de l'Observatoire des Sciences et des Techniques (OST), Economica, 2002.

A titre d'exemple, sur le périmètre de journaux retenu ici, l'impact relatif France serait nettement au-dessous de l'unité, à 0,96 pour "toutes disciplines". La différence est très accusée pour certaines disciplines dans lesquelles les collaborations internationales sont nombreuses, comme la recherche médicale et les sciences de l'univers. En effet, les articles en co-publication internationale sont en général nettement plus cités que les articles seulement français et, en compte de présence, la totalité du bénéfice de l'article en co-publication est attribuée à la France, alors qu'un comptage fractionnaire aurait pondéré ce bénéfice au prorata de la contribution française.

⁸L'impact de chaque journal, qui apprécie son audience et sa visibilité, est calculé par l'OST à partir des citations attribuées aux articles. L'impact d'un journal pour 1997 est ici défini comme le ratio du nombre de citations reçues par les articles (publiés en 1997 dans le journal), dans la période 1997, 1998, 1999 (donc sur 3 ans), au nombre d'articles du journal en 1997 (citation moyenne par article). Le nombre moyen de citations obtenues par les articles du journal (son impact) rappelle l'"impact factor" des journaux calculés à deux ans de l'ISI, mais le mode de calcul est différent (types de documents pris en compte, calcul au niveau document par "année citée", choix des fenêtres temporelles).

rapport à ce nombre de citations attendu. L'agrégation pour tous les articles d'un acteur définit respectivement l'impact espéré - "l'effet journal" - et le "ratio de citation relative" (ou RCR), qui dénote la sur- ou sous-visibilité de cet acteur dans les journaux considérés. Ceci traduit deux formes de concurrence : d'abord pour accéder à la publication dans une revue de fort impact ("est-on présent dans les journaux de bonne visibilité ?"), ensuite pour être visible et apprécié au sein de la communauté de recherche (supposée reflétée par un journal) à laquelle l'auteur s'adresse ("est-on bien ou moins bien visible dans les journaux où on est présent ?"). La "stratégie de publication" consiste notamment à se positionner dans ces deux formes de concurrence. Il ne faut pas oublier, à cet égard, que les articles de recherche fondamentale ont relativement plus de chances d'accéder à des revues de fort impact que les articles à caractère plus appliqué par exemple⁹. De même, le ratio de citation relative dépend du type de document concerné (article, note, synthèse, etc.) et une institution qui publie beaucoup d'articles de synthèse, par exemple, verra ici son ratio de citation relative augmenter.

Tableau 12 : Ratios de citation relative (CRC exclus)

	Biologie		Biologie Appliquée -			Sciences		Sciences pour l'Ingénieur		Toutes disc. confondues
	Fondamentale	Recherche Médicale	Ecologie	Chimie	Physique	de l'Univers	Autres*			
CEA	0,86	1,02	0,95	1,07	1,25	1,40	1,20	ns	1,12	
Cemagref	<i>0,70</i>	<i>0,00</i>	<i>1,23</i>	<i>0,64</i>		<i>0,75</i>	<i>0,64</i>	<i>ns</i>	<i>0,81</i>	
CNRS	1,00	1,15	1,14	1,01	1,07	1,18	1,10	ns	1,05	
Inra	0,94	0,96	1,17	1,11	<i>1,01</i>	1,23	<i>0,95</i>	<i>ns</i>	1,05	
Inserm	1,03	1,20	<i>1,14</i>	1,02	<i>0,83</i>	<i>0,36</i>	<i>0,62</i>	ns	1,13	
IRD-France	1,23	<i>1,41</i>	1,00	<i>0,75</i>	<i>1,81</i>	1,13	<i>0,84</i>	<i>ns</i>	1,02	
Universités	0,95	1,13	1,04	0,99	1,06	1,17	1,05	ns	1,04	
Autres FRA	1,00	1,18	1,11	1,04	1,14	1,08	1,03	ns	1,06	
France	0,99	1,14	1,10	1,00	1,07	1,15	1,05	ns	1,05	
Monde	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	

(*) *Mathématiques- et journaux multidisciplinaires ; le RCR composite n'a pas de signification*

année 1997

En italique : chiffres peu significatifs, correspondant à moins de 50 publications

ns : pour la colonne composite "Autres"

Le tableau 12 présente les ratios de citation relative (normalisés par rapport au monde) des différentes institutions, pour les différentes disciplines et toutes disciplines confondues, et il permet de préciser certaines composantes de la visibilité acquise par les institutions.

La visibilité du CEA en sciences de l'univers, physique et sciences pour l'ingénieur passe à la fois par une capacité à investir des journaux de bonne visibilité et par une bonne performance relative dans ces journaux.

Par contraste, en biologie fondamentale, le bon impact du CNRS est entièrement attribuable à la visibilité des journaux où il publie, le ratio de citation relative étant juste à la moyenne.

En biologie appliquée, le ratio de citation relative du Cemagref ou de l'Inra est du même ordre de grandeur que celui de l'Inserm ou du CNRS, la différence en impact réel provenant essentiellement d'une différence dans l'accès aux journaux les plus visibles.

⁹On pourra ainsi observer, dans le tableau 13 donné en annexe, que la capacité de l'Inserm, du CEA et du CNRS à accéder aux journaux de bonne visibilité est plus forte que celle d'institutions de recherche à vocation appliquée, qui ont davantage de difficultés à publier dans des journaux d'impact élevé, qui sont souvent des journaux à spectre large publiant des articles de recherche plutôt fondamentale. En biologie par exemple, le type de modèle étudié (humain, animal, végétal) peut largement conditionner l'accès à des revues de niveaux d'impact assez différents.

CONCLUSIONS

Ce premier exercice, décidé entre les partenaires pour tester la faisabilité et l'intérêt de travailler à des indicateurs bibliométriques déclinés par institution, nous semble conclusif. Il a montré qu'une collaboration active entre les représentants des institutions partenaires est efficace, et devrait permettre une appropriation collective de la démarche et des résultats par chacune d'entre elles. Les résultats originaux obtenus ici avec un jeu assez simple d'indicateurs sont stimulants, et la poursuite de l'exercice prometteuse. Elle devra s'appuyer sur des données de départ plus précises (repérages institutionnels au niveau des adresses et non plus du seul article, donnant accès notamment aux analyses territoriales) et plus récentes, couvrant un périmètre de journaux élargi. La mise en œuvre d'une batterie plus complète d'indicateurs sera utile. Enfin, d'autres institutions de recherche seront appelées à participer aux travaux. Ceci permettra de décrire le paysage des institutions de recherche plus complètement et de mieux cerner ses évolutions.

A - Impacts relatifs espérés

L'*impact espéré relatif* est le ratio de la part mondiale de l'acteur dans les citations espérées compte tenu de l'impact des journaux où il publie, à la part mondiale de l'acteur dans les publications (ratio calculé dans la discipline considérée). L'impact global peut être vu comme le produit de l'impact espéré et du ratio de citation relative (RCR) commenté. Le RCR est souvent différent de l'unité, ce qui montre que l'approximation souvent faite de la visibilité des recherches par le seul "impact espéré" n'est pas suffisante.

Tableau 13 : Impact relatif espéré par rapport au monde (CRC exclus)

	Biologie Fondamentale	Recherche Médicale	Biologie Appliquée - Ecologie	Chimie	Physique	Sciences de l'Univers	Sciences pour l'Ingénieur	Autres*	Toutes disc. confondues
CEA	1,20	1,30	1,78	1,28	1,27	1,41	1,23	ns	1,27
Cemagref	<i>ns</i>	<i>ns</i>	0,69	<i>ns</i>	-	0,57	1,08	<i>ns</i>	0,45
CNRS	1,16	1,50	1,37	1,10	1,12	1,06	1,20	<i>ns</i>	1,13
Inra	0,77	0,80	1,08	0,82	0,96	0,61	1,32	<i>ns</i>	0,90
Inserm	1,21	1,56	1,79	1,20	0,73	<i>ns</i>	1,65	<i>ns</i>	1,72
IRD-France	0,60	0,84	0,91	<i>ns</i>	<i>ns</i>	0,75	0,98	<i>ns</i>	0,79
Universités	0,98	0,91	1,19	1,06	1,09	0,95	1,21	<i>ns</i>	0,97
Autres FRA	0,96	1,08	0,86	0,97	1,05	0,86	1,04	<i>ns</i>	0,99
France	1,02	1,00	1,12	1,05	1,09	1,00	1,16	<i>ns</i>	1,03
Monde	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

* Mathématiques et journaux multidisciplinaires

année 1997

L'*impact relatif espéré par rapport au monde* est le ratio de la part mondiale de l'acteur en citations espérées, à la part mondiale de l'acteur en publications, ratio calculé dans la discipline considérée

L'*impact relatif espéré* caractérise la visibilité des journaux dans lesquels on publie

En italique : chiffres correspondant à moins de 50 publications

ns pour la colonne composite "Autres", ou lorsqu'il y a moins de 8 publications

On observe ainsi la capacité de l'Inserm, du CEA et du CNRS à accéder aux journaux de bonne visibilité. Les instituts de recherche appliquée ont davantage de difficultés à publier dans des journaux d'impact élevé, souvent des journaux de spectre large.

B - Distribution des citations

La distribution des citations, comme la plupart des distributions bibliométriques, est très asymétrique c'est-à-dire qu'une grande masse d'articles sont peu ou pas cités alors qu'un petit nombre le sont très fortement. Quelques articles (par exemple des articles de synthèse) peuvent fortement influencer les indicateurs agrégés comme les impacts moyens. D'un point de vue descriptif, il est souhaitable de décrire de manière un peu plus détaillée la distribution des articles de chaque institution par niveau d'impact, en regardant par exemple le nombre d'articles de cette institution dans les 5 % d'articles les plus cités au niveau mondial, puis dans les 5 % suivants, puis dans les 10 % suivants, puis ensuite dans des tranches de 20 % (quintiles).

Les tableaux 14a et 14b ci-dessous présentent ces distributions, par discipline, pour les institutions ayant une activité suffisante dans la discipline. Deux indicateurs y figurent :

- la part mondiale (en millième) de l'institution dans la classe d'impact - tableau 14a ;

- "l'indice d'activité" de l'institution dans la classe d'impact (qui rapporte cette part mondiale dans cette classe à la part mondiale toutes classes confondues) - tableau 14b : si par exemple une institution place 10 % de ses articles dans la classe des 5 % d'articles mondiaux les plus visibles, son indice d'activité est égal à 2 dans cette classe.

Il faut bien noter que le premier indicateur, la position par classe, décrit un positionnement brut de l'institution, par exemple l'activité dans une classe d'excellence, indépendamment de son activité dans les autres classes - alors que le second, l'indice d'activité, tient compte de l'activité globale de l'institution. Par exemple, à nombre égal de publications dans la classe "d'excellence" la plus visible, un acteur engagé dans des publications de transfert - qui sont généralement moins visibles - y aura un indice d'activité plus faible qu'un acteur publiant essentiellement dans des revues de recherche fondamentale.

Tableau 14a : Distribution des citations - parts mondiales par classe

<i>part mondiale dans la classe - pour mille*</i>		<i>Classe, de la plus visible (-1) à la moins visible (-7)</i>							
Discipline		-1	-2	-3	-4	-5	-6	-7	Total
		(5 %)	(5 %)	decile (10 %)	quintile (20 %)	quintile (20 %)	quintile (20 %)	quintile (20 %)	
Biologie fondamentale	CEA	1,6	4,8	4,4	3,5	3,0	2,4	1,8	2,8
	CNRS	42,9	42,1	39,3	39,6	38,8	33,3	25,8	34,5
	Inra	3,0	4,9	9,4	10,0	11,0	8,6	8,5	8,9
	Inserm	28,5	26,3	22,2	20,7	17,9	14,9	12,7	17,8
	Universités	41,6	45,5	45,6	49,2	52,1	49,5	45,7	47,7
	Autres FRA	5,2	6,2	6,1	6,5	6,0	5,3	4,7	5,6
	Monde	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0
Recherche Médicale	CNRS	15,0	14,5	14,1	12,0	9,7	7,2	5,1	9,1
	Inra	1,1	1,9	2,0	2,5	2,8	3,0	1,8	2,3
	Inserm	33,0	26,8	23,5	19,0	13,5	10,5	6,7	14,2
	Universités	49,0	43,6	46,3	42,7	41,6	48,3	47,2	45,5
	Autres FRA	8,5	7,9	6,0	6,2	5,5	5,4	5,1	5,8
	Monde	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0
Biologie Appliquée Ecologie	CNRS	46,9	41,6	30,1	28,8	24,5	20,2	16,2	24,0
	Inra	27,1	25,7	30,7	24,3	21,1	20,5	13,2	20,2
	IRD - France	0,5	2,8	2,8	2,7	2,2	2,6	1,8	2,2
	Universités	41,5	39,5	33,2	34,9	33,2	29,2	24,2	30,5
	Autres FRA	3,5	3,2	4,8	3,3	5,4	3,1	4,8	4,1
	Monde	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0
Chimie	CEA	4,3	5,5	4,6	5,3	3,8	2,5	2,3	3,6
	CNRS	55,4	61,4	60,8	63,8	56,5	52,5	39,6	53,1
	Inra	1,5	1,8	2,3	2,6	2,7	2,1	1,4	2,1
	Universités	47,1	55,9	60,3	61,9	56,3	54,6	41,8	53,0
	Autres FRA	5,1	3,9	4,4	4,0	4,9	4,4	4,6	4,5
	Monde	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0
Physique	CEA	24,3	18,8	16,3	15,5	11,8	9,1	5,8	11,5
	CNRS	80,3	78,9	77,1	75,9	65,6	60,0	45,0	62,7
	Universités	67,9	65,5	69,8	67,2	59,8	56,8	41,8	56,9
	Autres FRA	9,8	7,3	9,3	8,3	7,5	5,6	5,3	7,0
	Monde	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0
Sciences de l'Univers	CEA	20,5	11,5	9,4	7,4	5,4	2,2	2,6	5,7
	CNRS	104,9	90,1	78,6	68,5	62,5	55,5	53,0	64,3
	Inra	0,6	4,1	2,0	2,8	3,9	3,3	2,7	3,0
	IRD - France	3,6	2,7	4,7	4,4	3,7	3,9	4,8	4,2
	Universités	67,6	64,4	55,1	53,4	52,2	49,7	44,3	51,3
	Autres FRA	4,8	11,5	8,2	8,7	7,8	7,8	7,9	8,0
	Monde	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0
	Monde	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0
Sciences pour l'Ingénieur	CEA	11,1	7,0	5,0	5,4	4,5		2,6	4,3
	CNRS	55,9	47,8	44,2	41,7	34,7		26,2	34,2
	Universités	49,1	44,3	43,2	41,0	36,4		26,6	34,1
	Autres FRA	12,5	17,4	18,6	9,1	9,2		10,6	10,7
	Monde	1 000,0	1 000,0	1 000,0	1 000,0	1 000,0		1 000,0	1 000,0

Attention : pas d'additivité en colonne

année 1997

*ne sont gardées dans ce tableau que les institutions dont le total de présence dans la discipline dépasse strictement les 2 pour mille par rapport au monde

Tableau 14b : Distribution des citations - indice d'activité

Indice d'activité*		Classes, de la plus visible (-1) à la moins visible (-7)							Total	
Discipline		-1 (5 %)	-2 (5 %)	-3 decile (10%)	-4 quintile (20%)	-5 quintile (20%)	-6 quintile (20%)	-7 quintile (20%)		
Biologie fondamentale	CEA	0,57	1,70	1,55	1,26	1,07	0,86	0,64	1,00	
	CNRS	1,24	1,22	1,14	1,15	1,12	0,96	0,75	1,00	
	Inra	0,34	0,55	1,06	1,12	1,24	0,96	0,96	1,00	
	Inserm	1,60	1,48	1,25	1,16	1,01	0,84	0,72	1,00	
	Universités	0,87	0,95	0,96	1,03	1,09	1,04	0,96	1,00	
	Autres FRA	0,93	1,11	1,09	1,16	1,07	0,95	0,85	1,00	
	Monde	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	
Recherche Médicale	CNRS	1,65	1,60	1,56	1,32	1,07	0,79	0,56	1,00	
	Inra	0,50	0,81	0,86	1,08	1,21	1,30	0,80	1,00	
	Inserm	2,33	1,89	1,66	1,34	0,95	0,74	0,48	1,00	
	Universités	1,08	0,96	1,02	0,94	0,92	1,06	1,04	1,00	
	Autres FRA	1,48	1,37	1,04	1,07	0,96	0,93	0,89	1,00	
	Monde	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	
	Biologie Appliquée Ecologie	CNRS	1,95	1,73	1,25	1,20	1,02	0,84	0,67	1,00
Inra		1,34	1,27	1,52	1,20	1,05	1,01	0,65	1,00	
IRD - France		0,22	1,26	1,27	1,20	0,98	1,16	0,79	1,00	
Universités		1,36	1,29	1,09	1,15	1,09	0,96	0,79	1,00	
Autres FRA		0,84	0,77	1,17	0,81	1,31	0,75	1,16	1,00	
Monde		1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	
Chimie		CEA	1,18	1,52	1,26	1,47	1,05	0,68	0,64	1,00
	CNRS	1,04	1,16	1,15	1,20	1,06	0,99	0,75	1,00	
	Inra	0,73	0,88	1,10	1,25	1,28	1,03	0,69	1,00	
	Universités	0,89	1,05	1,14	1,17	1,06	1,03	0,79	1,00	
	Autres FRA	1,14	0,88	1,00	0,90	1,10	1,00	1,03	1,00	
	Monde	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	
	Physique	CEA	2,12	1,64	1,42	1,35	1,02	0,80	0,50	1,00
CNRS		1,28	1,26	1,23	1,21	1,05	0,96	0,72	1,00	
Universités		1,19	1,15	1,23	1,18	1,05	1,00	0,73	1,00	
Autres FRA		1,40	1,05	1,34	1,20	1,08	0,81	0,76	1,00	
monde		1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	
Sciences de l'Univers		CEA	3,60	2,02	1,65	1,30	0,95	0,38	0,46	1,00
		CNRS	1,63	1,40	1,22	1,07	0,97	0,86	0,82	1,00
	Inra	0,20	1,35	0,66	0,94	1,30	1,11	0,91	1,00	
	IRD - France	0,87	0,65	1,13	1,04	0,88	0,94	1,16	1,00	
	Universités	1,32	1,25	1,07	1,04	1,02	0,97	0,86	1,00	
	Autres FRA	0,60	1,44	1,02	1,09	0,98	0,98	0,99	1,00	
	Monde	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	
Sciences pour l'Ingénieur	CEA	2,59	1,63	1,18	1,26	1,06		0,62	1,00	
	CNRS	1,63	1,40	1,29	1,22	1,01		0,77	1,00	
	Universités	1,44	1,30	1,27	1,20	1,07		0,78	1,00	
	Autres FRA	1,17	1,62	1,73	0,85	0,85		0,99	1,00	
	Monde	1,00	1,00	1,00	1,00	1,00		1,00	1,00	

année 1997

*ne sont gardées dans ce tableau que les institutions dont le total de présence dans la discipline dépasse strictement les 2 pour mille par rapport au monde

Nous nous limiterons ici à quelques observations sur les grandes disciplines. En biologie fondamentale, l'Inserm, le CNRS et le CEA montrent une distribution favorable, étant d'autant plus actifs que la classe de visibilité est élevée, cependant que l'Inserm est bien présent dans les tranches les plus visibles de la recherche médicale. En biologie appliquée-écologie, le CNRS, l'Inra et les universités montrent une répartition favorable. La distribution des articles impliquant le CEA est tout à fait favorable, les profils du CNRS et des universités étant également bons. De même pour les universités, en sciences de l'univers et en sciences pour l'ingénieur.

C - Nomenclature disciplinaire

les disciplines OST comme agrégats des "spécialités scientifiques" du SCI	
1. BIOLOGIE FONDAMENTALE	
Anatomie, morphologie	Microbiologie
Biochimie, biologie moléculaire	Microscopie
Biologie cellulaire, histologie	Neurosciences
Biologie moléculaire et cellulaire	Nutrition, diététique
Biomatériaux	Narasitologie
Biométhodes	Physiologie
Biophysique	Psychologie
Biotechnologie et microbiologie appliquée	Sciences comportementales
Embryologie	Systèmes reproducteurs
Génétique, hérédité	Techniques du laboratoire
Génie biomédical	Virologie
Imagerie neuronale	
2. RECHERCHE MEDICALE	
Allergologie	Médecine intégrative et de complément
Andrologie	Médecine légale
Anesthésiologie	Médecine tropicale
Cancérologie	Médecine vétérinaire
Chimie, clinique et médecine	Neurologie clinique
Chirurgie	Odontologie
Dermatologie, vénérologie	Ophthalmologie
Endocrinologie	Orthopédie
Ethique médicale	Oto-rhino-laryngologie
Gastro-entérologie	Pathologie
Gérontologie	Pédiatrie
Gynécologie, obstétrique	Pharmacologie - pharmacie
Hématologie	Pneumologie
Immunologie	Psychiatrie
Maladies infectieuses	Radiologie, médecine nucléaire
Médecine cardiovasculaire	Réhabilitation
Médecine cardiovasculaire 2	Rhumatologie
Médecine clinique, autres	Santé publique
Médecine d'urgence	Santé publique 2
Médecine d'urgence 2	Toxicologie
Médecine de la dépendance	Transplantations
Médecine du sport	Urologie - néphrologie
Médecine expérimentale	
Médecine interne générale	
3. BIOLOGIE APPLIQUEE-ECOLOGIE	
Agriculture	
Agronomie générale (agriculture, sc. sols)	
Biodiversité, conservation	
Biologie générale	
Biologie, autres	
Bois et textiles	
Botanique, biologie végétale	
Ecologie	
Entomologie	
Horticulture	
Mycologie	
Ornithologie	
Sciences des productions animales	
Sciences et techniques agro-alimentaires	
Sciences et techniques des pêches	
Stations agricoles expérimentales	
Sylviculture	
Zoologie générale	
4. CHIMIE	
Chimie analytique	
Chimie appliquée	
Chimie générale	
Chimie minérale et nucléaire	
Chimie organique	
Chimie physique	
Cristallographie	
Electrochimie	
Matériaux/analyse	
Science des matériaux	
Science des matériaux - bois, papier	
Science des matériaux - céramiques	
Science des polymères	
Traitements de surface	
5. PHYSIQUE	

Acoustique	
Instrumentation	
Matériaux composites	
Optique	
Physico-chimie	
Physique appliquée	
Physique des fluides et plasmas	
Physique des particules	
Physique du solide	
Physique générale	
Physique mathématique	
Physique nucléaire	
Physique, autres	
Spectroscopie	
6. SCIENCES DE L'UNIVERS	
Astronomie et astrophysique	
Biologie marine - hydrobiologie	
Géographie	
Géologie	
Géosciences	
Limnologie	
Météorologie	
Minéralogie	
Océanographie	
Paléontologie	
Ressources en eau	
Sciences de l'environnement	
Technologies de l'environnement	
7. SCIENCES POUR L'INGENIEUR	
Biocybernétique	Informatique et chimie
Composants	Informatique et robotique
Revue de synthèse en informatique (CRC)	Informatique et robotique 2
Contrôle	Informatique/applications
Contrôle 2	Informatique/divers 2
Energie et carburants	Informatique/imagerie
Génie aérospatial	Informatique/théorie et systèmes
Génie chimique	Ingénierie/systèmes
Génie chimique et thermodynamique	Intelligence artificielle
Génie civil	Mécanique
Génie de la construction	Métallurgie
Génie électrique et électronique	Photographie, imagerie
Génie industriel	Recherche opérationnelle
Génie maritime	Science et technologie nucléaires
Génie mécanique	Sciences de l'information
Génie métallurgique et minier	Systémique
Génie minier	Technologies marines
Génie pétrolier	Télécommunications
Informatique	Téledétection et télécontrôle
Informatique/divers	
8. MATHEMATIQUES	
Mathématiques	
Mathématiques appliquées	
Mathématiques générales	
Mathématiques théoriques	
Mathématiques, autres	
Méthodes mathématiques (biologie et médecine)	
Méthodes mathématiques (sciences physiques)	
Méthodes mathématiques (sciences sociales)	
Statistique et probabilités	