dépasser les frontières

Institut de l'information scientifique et technique

facilite l'accès aux résultats issus des différents champs de la recherche mondiale

Publications françaises Open Access 2010-2014
Executive summary

Octobre 2016

Afin de disposer d'éléments concrets pour éclairer la décision publique, le groupe de travail BSN4 a demandé à l'Inist la réalisation d'une étude bibliométrique sur les publications françaises en Open Access (OA) en science technique médecine (STM) sur une période de 5 ans (2010-2014), pour en connaître le poids dans la production scientifique française et tenter d'estimer le montant des APC. A cet *executive summary*, sont associés trois documents qui constituent l'étude complète :

- OpenAccess_BSN4_Méthodologie_30septembre2016.docx
- OpenAccess_BSN4_EtudeBibliométrique_30septembre2016.docx
- OpenAccess_BSN4_Remarques_Préconisations_30septembre.docx

Un certain nombre des résultats sont consultables via le tableau de bord dynamique créé à l'aide de l'outil libre ezVIS[1], dont l'accès est réservé.

Une recherche est effectuée dans le WoS sur les bases de données Science Citation Index Expanded (SCI-EXPANDED) et Conference Proceedings Citation Index- Science (CPCI-S) de Thomson Reuters. Elle est limitée aux publications françaises (France et outre-mer) et aux documents suivants : *articles, letters, proceedings papers* et *reviews*[2].

Après homogénéisation des titres des périodiques et alignement des titres puis des ISSNs des périodiques du corpus France avec ceux de 5 listes de périodiques OA et « Hybride »[3], les publications françaises en STM sont réparties en fonction du statut du périodique dont elles sont issues :

Hybride : publication issue d'un périodique considéré comme « Hybride » d'après les listes de périodiques des « Grands Editeurs ».
SansStatut : publication issue d'un périodique considéré comme sans statut (périodique non édité par les « Grands Editeurs » et marqué ni OA, ni « Hybride »).
NiNi_GrandsEditeurs : publication issue d'un périodique édité par un « Grand Editeur » marqué ni OA, ni « Hybride ».
OA : publication issue d'un périodique considéré comme OA par une des listes de périodiques suivantes : JCR, DOAJ, ROAD, OA des « Grands Editeurs ».

	Nb de publis	%
Corpus France	**358 157**	***100***
Hybride	**131 717**	***37***
SansStatut	**116 268**	***32***
NiNi_GrandsEditeurs	**76 096**	***21***
OA	**34 076**	***10***

Tableau 1 : Répartition du nombre de publications françaises par statut de périodiques 2010-2014

A ce stade, près de 70% des publications (37% « Hybride » et 32% « SansStatut ») ont un statut qui doit être précisé.

Après l'interrogation de chacun de leur DOI via l'API CrossRef, les résultats finaux concernant les 34 660 publications sont les suivants:

[1] https://github.com/madec-project/ezvis/

[2] Equation de recherche finale : (CU=(FRANCE OR MARTINIQUE OR "FR POLYNESIA" OR "FRENCH GUIANA" OR REUNION OR GUADELOUPE OR "NEW CALEDONIA" OR WALLIS OR MAYOTTE) and py=2010-2014) AND DOCUMENT TYPES: (Article OR Letter OR Proceedings Paper OR Review)
Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2016

[3] Les 5 listes de périodiques retenues :

1. La liste des périodiques considérés comme OA dans le JCR (Journal Citation Reports produit par Thomson Reuters) de chaque année
2. La liste de périodiques du DOAJ (Directory of Open Access Journals) au 30 mai 2016. Elle précise l'année à partir de laquelle le périodique est devenu OA.
3. La liste des périodiques ROAD (Répertoire des ressources scientifiques et universitaires en libre accès) en 2015.
4. La liste de périodiques OA de 7 éditeurs, appelés dans l'étude « Grands Editeurs », (Elsevier, Springer, Taylor &Francis, Wiley Blackwell, Nature, IEEE, ACS) en 2015.
5. La liste de périodiques « Hybride » des 7 « Grands Editeurs » ci-dessus en 2015.

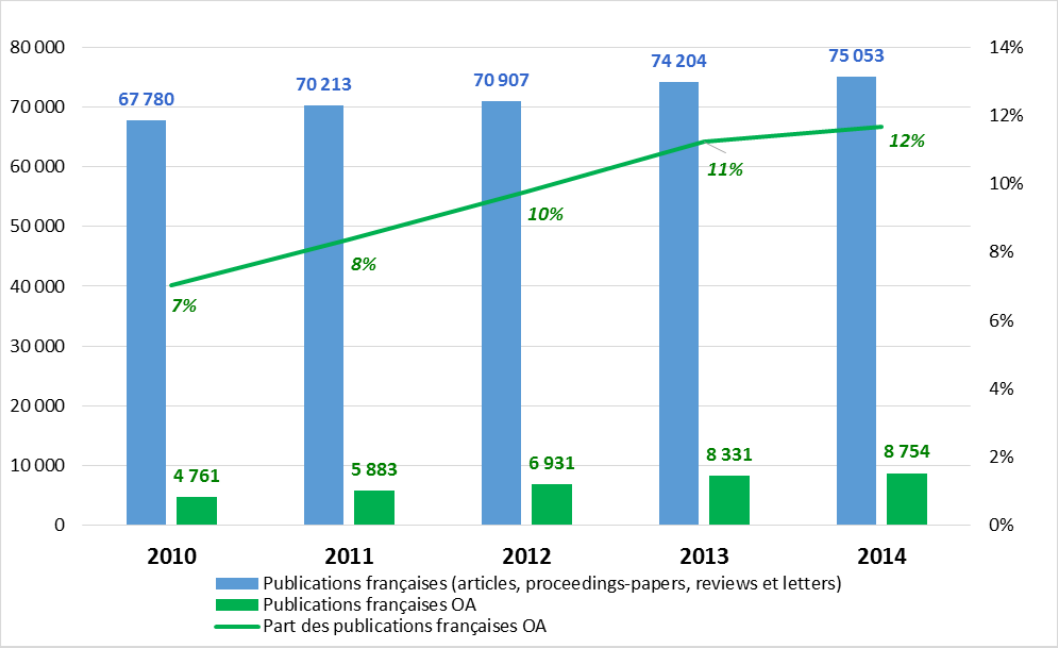

Figure 1 : Evolution du nombre de publications françaises en STM

La part de l'OA progresse régulièrement pour atteindre **12% (8 754 publications) en 2014** avec un taux de croissance sur la période 2010-2014 très élevé : +84%.

Un zoom sur la répartition des publications OA en fonction du statut du périodique montre que, bien que croissante sur la période, la **part des publications issues de périodiques « Hybride » (2% sur la période** avec 584 publications[4]) **est faible.**

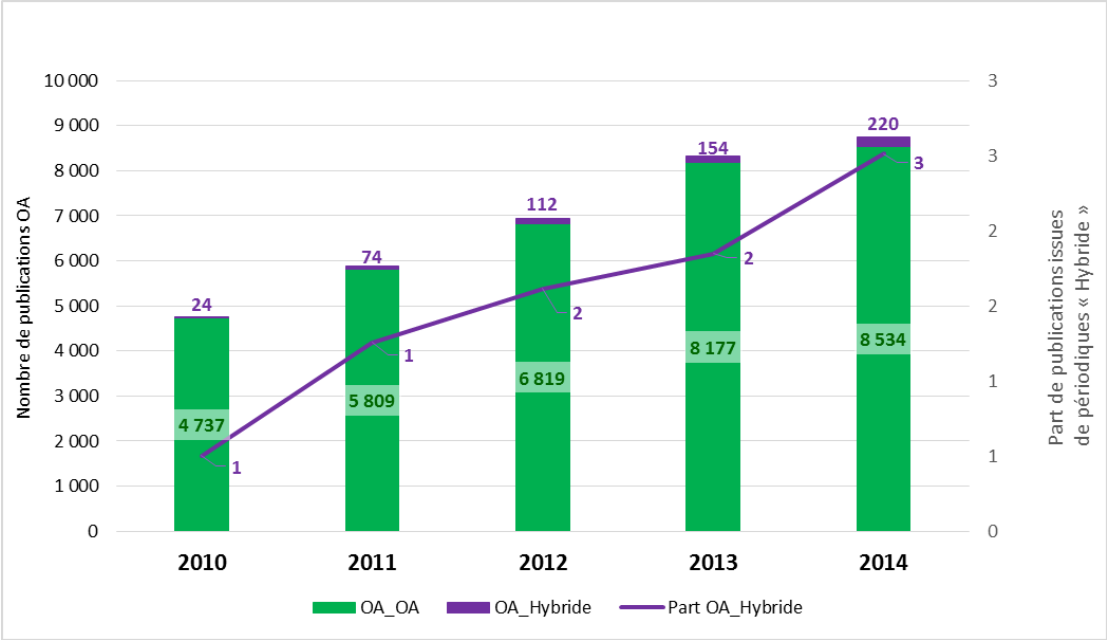

Figure 2 : Répartition du nombre de publications par statut des publications OA

Ce faible taux est lié à la difficulté de repérer ces types de documents et notamment à l'insuffisance des données présentes actuellement dans le réservoir CrossRef. Différentes pistes pour améliorer ce repérage sont envisageables :

[4] Sur près de 300 000 DOI interrogés dans l'application CrossRef, 584 publications ont les critères permettant de les considérer commes des publications OA issues de périodiques « Hybride ».

- Vérifier manuellement le statut de périodiques actuellement « SansStatut » les plus représentés (soit une soixantaine).
- Fournir les DOI de notre étude aux auteurs finlandais de l'étude longitudinale consacrée à la question des publications issues de périodiques « Hybride »[5].
- Utiliser des données de PubMed Central, avec pour conséquence de favoriser le domaine biomédical.
- Utiliser doai.io quand cette application proposera une API.
- Inciter les éditeurs à mieux alimenter CrossRef.

Les principaux résultats de l'étude bibliométrique concernant les **34 660 publications pour la période 2010-2014 sont présentés ci-dessous.**

- La répartition par type de documents des publications françaises OA en STM reflète celle des publications françaises globales, avec une forte prédominance des *articles*.

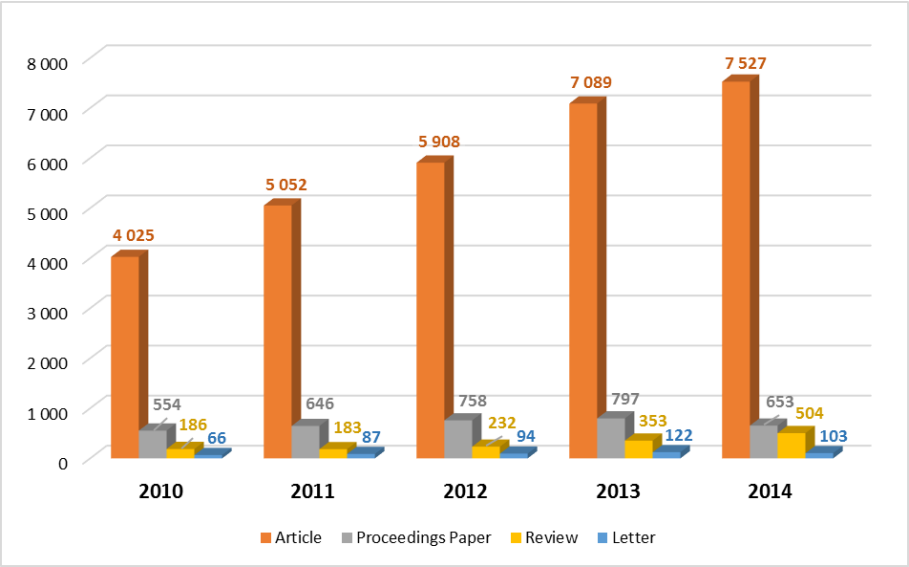

Figure 3 : Répartition du nombre de publications OA par type de documents

- La répartition par éditeur des publications issues de périodiques OA, montre qu'avec 12%, les « Grands Editeurs » représentent une faible part. En ce qui concerne les autres éditeurs, 2 d'entre eux, *Public Library Science* et *Biomed Central LTD*, se détachent nettement avec respectivement 24 et 14%, soit à eux 2, plus de 3 fois la part des « Grands Editeurs ». Une grande dispersion des autres éditeurs est à signaler avec plus de 600 éditeurs représentant 20% de la part des publications OA.
- La répartition par éditeur des 584 publications OA issues de périodiques « Hybride » montre qu'*American Physical Society* et *ACS* représentent respectivement plus 60 et 35% de ces publications sur la période.
- La répartition, par périodique, des publications OA issues de périodiques OA montre que la moitié des publications françaises OA en STM sont issues de seulement 2% des périodiques OA (26 titres), et l'autre moitié dans les 98% restant (1 122 titres). Avec 6 776 publications sur 34 076, *PLOS ONE* représente à lui seul 20% des publications avec une forte progresssion sur la période.
- La répartition, par périodique, des publications OA issues de périodiques « Hybride » montre que sur les 55 titres, *Physical Review Letters* et *Physical Review D* se détachent nettement sur

[5] Laakso, M., & Björk, B.-C. Hybrid open access—A longitudinal study. Journal of Informetrics (2016), http://dx.doi.org/10.1016/j.joi.2016.08.002

la période (avec respectivement 32 et 23% des publications) tout en voyant leur nombre et part de publications diminuer.

- La répartition dans les 22 disciplines ESI[6] montre la prédominance du domaine biomédical avec plus de la moitié des publications. En 2012, la très forte augmentation du nombre de publications « *Multidisciplinary* » et en parallèle la très forte diminution de publications « *Biology & Biochemistry* » sont liées au changement de marquage du périodique *PLOS ONE* (passé de « *Biology & Biochemistry* » à « *Multidisciplinary* »).

Disciplines ESI	2010		2011		2012		2013		2014		2010-2014	
	Nb	%	Nb	%	Nb	%	Nb	%	Nb	%	Nb	%
Corpus France OA	**4 761**	*100*	**5 883**	*100*	**6 931**	*100*	**8 331**	*100*	**8 754**	*100*	**34 660**	***100***
Clinical Medicine	1 054	*22*	1 219	*21*	1 261	*18*	1 532	*18*	1 910	*22*	**6 976**	***20***
Biology & Biochemistry	1 278	*27*	1 796	*31*	763	*11*	945	*11*	1 191	*14*	**5 973**	***17***
Physics	914	*19*	1 034	*18*	1 336	*19*	1 273	*15*	1 341	*15*	**5 898**	***17***
Multidisciplinary	42	*1*	34	*1*	1 715	*25*	2 099	*25*	1 769	*20*	**5 659**	***16***
Engineering	467	*10*	796	*14*	965	*14*	1 057	*13*	1 052	*12*	**4 337**	***13***
Geosciences	409	*9*	588	*10*	635	*9*	874	*10*	777	*9*	**3 283**	***9***
Molecular Biology & Genetics	463	*10*	459	*8*	546	*8*	723	*9*	881	*10*	**3 072**	***9***
Chemistry	328	*7*	411	*7*	401	*6*	486	*6*	525	*6*	**2 151**	***6***
Plant & Animal Science	317	*7*	288	*5*	322	*5*	386	*5*	517	*6*	**1 830**	***5***
Microbiology	262	*6*	300	*5*	313	*5*	466	*6*	469	*5*	**1 810**	***5***
Environment/Ecology	163	*3*	329	*6*	313	*5*	365	*4*	355	*4*	**1 525**	***4***
Materials Science	104	*2*	234	*4*	216	*3*	380	*5*	334	*4*	**1 268**	***4***
Immunology	153	*3*	173	*3*	157	*2*	261	*3*	346	*4*	**1 090**	***3***
Mathematics	158	*3*	198	*3*	215	*3*	254	*3*	245	*3*	**1 070**	***3***
Social Sciences, general	132	*3*	154	*3*	228	*3*	220	*3*	228	*3*	**962**	***3***
Neuroscience & Behavior	89	*2*	96	*2*	111	*2*	204	*2*	337	*4*	**837**	***2***
Computer Science	122	*3*	135	*2*	157	*2*	161	*2*	155	*2*	**730**	***2***
Agricultural Sciences	106	*2*	129	*2*	82	*1*	119	*1*	107	*1*	**543**	***2***
Space Science	32	*1*	104	*2*	110	*2*	114	*1*	100	*1*	**460**	***1***
Pharmacology & Toxicology	57	*1*	53	*1*	64	*1*	99	*1*	122	*1*	**395**	***1***
Psychiatry/Psychology	40	*1*	37	*1*	55	*1*	97	*1*	143	*2*	**372**	***1***
Economics & Business	0	*0*	3	*0*	3	*0*	5	*0*	11	*0*	**22**	***0***

Tableau 2 : Répartition du nombre de publications OA par discipline ESI

[6] La somme du nombre de publications est supérieure au nombre total annuel car un périodique – et par répercussion une publication – peut être classé dans plusieurs disciplines ESI.

Le calcul des montants des APCs s'appuie d'une part sur les données des fichiers du DOAJ et de l'Inserm et d'autre part sur l'hypothèse que le pays (puis selon la même logique l'organisme) payeur est mentionné dans le champ « *Reprint Author* » (RP).

La part des publications RP France est élevée. D'après ce mode de calcul, **la France paierait des APCs pour plus de 60% de ses publications OA**.

Avec un taux de progression moyen de plus de 20% par an, le montant des APCs des publications OA RP France a plus que doublé sur la période 2010-2014 pour atteindre **plus de 6 Millions € en 2014** alors que son nombre de publications OA RP France a lui été multiplié par moins de 1,5.

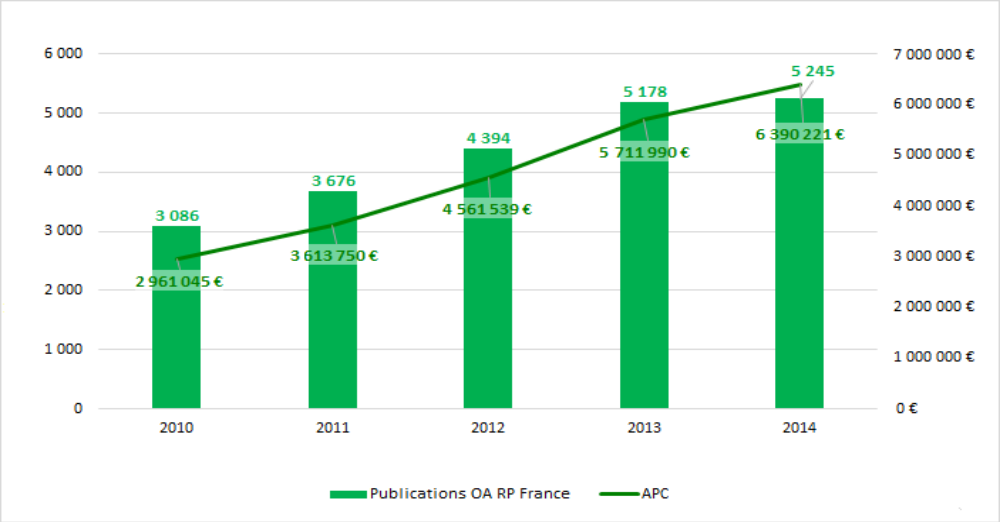

Figure 4 : Evolution du nombre de publications OA RP France et des montants APCs

Un zoom sur 5 organismes de recherche (CEA, CNRS, Inra, Inserm, ParisTech) est effectué. Si le repérage du nombre de publications OA pour les organismes est confirmé pour l'Inra et l'Inserm, une comparaison avec les coûts réels payés par l'Inserm à BioMed Central pour l'année 2014 met en évidence un surcoût. Tous ces chiffres fortement surestimés sont donc à nuancer. Une approche exploitant les données comptables est à approfondir.

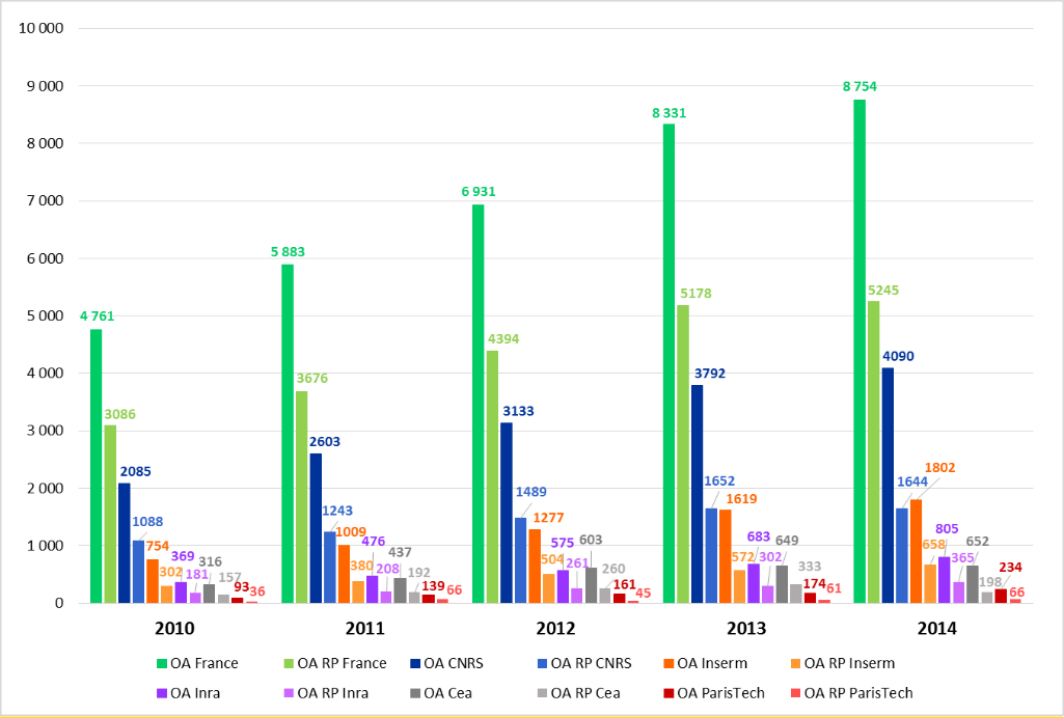

Figure 5 : Répartition du nombre de publications OA France et OA RP France CNRS, Inserm, Inra, CEA ou ParisTech

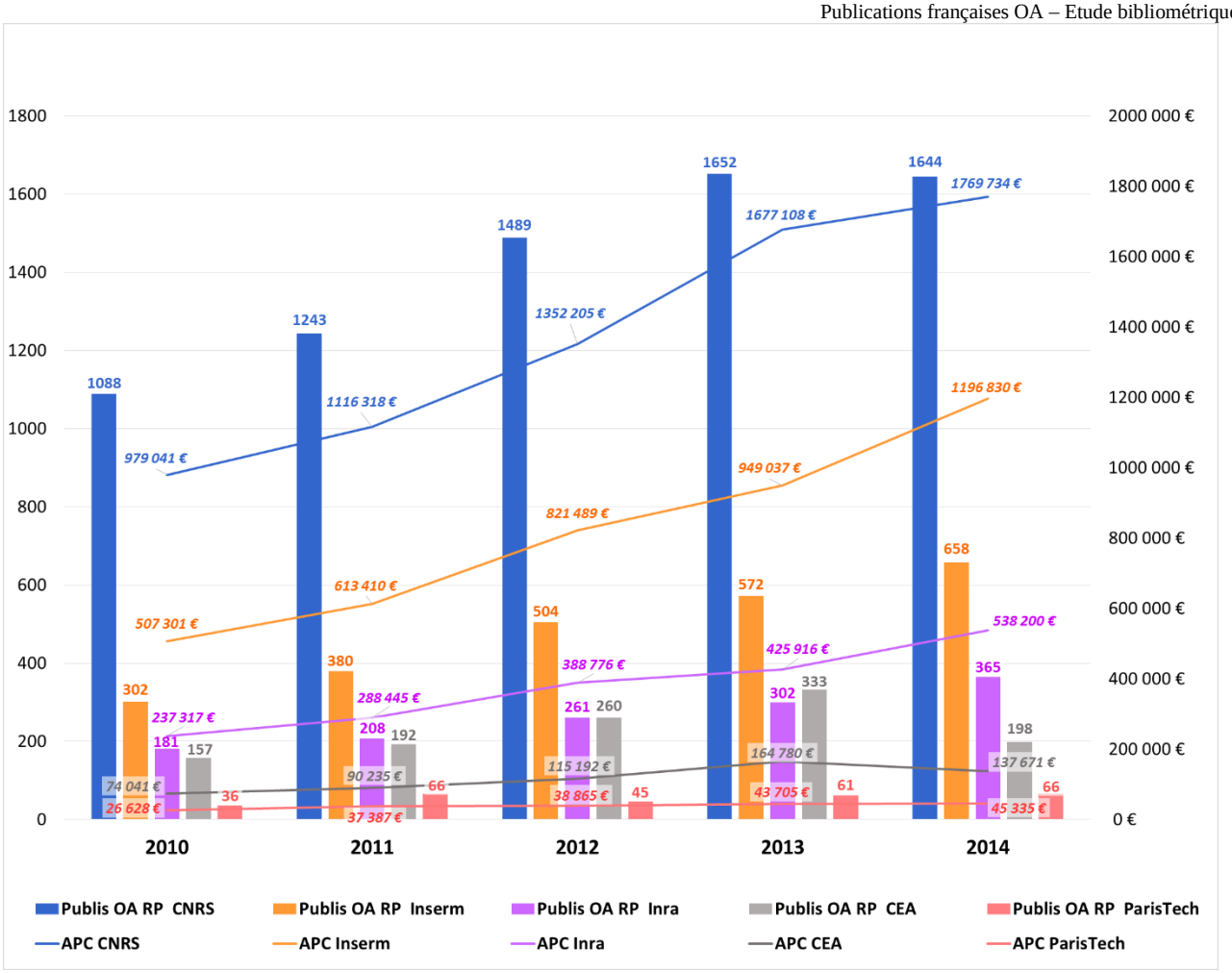

Figure 6 : Répartition du nombre de publications OA RP CNRS, Inserm, Inra, CEA et ParisTech et des montants APC

Conclusions

L'étude montre ou confirme :

- l'augmentation de la production scientifique en Open Access en France
- la difficulté de repérer les publications OA issues de périodiques « Hybride »
    - la nécessité d'inciter les éditeurs à alimenter CrossRef ou à mieux marquer les publications OA issues de périodiques « Hybride » dans les bases de données bibliographiques,
    - la nécessité d'un réservoir unique d'une liste de périodiques auxquels seraient associées leurs différentes caractéristiques (titres, ISSNs, statut et montant des APC par année, éditeur …) pour chaque année pour éviter la manipulation de multiples fichiers,
- la difficulté à affecter un coût à un organisme de recherche avec comme conséquence une surévaluation des montants de APC estimés
    - la nécessité de combiner les données comptables aux données bibligraphiques.