

Monitoring the Transition to Open Access

A report for the Universities UK
Open Access Co-ordination Group

August 2015



The
University
Of
Sheffield.



Contents

Acknowledgements	2
Executive Summary	3
1. Introduction	11
2. Open Access options available to authors	13
3. Accessibility: take-up of OA publishing models and posting of articles	28
4. Usage of OA and non-OA articles	42
5. Financial sustainability: universities	49
6. Financial sustainability: learned societies	64

Annexes

A Publishers responsible for journals in which UK authors frequently publish	73
B Journals with the most articles published by authors with a UK affiliation	74
C APCs and licensing for OA and hybrid journals	75
D Posting and embargo policies	79
E Methodology and quality assurance for the assessment of accessibility	80
F Costs to universities for seven subscription and three fully-OA publishers	84
G Methodology for the assessment of financial sustainability for learned societies	85
H Financial analysis of sampled learned societies	87
I Sample of 30 learned societies	96
J References	97
K Abbreviations and glossary	100
L REF main panels	102
M Desirable improvements to data quality and availability	103

Acknowledgments

The studies upon which this report is based were undertaken by a team led by Michael Jubb of the Research Information Network (RIN) and comprising

Stéphane Goldstein, *RIN*

Mayur Amin, Andrew Plume, Stephanie Oeben and a technical team led by M'Hamed Aisati, *Elsevier*

Stephen Pinfield, Peter Bath and Jennifer Salter, *University of Sheffield*

Rob Johnson and Mattia Fosci, *Research Consulting*

Each organisation took lead responsibility for one or more of the key strands of work:

- ❑ Open Access options available to authors: RIN
- ❑ Accessibility: take-up of OA publishing models and posting of articles: Elsevier
- ❑ Usage of OA and non-OA articles: RIN
- ❑ Financial sustainability for universities: University of Sheffield
- ❑ Financial sustainability for learned societies: Research Consulting

We stress, however, that members of the team as a whole provided critical levels of input into all the strands of work, and that we take collective responsibility for this report.

We also wish to thank all those publishers, and staff in UK universities, in learned societies, and in Jisc who helped us by providing data and answering our questions. This report could not have been produced without them.

Finally, we wish to thank those members of the Universities UK OA Co-ordination Group who made useful comments and suggestions in response to early presentations of our findings.



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Executive Summary

This study was commissioned in response to a recommendation of the Finch Group in its second report in 2013 that reliable indicators should be gathered on key features of the transition to open access (OA) in the UK. The findings presented here are thus a first attempt at generating such indicators covering five sets of issues:

- ❑ *OA options available to authors*: the numbers of fully-OA and hybrid journals, along with issues such as the level of article processing charges (APCs), the availability of CC-BY and other licences, and the length of embargo periods
- ❑ *Accessibility: authors' take-up of OA options*: the numbers - and the proportions of the overall population – of articles accessible on OA terms via different routes
- ❑ *Usage*: the levels of usage of OA articles as compared to those that are not accessible on OA terms
- ❑ *Financial sustainability for universities*: the amounts paid by UK universities in subscriptions and in APCs; and
- ❑ *Financial sustainability for learned societies*: the overall income and expenditure – as well as the volumes of journal-related income and expenditure – of UK learned societies which have some publishing income.

There are of course other issues highly relevant to the monitoring of progress towards OA, including such matters as the quality of services provided by publishers to authors and readers; and we hope that these will be addressed in subsequent studies.

We are aware that the data we have been able to gather presents a number of challenges to which we draw attention in the body of this report. We make some recommendations that could lead to improvements in the quality of the data; and we are also aware that in subsequent exercises, improvements could be made in the methodologies we have adopted. Hence we have tried to be both clear and cautious in presenting our findings, which we summarise below.

OA options available to authors

Publishing models

OA publishing options are now widely available: two-thirds of the world's journals offer an OA option of some kind; and more than three-quarters of the journals in which UK authors publish do so. By far the largest group of journals have adopted the hybrid model: just under half of all journals across the world operate in this way, and nearly two-thirds of those in which UK authors publish their work.

As a corollary, the numbers and proportions of subscription-only journals fell between 2012 and 2014; with a particularly sharp fall among the journals in which UK authors publish. But the publishers with whom UK authors most commonly publish their work show markedly varied profiles in their adoption of fully-OA, hybrid and subscription-only publishing models.

APCs and other charges

Levels of APCs vary widely. Only a small minority of fully-OA journals charge more than £2,000; but below that level there is wide variation between both journals and publishers. The great majority of hybrid journals charge APCs of between £1,000 and £2,000; only small minorities, concentrated in a few publishers, charge either less than £1,000 or more than £2,000.

Some journals levy other charges, irrespective of and unrelated to OA, in the form of submission or publication fees, page charges and charges for the use of colour in figures in the printed versions of articles. In some cases, where an OA option is also available and taken up, these other charges can exceed the level of the APC.

Licensing

Creative Commons licences have become widely, but not universally, accepted as a mechanism for promoting OA. Most journals - in social sciences and humanities (HSS) as well as in STEM subject areas – at least *allow* publishing under a Creative Commons CCBY licence, though there is widespread variation as to whether it is a default or an option. Policies vary also as to whether the CCBY licence or the more restrictive licences that preclude commercial use (CCBYNC) and/ or the creation of derivatives (CCBYND) are employed. Some journals allow Creative Commons licences, particularly CCBY, only when it is a funder requirement.

Posting policies and embargoes

Posting policies for subscription-based articles are complex, with considerable variations between journals and publishers; and full details are often difficult to interpret, or even to find. Policies are often modified in response to funder requirements. In general, policies are more permissive for pre-prints and authors accepted manuscripts (AAMs) than for versions of record (VoRs). Four-fifths of journals allow the posting of AAMs on personal websites with embargoes between zero and six months. But policies are progressively more restrictive for posting in institutional repositories (IRs), subject repositories or other services. Only a small minority of subscription-based journals – mostly in the physical sciences – allow VoRs to be posted on any website.

Accessibility: authors' take-up of OA options

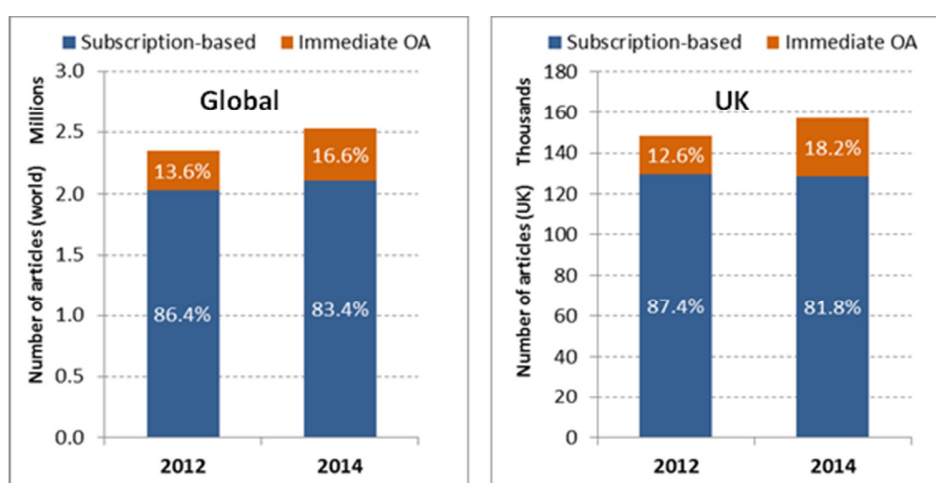
Publishing Models

The number of articles published globally in journals with an immediate OA publishing model (fully-OA journals, and OA articles in hybrid journals) grew faster between 2012 and 2014 than articles in subscription-based journals, although both increased in absolute terms. Growth was fastest in the take-up of the OA option in hybrid journals; but global take-up of publishing in fully-OA journals that do not charge an APC was static. In sum, journals with an immediate OA model accounted for just under 17% of global articles in 2014 (compared to 14% in 2012).

In the UK, take-up of immediate OA models grew even faster than the global average; but publication in subscription-based journals was essentially static (-0.4%). As a result, whereas UK take-up of immediate-OA publishing models was slightly below the world average in 2012 (13% as compared to 14% globally),

it had moved ahead of the world average by 2014 (over 18% as compared to under 17%). But the UK's profile of OA take-up is significantly different from the global averages: its use of OA in hybrid journals and of delayed OA journals is more than twice the world average in both cases, while its take-up of fully-OA journals with no APC (Gold-no APC) is less than half the world average and falling. It is also noticeable that UK authors show a preference for publishing in journals with higher citation rates in their field (as measured by the field-weighted citation impact (FWCI)) and that citations for hybrid and delayed OA journals are higher than APC-based fully OA journals (Gold-APC) which, in turn, are cited more than Gold-no APC journals (see table 9b).

Summary Figure 1. Journal publishing models employed by Global and UK authors



Take-up varies significantly by subject area, both globally and in the UK. When analysed in accordance with the four main panels used in the UK Research Excellence Framework (REF) it is highest in the health and life sciences (Panel A) and significantly lower in other subject areas. But again, the UK shows differences from the world averages: UK take-up is higher than the global average in the health and life sciences (Panel A), and in the physical sciences and engineering (Panel B); but lower in the social sciences, and in the arts and humanities (Panels C and D)¹.

Postings

Although this study provides data on levels of postings across immediate-OA and subscription-based business models (see Section 3 for full details), the summary below focuses mostly on subscription-based models, given that articles in OA journals (including OA articles in hybrid journals) are already openly accessible upon publication.

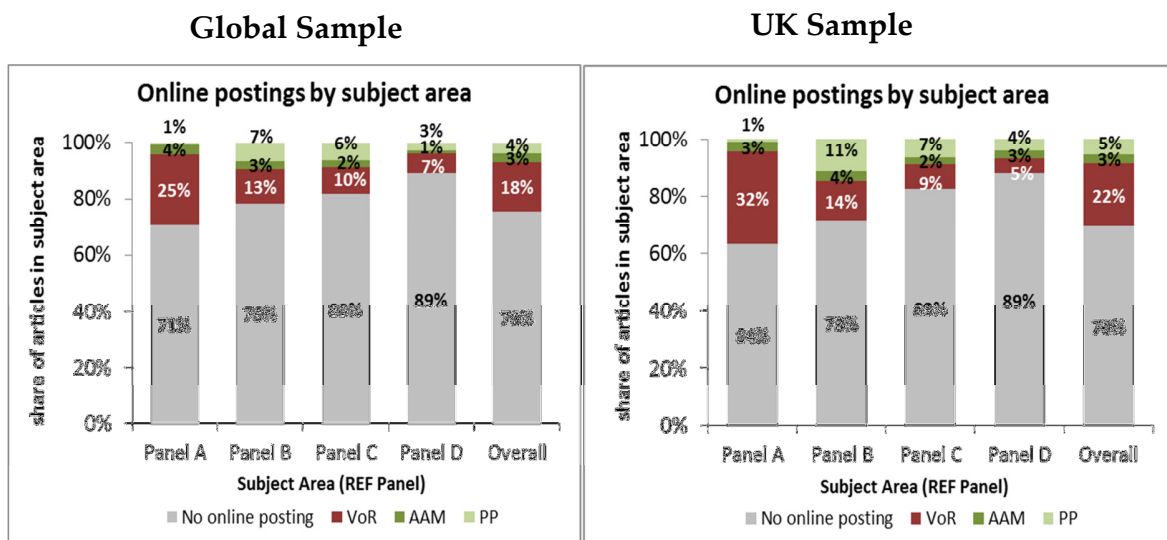
Versions of 15% of the papers published globally in the past two years – and 19% of papers published by UK authors - have been posted and are accessible online and in accordance with journal policies. UK postings are most strongly ahead of the world average in health and life sciences.

¹ A list of the disciplines and subject areas covered by the four main REF panels is at Annex L.

These percentages include postings of articles that were already openly accessible immediately on publication, on the publisher's site; excluding these we estimate that 9% of world and 10% of UK papers published in subscription-based journals are publicly accessible online in versions posted on various online sites and in accordance with journal policies.

We also estimate that a further 11% of UK articles, and 9% of global articles have been posted 'illicitly' in the sense that they were not in accordance with the terms of journal policies. The vast majority of 'illicit' postings (>90%) were of VoRs posted in the main on social sharing sites (e.g. ResearchGate), contrary to the policies of the relevant journals. But if we include in our calculation illicit postings as well as those in accordance with journal policies, a total of 19% of global papers in subscription-based journals, and 23% of such papers with UK authors, are freely accessible in repositories and other online sites. Subject repositories such as PubMedCentral (PMC) and social sharing sites such as ResearchGate are the predominant locations of posted articles; and articles become accessible progressively over the two years post-publication, with most becoming accessible by 12 months.

Summary Figure 2. Variation in online posting by REF panel subject area (includes postings related to both immediate OA and subscription-based journals as well as illicit postings)



Estimated total proportion of OA articles

Taking both OA publishing models and postings together, we estimate that globally 18% of articles published in the last two years were openly accessible immediately on publication. If illicit postings of articles (not in accordance with journal policies) are included, the proportion increases to 19%. The proportions also rise over time so that 20% (23% including illicit postings) are freely accessible within six months, 25% (30% incl. illicit) within 12 months and 27% (34% incl. illicit) within 24 months.

For UK articles, 20% of articles published in the last two years (22% including illicit postings) were openly accessible immediately upon publication, rising to 24% (28% incl. illicit) within six months, 32% (38% incl. illicit) within 12 months and 35% (43% incl. illicit) within 24 months².

SummaryTable1. Total proportions of OA content globally and for the UK

(a) Total proportion of Open Access content - 2014

GLOBAL	Cumulative %				
Months (after publication)	0	6	12	24	
<i>Sampled months*</i>	1-2	7-8	13-14	25-26	
Gold-APC	9.6%	9.6%	9.6%	9.6%	All articles immediately OA upon publication
Gold-noAPC	4.6%	4.6%	4.6%	4.6%	
Gold-Hybrid	2.4%	2.4%	2.4%	2.4%	
Delayed OA		0.8%	4.1%	4.5%	A further 1% accessible later for embargoes >24 months
OA Postings (excl. illicit postings)** ⁺	1.3%	2.5%	4.2%	6.2%	Subscription articles only
TOTAL (excl. illicit postings)**⁺	17.9%	19.9%	24.9%	27.3%	An additional 3% accessible free at publisher sites***

(b) Total proportion of Open Access content - 2014

UK	Cumulative %				
Months (after publication)	0	6	12	24	
<i>Sampled months*</i>	1-2	7-8	13-14	25-26	
Gold-APC	9.3%	9.3%	9.3%	9.3%	All articles immediately OA upon publication
Gold-noAPC	2.1%	2.1%	2.1%	2.1%	
Gold-Hybrid	6.5%	6.5%	6.5%	6.5%	
Delayed OA		1.8%	8.7%	9.6%	A further 1.5% accessible later for embargoes >24 months
OA Postings (excl. illicit postings)** ⁺	2.0%	4.2%	5.5%	7.5%	Subscription articles only
TOTAL (excl. illicit postings)**⁺	19.9%	23.9%	32.1%	35.0%	An additional 3% accessible free at publisher sites***

* for postings and Gold-Hybrid only. Up to two months were sampled to allow for lead time for posting to be discoverable and indexed by search engines

** includes posting for subscription content only, deduplicated for multiple postings and for content also available via delayed OA

*** usually for promotional or public service purposes

⁺ illicit postings are those that do not comply with journal policies. The vast majority (>90%) of 'illicit' postings was the result of publishers' versions of record being posted, contrary to the policies of those journals (which allow posting of AAMs but not VoRs)

⁺⁺ Including illicit postings the total proportion on openly accessible content after 24 months would be 34% (Global) and 43% (UK). Full details are given in Section 3, Table 11

² Our study is about OA, of course, and not about other access initiatives. These figures thus represent accessibility in the developed world, and we do not deal in this study with the large corpus of literature that is freely-accessible to users in developing countries via programmes such as Research4Life.

Usage of OA and non-OA articles

A key aim behind moves to promote OA is to make articles freely accessible, so that they can be read and used by more people, from a wider range of sectors. It is important, therefore, to examine whether the policies now being implemented do indeed have that effect. But in addressing the question whether OA does in fact lead to more usage, we need to take account of articles that can be found on a number of platforms, and there is as yet no service that aggregates usage data from anything like the full range of sites. Moreover, there is no data that enables us to determine the demographics of users, or to distinguish between users from the higher education sector and elsewhere.

Data from publishers indicates that downloading from their platforms is indeed higher for OA than for non-OA articles. But patterns across different journals, both fully-OA and hybrid, vary hugely: for one publisher, usage of OA articles in hybrid journals ranged from ten times that for non-OA articles to equality between the two kinds of articles. Similarly, downloads in UK universities of articles from more than 150 publishers shows no obvious pattern in the ratio of downloads of OA as against non-OA articles.

Downloads of articles from UK IRs appear to be highly skewed towards a small number of very popular journals and articles. But downloads from UK IRs are dwarfed by those from major subject repositories, especially PubMedCentral. Unfortunately, we were not able to gather any usage data from sharing sites such as ResearchGate, where it is suggested that usage has risen greatly in recent years.

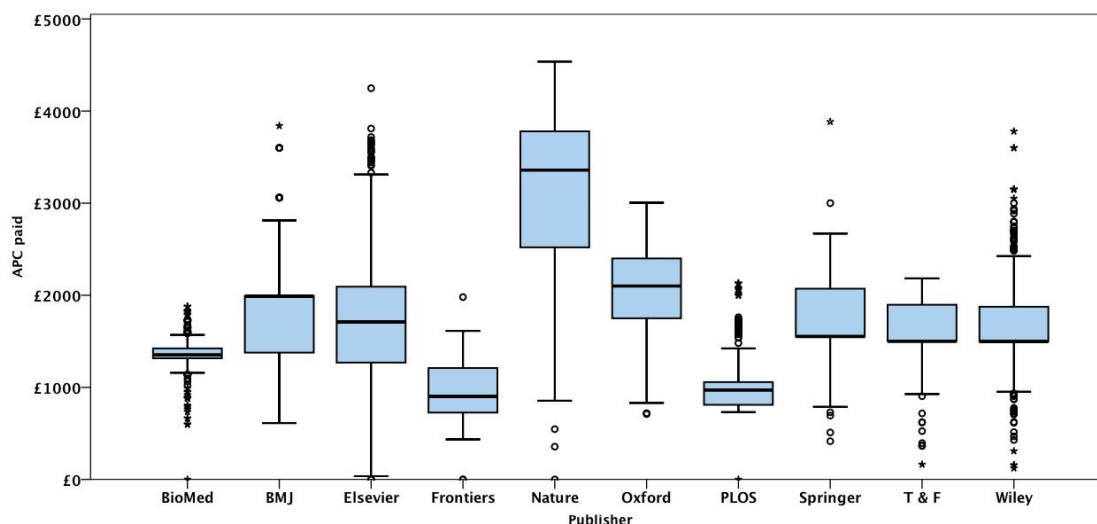
Our conclusion is that there is an urgent need for more openly-accessible usage data, or at least wider adoption of the models developed by the Publisher and Institutional Repository Usage Statistics (PIRUS) project – and since adopted by the COUNTER organisation - for gathering and aggregating article-level download data from a range of sites. Without such a service, which would need to encompass the major subject repositories as well as the sharing sites, aggregating usage data for OA and non-OA articles will remain partial at best. And it appears that there is little prospect of being able to gather data that would enable us to answer questions about the demographics of usage.

Financial sustainability: universities

The APC market is currently complex, with variable pricing, discounts, and additional charges, and these complexities are reflected in the actual payments made by universities. Analysis of the data from 24 universities in the UK on payments in 2014 of APCs and of subscriptions demonstrates that centrally-managed APC expenditure has continued to rise steeply (555% since 2012). But there was considerable variation in the levels of payments across different institutions, reflecting differences in levels of research activity and in institutional policies.

The largest number of payments was made for articles in the health and life sciences; and to commercial publishers.

Summary Figure 3. Range of APC payments for the top 10 publishers

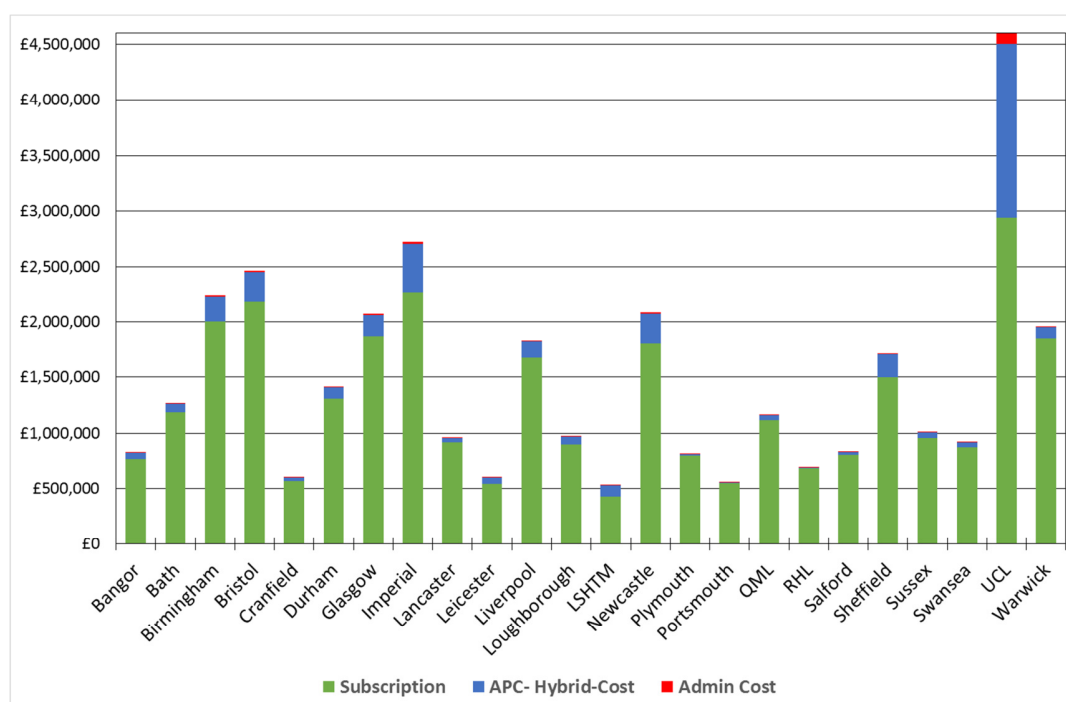


The level of payments varied from £0 to £4,536 with a mean of £1,586; and hybrid journal APCs were more expensive than those for fully-OA titles. Some of this variation may be explained by variations in the citation impact of different journals: for there is a correlation between APC price and the citation rates of journals as measured by the field-weighted citation impact (FWCI).

For a sample of 24 universities and seven major publishers that are mainly subscription-based, APCs now constitute 12% of universities' total expenditure on journals, with 1% for the administration of APCs, and 87% for subscriptions. When three major fully-OA publishers are added, APCs rise to 14% of costs, with subscriptions falling to 85%.

It is important that the trends and profiles of universities' expenditure should continue to be monitored, and we make a number of recommendations as to how data gathering and analysis could be made more effective for the future.

Summary Figure 4. Costs to universities for seven publishers, 2014



Financial sustainability: learned societies

Nearly 280 learned societies in the UK publish scholarly journals and conference proceedings, and we estimate that out of their total revenues of c£1.2bn, some £318m (26%) derives from publishing. In many cases the journals published by these societies are among the leading journals in their field internationally. Societies are evenly distributed across the four subject areas defined by the main REF panels; but both overall and publishing revenues are highly concentrated in the health and life sciences, and physical sciences and engineering. Most societies (63%) publish a single journal, but a sizeable minority (22%) publish three or more. Just under a quarter of societies (24%) publish on their own account, but the majority make use of the services of commercial publishers and university presses.

Analysis of the published accounts of a sample of societies indicates that the proportion of revenues derived from publishing varies widely, and there is no simple correlation between the proportion of revenues derived from publishing and a society's size or disciplinary focus. Levels of surplus and deficit from publishing also vary widely, with some societies showing a deficit while societies in the social sciences show on average a strong surplus.

The published accounts provide no evidence that up to the end of 2013 OA had any adverse impact on societies' publishing revenues and overall financial health. This is not unexpected, since RCUK's OA policy came into effect only in April 2013, and initially required only partial compliance for UK publications supported by RCUK funding. Moreover, many societies generate a high proportion of their revenues overseas, and have long-term agreements with commercial publishers. Hence the impact of the transition to OA may not be evident in societies' finances for a number of years. Nevertheless, further detailed work is required on the finances of a sample of societies, to provide a fuller picture than can be derived from their published accounts.

1. Introduction

1.1 The study, its aims and its limitations

The Finch Group's two reports (Finch, 2012 and 2013) both stressed the importance of establishing a process to collect authoritative indicators of key features of the transition to OA in the UK. The Group also stressed that funders, universities, publishers and learned societies should work together in that enterprise, in order to sustain the confidence of all parties in the process and its results. A working group convened by the Research Information Network (RIN), set out proposals (Research Information Network 2014) for a set of indicators that could be generated on a regular basis.

This study was commissioned by Universities UK's Open Access Co-ordination Group as a first exercise in collecting the data and generating indicators of the kind outlined in the working group's report, covering five sets of issues:

- ❑ *OA options available to authors:* the numbers of fully-OA and hybrid journals, along with issues such as the level of article processing charges (APCs), the availability of different kinds of licences (particularly CCBY) and the length of embargo periods
- ❑ *Accessibility: authors' take-up of OA options:* the numbers - and the proportions of the overall population - of articles published and accessible on OA terms via different routes
- ❑ *Usage:* the levels of usage of OA articles as compared to those that are not accessible on OA terms
- ❑ *Financial sustainability for universities:* the amounts paid by UK universities in subscriptions and in APCs; and
- ❑ *Financial sustainability for learned societies:* the overall income and expenditure - as well as the volumes of journal-related income and expenditure - of UK learned societies.

A final set of issues discussed in the RIN working group's report related to the quality of services provided by publishers to authors and readers. This was, however, excluded from the scope of our work. Like a number of other issues to which we draw attention in the body of this report, however, we believe that it is important that the quality of the services provided by publishers should be the subject of study during the course of the transition to OA.

We stress that this is an initial exercise. Although we draw on some previous work, we are breaking new ground in a number of areas. In line with the Finch recommendations, our aim has been to establish a reasonably authoritative baseline from which trends can be ascertained in subsequent studies. But we are aware that the data we have gathered presents a number of challenges, to which we draw attention in the body of this report. We make some recommendations that could lead to improvements in the quality of the data (see Annex M); and we are also aware that in subsequent exercises, improvements could be made in the methodologies we have adopted. For we agree with the Finch Group's recommendation that a key aim for all those concerned with the future of scholarly communications should be to build an authoritative set of data on trends in the development and take-up of OA in the UK. To that end, we have sought to be as explicit as possible about our methodologies, so that they can be replicated, and where necessary improved on, in subsequent exercises. See our recommendations relating to desirable improvements to data availability and gathering at Annex M.

In that context, we have been tried to be both clear and cautious in presenting our findings, and to acknowledge the limitations both of the data and of our methodologies. In the spirit of openness, we are also making the key sets of data we have gathered available (anonymised in some cases) for other researchers to analyse, since we are aware that they may be open to other interpretations. They will be made accessible on Figshare.

1.2. Previous work on OA

We are of course aware of previous work on some of the issues we deal with in this report. Bo-Christer Bjork and his colleagues at the Hanken School of Economics in Finland and in the US have published a large corpus of work on matters including the availability of different kinds of OA options, the take-up of those options, and the costs of APCs. Other significant studies have been undertaken by Yassine Gargouri and his colleagues at the Université de Québec and the University of Southampton, and Eric Archambault and his colleagues at Science Metrix. A key problem with this corpus of work – as Bjork himself has pointed out – is that the use of different bibliographic databases, different time-windows and different methods make comparisons and the analysis of trends difficult. Estimates, for example, of the numbers and proportions of articles deposited and accessible via author websites, IRs, subject repositories and other sites thus differ widely. Nevertheless, we refer to these and other studies at appropriate points in the body of this report; but we stress that a full literature review was beyond the scope of this study. A full list of references is at Annex J.

1.3 Terminology

The terminology used in relation to OA can be confusing and ambiguous. In order to minimise confusion, we have tried to be consistent in this report in using the terminology and labels outlined below to describe various publishing options (and their groupings) and versions of articles posted in various locations.

PUBLISHING MODELS

	Label	Publishing Model	Notes
Immediate OA	Gold-APC	APC only	Journals that are fully OA (Open Access) and require an article publishing charge (APC). Mostly Online only
	Gold-no APC	Funded by means other than APC	Journals that are fully OA (online) and do not require an APC. These are usually centrally subsidised or sponsored by various bodies and/or supported by subscriptions to print copies. Online only or Online + Print
	Gold-Hybrid	APC opted	Hybrid uptake, i.e. those articles in Hybrid journals that are OA where authors have opted to pay an APC. Online + Print
Subscription-based	Delayed OA	Subscription	Subscription model with delayed open access to final articles (VoR) on publisher sites
	Hybrid Potential	Subscription	Articles in Hybrid journals excluding those that are OA via an APC (Gold-Hybrid), i.e. subscription-based articles in Hybrid journals
	Subscription only	Subscription	Subscription only journals
sub-totals	Immediate OA	The sum of 'Gold-APC' + 'Gold-noAPC' + 'Gold-Hybrid'	
	Subscription-based	The sum of 'Hybrid Potential' + 'Subscription only' + 'Delayed OA'	
	Subscription	Subscription only + Hybrid potential (includes articles that are free at publisher sites for promotional or public-good purposes)	
	Hybrid-Total	The sum of 'Hybrid Potential' + 'Gold-Hybrid'	

POSTED VERSIONS

Label	Description	Notes
PP	Preprint	Preprints (pre-refereed versions)
AAM	Accepted Author Manuscript	Accepted versions of author manuscripts (post-refereed)
VoR	Version of Record	Final published journal articles (sometimes referred to as PJA) or Version of Record (VoR)

2. OA Options available to authors

2.1. Questions and aims

Our aim has been to build an initial evidence base on the options available to authors who wish to publish on OA terms, in order to meet funders' policy requirements, or for any other reason. This work thus seeks to elucidate aspects of the 'supply' of OA options, including:

- ❑ the numbers of fully-OA and hybrid journals
- ❑ the levels of APCs and other charges
- ❑ the length of embargo periods prescribed by publishers of different subscription-based journals
- ❑ the extent to which publishers either prescribe or allow use of the CCBY and other Creative Commons licences.

We have sought to establish a baseline of information on each of these questions, from which future exercises might identify trends. A summary of our findings is at the end of this section.

2.2 Context and previous studies

The growth of fully-OA journals can be tracked via the Directory of Open Access Journals (DOAJ) (<https://doaj.org/>), though new criteria and processes for inclusion in the Directory were introduced in 2013, which led to several hundred titles being removed. As of 30 May 2015, over 10,500 journals are listed in DOAJ³.

Bjork (2012) has tracked the growth in the numbers of hybrid journals, identifying over 4,300 in early 2012; and the number has risen very significantly since then, along with take-up of the OA option in such journals. Some hybrid titles, most notably *Nature Communications*, have made the transition to full OA⁴. Nevertheless, the hybrid model is regarded with suspicion by some commentators, including librarians; and some funders refuse as a matter of policy to meet the costs of APCs for hybrid journals.

Bjork and his colleagues have also drawn attention to the significance of 'delayed OA' journals, which make all their contents freely accessible on the publisher's platform after an embargo period. Several learned society publishers which use the HighWire platform employ this model, and some 2.5 million articles are freely available on that platform. Laakso and Bjork (2013) identified nearly 500 journals – most of high status - using this model; and calculated that 78% of the 111,000 articles they published in 2011 were freely-accessible within 12 months of publication. The terms on which articles are made accessible, however, do not typically include a Creative Commons (CC) or similar licence, and the 'delayed OA' term – though widely used (as it is in this report) - is thus somewhat problematic.

Determining means and medians for APCs for journals of different types is made more difficult by the wide range of discounts and surcharges levied by some publishers. But various studies (Inger and Gardner 2013; Bjork and Solomon 2012a, 2012b; Solomon and Bjork 2012a, 2012b; 2014; Morrison et al 2014; Kingsley 2014; Ware and Mabe 2015) have demonstrated that APCs are on average higher for hybrid than for fully-OA journals (excluding those of the latter that do not charge any APC at all); and that the APCs

³ Not all of the journals listed in DOAJ are published in English, and some publish relatively small numbers of articles

⁴ Further examples of journals making such a transition are given in Ware and Mabe *The STM Report: an overview of scientific and scholarly journal publishing*, fourth edition, 2015 pp 95-96.

for fully-OA journals published by fully-OA publishers are on average lower than for such journals published by mixed-model publishers. Such analyses have heightened the sensitivities surrounding the so-called ‘double-dipping’ issue: concerns that publishers are receiving both APCs and subscription revenues for hybrid journals. They have also helped to stimulate the search for arrangements to limit the cost of APCs (Bjork and Solomon 2014) and/or to reach arrangements under which APCs and subscription costs are negotiated and managed together (Lawson 2015). Major publishers including Wiley, Springer, and Taylor and Francis have reached different agreements with Jisc on this basis. The analysis later in this report (Section 5) on financial sustainability for universities is based on actual prices paid, and provides new insights into relationships between price and citation impact, and a richer context for some of the differences in APC levels noted above.

Alongside the growth of OA publishing, the number of open access repositories as recorded both by OpenDOAR (www.opendoar.org) and the Registry of Open Access Repositories (<http://roar.eprints.org>) has also been growing. OpenDOAR currently (May 2015) includes over 2,800 repositories, just under 2,000 of which include journal articles (though not necessarily full-text). Bjork et al (2013) and Pinfield et al (2015) have used the data from these directories and other sources to plot the pattern of development by repository type and location, as well as the underlying technical infrastructure.

Statistics from SHERPA/RoMEO (<http://www.sherpa.ac.uk/romeo/>) show that of the 1,834 publishers covered by the database (May 2015), 76% allow the deposit of some version of published articles in a website or repository of some kind; but the database does not store information about permitted locations of deposit, or embargo lengths, in a structured way. Laakso (2014) has created a manually-coded database (which he has generously made available to us) of the detail of posting policies for the hundred largest journal publishers by article output. His publisher-level analysis indicates that for just over 80% of 1.1m articles published in 2010, the policies allowed for AAMs or VoRs to be posted and made accessible up to one year after publication; and that the policies were more permissive for posting on personal websites and institutional repositories (IRs) than for subject repositories. He concluded that then-current (2013) estimates of the numbers of articles actually accessible from repositories and author websites indicated that there was considerable *potential* for deposit that was not being exploited by authors and their institutions.

2.3. Data sources and methodology

We have employed three approaches to gathering data, based on

- a) the publishing models adopted by the journals included in the SCOPUS database
- b) data from a sample of publishers⁵, including all those responsible for the journals that are most popular with UK authors in each of the subject areas covered by the four main REF panels
- c) data relating to individual journal titles which show high numbers of publications from authors with a UK affiliation

Our objective through the first approach was to gain an overall picture of the choices available to researchers both globally (covering the publishing models of the 22,000 journals in the SCOPUS database) and in the UK (covering the publishing models of the 13,500 of those journals in which UK authors

⁵ Publishers may, for the purposes of this section, cover a number of imprints. Thus data for Springer includes journals under the BioMedCentral imprint. But the merger between Springer and Nature Publishing Group occurred while we were undertaking this study, and we have not sought to amalgamate them.

published during 2014). This was based on desk research on publishers' price lists and catalogues, individual journal websites, and data from the DOAJ.

Through the second approach we sought to examine the varying profiles of different publishers in terms of publishing and business models, and their policies on matters such as licensing and the posting of different versions of articles. Here we focused on a sample of publishers, covering those responsible for journals that were among the 25 most popular with UK authors between 2010 and 2014 in each of the four broad subject areas covered by the main REF panels. These publishers include the eight largest recorded in the SCOPUS database, some with a number of imprints (such as Cell Press within the Elsevier portfolio); but also a number of smaller ones (four of them publishing only one journal each). Together they are responsible in total for some 11,500 journals and over a million articles annually. The larger publishers provided aggregate data on the numbers of journals employing different publishing models, levels of APCs, licensing, and posting policies. For the smaller publishers we were able to rely for the most part on desk research. A full list of the publishers is at Annex A. We do not name the publishers in the tables in this Section, since they provided data to us in confidence, and our purpose is in any case to identify broad patterns. Where particularly salient features of individual publishers are in the public domain, however, we draw attention to them in the text.

Through the third approach we sought in more detail to ascertain the models and policies of the journals in which UK authors most frequently publish, thus providing some insight into the choices available and the choices made. The data was gathered through detailed analysis of the information available on each journal's website. The number of UK peer-reviewed articles in the 25 journals in each of the four broad subject areas ranged from 3,290 in the arts and humanities, to 39,290 in the physical sciences; and the proportions of all UK-authored publications between 2010 and 2014 recorded in SCOPUS accounted for by the 25 most frequently-used titles ranged from 11% in the health and life sciences to 31% in the arts and humanities. A full list of the journals is at Annex B.

We should stress at this point that detailed information on matters such as the posting of different versions of articles (pre-prints, AAMs, and VoRs)) on different kinds of sites (personal websites, repositories and so on) is not always easy to find, or in some cases to interpret. In some cases we have been unable to find clear answers on specific points such as the posting of different versions of articles in different locations. Moreover, some of the models and policies adopted by publishers do not fit readily into standard categories, and while we have drawn attention to specific examples, we are aware that the numbers we present may in some instances simplify what is in fact a more complex picture. In dealing with those complexities we have identified, we have in general focused on policies as they affect UK authors, particularly in the HE sector, where there are important interfaces between the policies of journals and publishers on the one hand, and of major research funders on the other.

2.4. Choices available to authors

2.4.1 Publishing models

The publishing models adopted by the 22,000 journals in the SCOPUS database in 2014, and the 13,500 of those journals in which authors with a UK affiliation published at least one article that year are shown in Tables 1a and 1b respectively. The tables also show for comparative purposes estimates for 2012, and the change between the two, in terms of the number of *journals* (not the change in share) within each publishing model.

**Table 1a. Journal publishing models:
all journals**

		Journals		
		Global % of journals	Global % of journals	Global Growth*
	Publishing model	2012	2014	CAGR 12-14
Potential Immediate OA	Gold-APC	7.8%	8.2%	4.4%
	Gold-no APC	8.7%	8.6%	0.8%
	Hybrid-Total	45.5%	49.0%	5.5%
Subscription based	Delayed OA	2.3%	2.4%	5.4%
	Subscription only	35.6%	31.7%	-4.0%
	Immediate OA	62.1%	65.8%	4.7%
	Subscription-based	37.9%	34.2%	-3.4%
	Total No. Journals	21,741	22,486	1.7%

**Table 1b. Journal publishing models:
journals used by UK authors**

		Journals		
		UK % of journals	UK % of journals	UK Growth*
	Publishing model	2012	2014	CAGR 12-14
Potential Immediate OA	Gold-APC	7.2%	7.3%	1.4%
	Gold-no APC	6.2%	5.6%	-4.5%
	Hybrid-Total	59.9%	63.8%	3.9%
Subscription based	Delayed OA	3.5%	3.8%	4.3%
	Subscription only	23.1%	19.5%	-7.7%
	Immediate OA	73.4%	76.8%	3.0%
	Subscription-based	26.6%	23.2%	-6.0%
	Total No. Journals	13,411	13,585	0.6%

The tables show that in 2014:

- ❑ just under 17% of journals in the world, and 13% of those in which UK authors actually published, were fully-OA (Gold charging an APC, or Gold with no APC)
- ❑ 49% of journals in the world, and 64% of those in which UK authors published, followed the hybrid model (though authors did not necessarily, of course, take up the OA option)
- ❑ ‘delayed OA’ journals (providing access on the publisher’s site after an embargo period) constituted over 2% of journals globally, and nearly 4% of those in which UK authors published
- ❑ in total, just under two-thirds of journals globally provided an OA option, and just over one-third were subscription-only. But more than three-quarters of the journals used by UK authors offered an OA option, with less than a quarter operating subscription-only
- ❑ subscription-only journals fell between 2012 and 2014 both in numbers and as a proportion of all journals, with the fall particularly marked among those in which UK authors publish

The data from our sample of individual publishers enables us to examine their varying profiles in terms of their adoption of different publishing models. The sample includes publishers responsible for around half the journals included in the SCOPUS database; but it is not, of course, a random sample, since it focuses on the publishers (not the journals) with which UK authors tend to publish their articles. It is therefore not surprising that the overall proportions of fully-OA, hybrid and subscription-only journals are somewhat different from those shown in Table 1a above (though they are not markedly different from the totals for journals used by UK authors shown in Table 1b).

The notable finding illustrated in Table 2, however, is the extent of variation between publishers in their adoption of different publishing models. There are significant differences between the four largest

publishers; and while there is a significant group of medium-sized and small publishers all of whose journals are hybrid, there are others showing much lower rates of adoption of that model. And a smaller group of publishers, two in the humanities and one in medicine and life sciences, publish all their journals as subscription-only.

Table 2. Publishing models adopted by 32 publishers with whom UK authors publish articles

Publisher	All journals	subscription only	%	Fully OA	%	Hybrid	%
1	2680	556	21%	516	19%	1608	60%
2	2554	524	21%	392	15%	1638	64%
3	2026	100	5%	42	2%	1866	92%
4	1647	340	21%	44	3%	1257	76%
5	826	45	5%	36	4%	745	90%
6	356	114	32%	12	3%	230	65%
7	338	42	12%	29	9%	267	79%
8	300	0	0%	0	0%	300	100%
9	230	15	7%	3	1%	212	92%
10	113	34	30%	36	32%	43	38%
11	70	45	64%	20	29%	5	7%
12	68	20	29%	4	6%	44	65%
13	54	4	7%	10	19%	40	74%
14	47	0	0%	0	0%	47	100%
15	40	0	0%	0	0%	40	100%
16	39	0	0%	0	0%	39	100%
17	31	31	100%	0	0%	0	0%
18	18	8	44%	2	11%	8	44%
19	18	9	50%	5	28%	4	22%
20	11	0	0%	3	27%	8	73%
21	11	0	0%	2	18%	9	82%
22	10	0	0%	0	0%	10	100%
23	7	0	0%	7	100%	0	0%
24	5	5	100%	0	0%	0	0%
25	5	1	20%	1	20%	3	60%
26	4	3	75%	1	25%	0	0%
27	3	0	0%	0	0%	3	100%
28	2	0	0%	1	50%	1	50%
29	1	1	100%	0	0%	0	0%
30	1	0	0%	0	0%	1	100%
31	1	0	0%	0	0%	1	100%
32	1	0	0%	0	0%	1	100%
Total	11517	1897	16%	1166	10%	8430	73%

Data relating to the 25 journals which had most articles from UK authors between 2010 and 2014 in each of our four subject areas provides further evidence of the availability of OA options, and to some extent of take-up. We recognise the limitations of the small sample; but Table 3 shows some notable differences between our four subject areas. Thus the availability and take-up of fully-OA journals varies from seven in the health and life sciences and three in the physical sciences and engineering, to none in the social sciences or the arts and humanities. Conversely, numbers of hybrid journals ranged from 13 in health and life sciences to 20 or 22 in the other three subject areas; and the number of subscription-only journals from two in the physical sciences and engineering to four in the health and life sciences. Thus the health and life sciences show both the highest number of both fully-OA and of subscription-only journals in our sample⁶.

The census of journals and articles presented in Section 3 shows that UK authors tend in general to publish in journals where the citation impact (as measured by the Field Weighted Citation Index (FWCI); see definition in Annex E) is higher than the average for their field. Nevertheless, it is notable that data we present in Annex B also suggests that in the humanities and social sciences (HSS), many authors are willing to publish in journals where the FWCI is less than the average for their field. The subscription-only journals that are popular with UK authors include high-status titles such as *Nature* but also journals that are important in a wide range of subjects and fields such as *Blood*, the *British Journal of Nursing*, the *Astrophysical Journal*, the *Political Quarterly* and the *Classical Quarterly*.

Delayed free access on the publisher's platform is a common feature of our sample of journals in STEM areas, but not in HSS. In the health and life sciences, 12 of the journals that are not fully-OA make all articles freely accessible after a delay, usually of 12 months, but in two cases after six months. In the physical sciences and engineering the delays are of 12 months for most of the nine journals that operate this system, but for one it is 6 months, and for another 36 months⁷. In the HSS disciplines, the only journal that provides delayed OA is the *Notes and Records of the Royal Society*, with a delay of 24 months.

Table 3. Publishing models for 25 journals in each of four subject areas

Subject area	Fully Gold Journals	hybrid journals	Subscription journals	Delayed OA
Health and life sciences	7	13	5	12
Physical sciences and engineering	3	20	2	8
Social sciences	0	22	3	0
Arts and humanities	0	22	3	1

2.4.2 APCs

Data on levels of APCs (excluding VAT) from the twenty of our sample of publishers who publish at least one fully-OA journal and the 28 who publish hybrid journals is given in Annex C. In sum, they show that 59% of the fully-OA journals charge an APC of £1,000 or less; and indeed the majority of those charge no

⁶ The number of publishers responsible for the 25 journals also differed across the four fields, from seventeen in the health and life sciences to seven in the social sciences.

⁷ For one publisher, the American Chemical Society, delayed OA after 12 months is in effect a cheaper option for immediate OA, since a fee is charged for this option. In all other cases, the access requires no payment.

APC at all⁸. A further 38% charge an APC of between £1,001 and £2,000, the great majority of those at £1,500 or less. Only a small minority of fully-OA journals, concentrated in one publisher (Nature Publishing Group) charge APCs of more than £2,000⁹. At the lower level of up to £2,000, however, again there is no obvious pattern in the banding of APCs by price across our sample of publishers.

The pattern for hybrid journals is significantly different. By far the majority (88%) charge between £1,000 and £2,000, with much smaller minorities charging either £1,000 or less, or more than £2,000 (5%).

Among publishers with significant numbers of journals, only the American Chemical Society and Nature Publishing Group charge APCs of £2,000 or more for the majority of their journals. The pattern across our sample of publishers is not immediately obvious; but the analysis we present in Section 5 suggests that there is some relationship between the APCs actually paid by UK universities, and the citation profile (as indicated by the FWCI) of the relevant journals.

APC levels for our sample of journals in which UK authors most frequently publish varied, as shown in Table 4, from under £300 to £3,333, with means and medians between c£1500 and c£1800; and there is no obvious pattern in either the averages or the range of published levels of APCs in the different subject groups. But with all the data on APCs, it is important to stress that published APC levels may not reflect what is actually paid. Thus Table 4 indicates that some of the journals come from publishers that offer a membership or similar scheme providing discounts on APCs, and that these are common in the HSS subject areas, although less so in STEM. This reflects in part the smaller numbers of publishers for the most popular journals in the HSS as compared with STEM disciplines; and perhaps also a higher preference of UK HSS authors to publish in UK-based journals, where publishers have felt more pressure than in the US to introduce such schemes. Some publishers, including the Institute of Physics, Wiley, and Taylor and Francis, have also reached agreements with Jisc under which a proportion of the costs to institutions of APCs is offset against the cost of subscriptions, or vice versa. The agreements take a number of forms, and it is too early to assess their sustainability or their overall impact.

Table 4. APCs for 25 journals in each of four subject areas

Subject area	Gold and hybrid jnls	APC mean (£s)	APC Median	APC range (£s)	Membership scheme	Offset agreement
Health and life sciences	20	1466	1687	850-3333	7	1
Physical sciences and engineering	23	1466	1451	286-3150	8	4
Social sciences	22	1750	1578	800-2333	15	12
Arts and humanities	22	1788	1653	800-2000	14	13

⁸ It is important to distinguish between the numbers and proportions of fully-OA *journals* that do or do not charge an APC, and of *articles* published in such journals. Data from DOAJ suggests that two-thirds of its *journals* do not charge an APC; but two-thirds of the *articles* are published in journals that do charge an APC

⁹ All the figures shown here exclude VAT, which adds 20% to the charges for APCs in the UK. US \$ have been converted to UK £ at an exchange rate of 0.65

For these and other reasons, the figures in the tables do not fully reflect a highly-complex picture. The discounts offered under membership schemes can be significant: up to 25% or even 30% with publishers including Wiley, or Taylor and Francis, and also with smaller publishers such as the Royal Society. The American Chemical Society's discounts for members of the society, and for authors at institutions with a subscription to its full range of journals, are also substantial. (Conversely, it imposes surcharges on authors wishing to use a CC licence.) For UK authors and institutions, the discounts under the Jisc-negotiated offset agreements are likely to have an even bigger impact on the effective level of APC prices. But the extent to which authors are aware, before they decide to publish OA, either of the publicly-stated APC or of such discounts and offset arrangements, is unclear, though there is anecdotal evidence that the Royal Society of Chemistry's Gold for Gold scheme has achieved resonance in the UK chemistry community. The key point to stress, however, is that the figures summarised in Table 4 relate to the core published levels of APCs, and do not take account of a highly-complex range of discounts and surcharges (for the use of CC licences, for example), or the levying of submission or publication fees, or page or colour charges, as set out in Section 2.4.3. Data on the APCs actually paid by UK universities are presented in Section 5.

2.4.3 Other author charges

Some journals levy other charges, irrespective of and unrelated to OA, in the form of submission or publication fees, page charges and charges for the use of colour in figures in the printed versions of articles. In some cases, where an OA option is available and taken up, these other charges can exceed the level of the APC. We have gathered data on these for our sample of journals in which UK authors most frequently publish, and the results are shown in Table 7. This shows (again recognising the limitations of the sample) that levying publication and page charges is a feature of a fifth of the top 25 popular titles in the health and life sciences, and a quarter of those in the physical sciences and engineering; and charges for colour printing are imposed in just under half of the titles in the physical sciences. By contrast, publication and page charges do not feature at all in titles in the HSS disciplines; but charges for colour printing are levied by a substantial proportion of titles in those disciplines, all published by Taylor and Francis/Routledge, which charges £250 per figure for print.

Submission fees are relatively rare, and in our sample were levied by only three titles, all in the health and life sciences. They range from \$50 in *Blood* and the *Journal of Immunology* (which waives it for members of the American Association of Immunologists) to \$130 for the *Journal of Neuroscience* (which levies a publication charge also). Publication fees range from \$500 to \$1720, but in the case of the American Physical Society's journals they are levied only if manuscripts are submitted not using one of the publisher's prescribed formats.

Page charges are complex too. Journals published by some American learned societies (the majority of journals levying such charges in our sample) levy charges on every page, at levels that may increase once a page or word-limit has been reached¹⁰. Conversely, charges may be reduced for society members (for example the American Society for Biochemistry and Molecular Biology's *Journal of Biological Chemistry*). Nevertheless, the effective cost of publishing a paper in a journal that levies page charges may be as much as a third or more higher than the cost of the APC. And charges for colour figures may have an even larger impact; inclusion of just three colour figures in the American Physical Society's *Physical*

¹⁰ The two UK-based journals in our sample that levy page charges (*Proceedings of the Royal Society B: Biological Sciences*, and *Nucleic Acids Research* (OUP) do so only once a page limit has been exceeded.

Review B would imply a charge higher than the APC of \$1700; and the charges could be avoided only if authors agreed to forgo the use of colour in the printed version of their article.

Table 5. Publication, page and colour charges for 25 journals in each of four subject areas

Subject area	Submission and/or publication fee	range (£s)	Page charges	range (£s)	Colour charges	range per figure (£s)
Health and life sciences	5	33-1147	5	41-200	5	33-413
Physical sciences and engineering	7*	333-1000	6	20-200	12	83-633
Social sciences	0	n/a	0	n/a	3	250
Arts and humanities	0	n/a	0	n/a	9	250

2.4.4 Licences

Data on the availability of the CCBY licence from the twenty of our sample of publishers who publish at least one fully-OA journal and the 28 who publish hybrid journals is given in Annex C. In sum, they show that over 99% of the fully-OA journals published by our sample of publishers either use the CC-BY licence as the default, or provide it as an option for authors, with a majority adopting it as the default or requirement. And data from the DOAJ and from the Open Access Scholarly Publishers Association (Redhead, 2015) indicates that use of the CCBY licence in fully-OA journals published by the Association's members has risen sharply since 2010, although publishers including BMJ and NPG still make extensive use of other CC licences which preclude commercial use (NC) or the creation of derivatives (ND). There is considerable variation among publishers as to whether they simply allow CCBY or another CC licence as an option as distinct from using it as a default.

The pattern for hybrid journals is rather different. Again most journals allow for the use of a CCBY licence, in some cases (as with *The BMJ*) only where funders require its use; and in others, (as with the journals of the American Chemical Society) if authors pay a higher level of APC. Around 20% of journals use CCBY as the default. Once more, there is no obvious pattern in the variations in policy between publishers.

When we look at licence terms for our sample of OA and hybrid journals in which UK authors most frequently publish, some interesting patterns emerge. Table 6 indicates that the CCBY licence is widely on offer in the HSS disciplines, although its use is particularly controversial among HSS researchers as compared with those in some STEM subjects. This does not, of course, imply that the CCBY licence is widely taken up by HSS authors who opt for OA publication, though those who publish with journals that require the use of CCBY have no choice. And such journals include those published by both fully-OA publishers such as PLOS and BioMedCentral and also publishers of hybrid journals, such as Springer and

the Portland Press; while others such as BMJ group specifically allow use of CCBY when funders require it. Conversely a few publishers, including the American Chemical Society, charge extra for authors who wish to use a CCBY licence; and for several titles, especially for conference proceedings, it is wholly unclear, even when OA publishing is available, whether the CCBY licence is even on offer.

Table 6. Licensing for fully-OA and hybrid journals in each of four subject areas

Subject area	Fully Gold Journals	hybrid journals	CCBY licence offered for Gold
Health and life sciences	7	13	17
Physical sciences and engineering	3	20	17
Social sciences	0	22	22
Arts and humanities	0	22	22

2.4.5 Volumes of OA articles

Table 7 shows the numbers of OA articles published in 2014 in the fully-OA and hybrid journals of a subset of our sample of publishers who were able to provide data of this kind. Together these publishers were responsible for over 1 million articles, over 12% of which were published in fully-OA journals, and 2% on OA terms in hybrid journals. But the ratios varied widely, even if we exclude the one fully-OA publisher (PLOS). Some of the differences are explicable in terms of publishers' varying subject focus: the low ratios of some publishers reflect the relatively low rates of take-up of OA in the HSS subjects. But other differences seem to reflect the particular policies and strategies of individual publishers: thus the relatively high rates of some publishers may be explained by the relatively early steps they took to launch OA journals, and to shift from subscription-only to hybrid.

Table 7. Volumes of OA articles for selected publishers in 2014

Publisher	articles in fully-OA journals	CCBY	OA articles in Hybrid journals	CCBY	Subscription-only articles 2014	All articles	All OA articles as % of all articles published
1	0	0	866	425	40194	41060	2%
2	1601	112	677	64	10634	12912	18%
3	162	158	328	300	32684	33174	1%
4	34300	Not known	8900	Not known	365700	408900	11%
5	0	0	27	27	9000	9027	0%
6	8406	8406	807	807	22303	31516	29%
7	5239	1187	1259	621	5515	12013	54%
8	3191	1768	2013	1225	35427	40631	13%
9	30817	30817	0	0	0	30817	100%
10	118	118	419	419	1910	2447	22%
11	1170	444	1170	0	58000	60340	4%
12	40466	40263	6507	6152	210119	257092	18%
13	1037	940	1063	290	74093	76193	3%
Totals	126,507	84,213	24,036	10,330	865,579	1,016,122	15%

2.4.6 Posting policies and embargoes

Policies relating to the posting of articles on author websites, institutional and subject repositories, and other locations vary widely across publishers and individual journals, according to the version to be deposited, the location of deposit, and embargo periods. The pattern of variation is complex, and in a number of cases the policies are not clear, even after diligent search. But in general terms, policies are more permissive for pre-prints and AAMs than for VoRs and there is a similar gradation in moving from postings on author websites through institutional and subject repositories to sites that are seen as commercial operations, such as the Social Science Research Network (SSRN).

Preprints

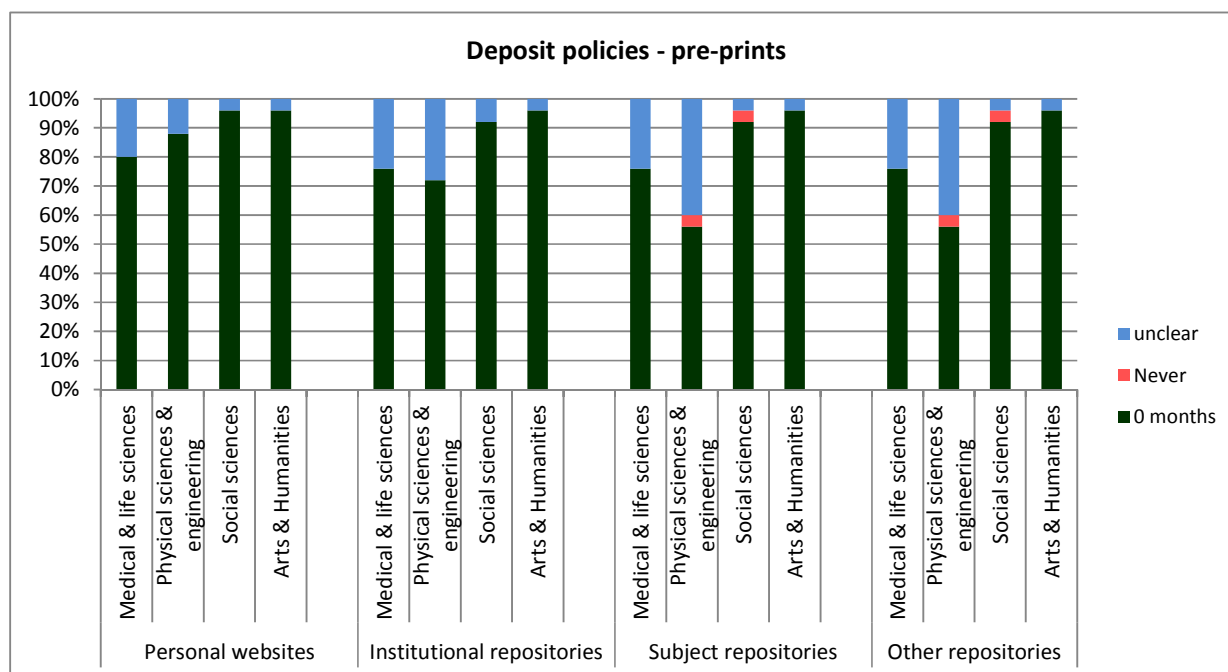
The great majority of journals from our sample publishers allow the posting of pre-prints, especially on personal websites: even allowing for those where policies are not clear, nearly 90% allow posting.

Explicit policies about the posting of pre-prints can be found for most but not all of our sample of journals in which UK authors frequently publish; and where policies are in place, they allow for posting at least on personal and institutional websites (though in many cases it is unclear whether posting is allowed elsewhere). Several journals, including the American Society of Hematology's *Blood* and the journals published by the American Chemical Society point to the risks of prior publication, particularly if a paper is posted in commercially-funded repositories and services¹¹; a small number of journals explicitly prohibit posting in such repositories. The American Geophysical Union's *Geophysical Research Letters*, for

¹¹ The *Modern Law Review* requires that all versions of a submitted paper, and related papers, are removed from the internet while the submitted version is under review, in order 'to preserve the anonymity of the refereeing process'.

example, in effect advises authors to remove preprints from publicly-available servers once a paper has been accepted for publication. More usually, journals require that once a paper has been accepted, any posted pre-print should be accompanied by a statement that it has been so accepted; and once it has been published by a citation and link to the published version (sometimes but not always through a DOI). The format and wording of these statements is usually prescribed, though the wording and formats vary; and some publishers, including Wiley and Elsevier do not prescribe any particular format.

Figure 1. Pre-print deposit policies for journals in each of four subject areas



Author accepted manuscripts (AAMs)

For AAMs, the position is more complicated. Laakso (2014) reported that over 80% of journals in 2010 allowed authors to post AAMs on personal websites or IRs, and that 67% allowed access from a personal website with an embargo between zero and six months. Embargo periods rose for IRs (only 51% allowed an embargo between zero and six months); and policies were much more restrictive when it came to posting in subject repositories or other sites. Findings from the data we have gathered from our sample of publishers are shown in Annex D. They are in line with Laakso's, indicating that 77% of the journals from our sample publishers (excluding fully-OA journals) allow posting and access via personal websites with an embargo between zero and six months, but only 18% allow posting and similar access via an IR, and 7% via a subject repository. A further 61% allow access via an IR or subject repository after 12 months, but for over a fifth of the journals, the embargo on access via subject repositories lasts for up to 24 months, and some journals do not allow posting at all.

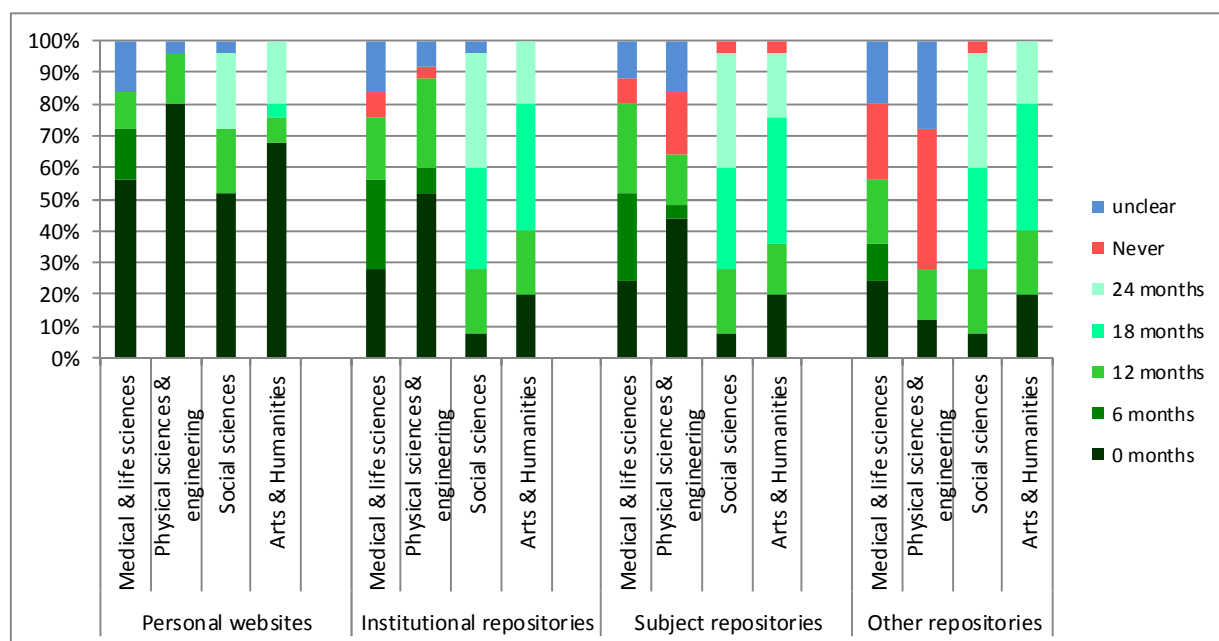
Several caveats should be noted in relation to these figures. First, policies for a minority of journals are not wholly clear; and some larger publishers who publish on behalf of societies record policies in some cases simply as 'non-standard'. Second, some journals and publishers – including, for example, BMJ, Elsevier and the Portland Press - modify their policies in response to funder mandates both from the US

(particularly the NIH) and from the UK. Nevertheless, there are significant differences in the pattern of policies across different publishers. Most are relatively liberal in relation to embargoes for articles posted on personal websites, but publishers including Wiley and Oxford University Press (which forbids authors to replace a pre-print with an AAM) are more restrictive. Similarly, while many publishers lengthen embargoes for articles posted in IRs as distinct from personal websites, others, including SAGE, Emerald, and BMJ do not, lengthening them only in relation to subject or other repositories.

When we examine the policies for our sample of journals in which UK authors frequently publish, we find the same tightening of policies - particularly in the STEM subjects - as we move from personal websites and IRs to subject repositories and other sites. As Figure 2 indicates, the published policies allow for posting at least on personal websites and IRs (with one exception for IRs, where the American Geophysical Union's *Geophysical Research Letters* specifies the posting only of the VoR); and typically some kind of statement about the nature and location of the published version is required. But in the two STEM groups of disciplines, 17 journals explicitly prohibit posting in repositories other than IRs or not-for-profit subject repositories.

Embargo periods also tend to rise as we move from personal websites to repositories, and from the STEM disciplines to HSS. These findings are in line with Laakso (2014) who reported that 78% of journals in the physical sciences, but only 45% of those in the social sciences allowed deposit in at least one location with no embargo; and that while 78% of journals allowed posting on a personal website, only 33% allowed posting in a subject repository. In our much smaller sample, a majority of journals, with the exception of those in the social sciences, impose no embargoes on AAMs posted on personal websites. But for postings in repositories, embargo periods rise dramatically, particularly in the HSS subjects, where the typical periods are between 18 and 24 months. Nevertheless, just over half the journals in our sample in the physical sciences and engineering allow postings in IRs with no embargo; and across the STEM disciplines the embargoes do not exceed 12 months. In a significant number of cases, however, these embargo periods relate specifically to articles by authors in UK universities, since the publishers have taken account of the policies of the UK Funding Councils, the Research Councils and the Wellcome Trust.

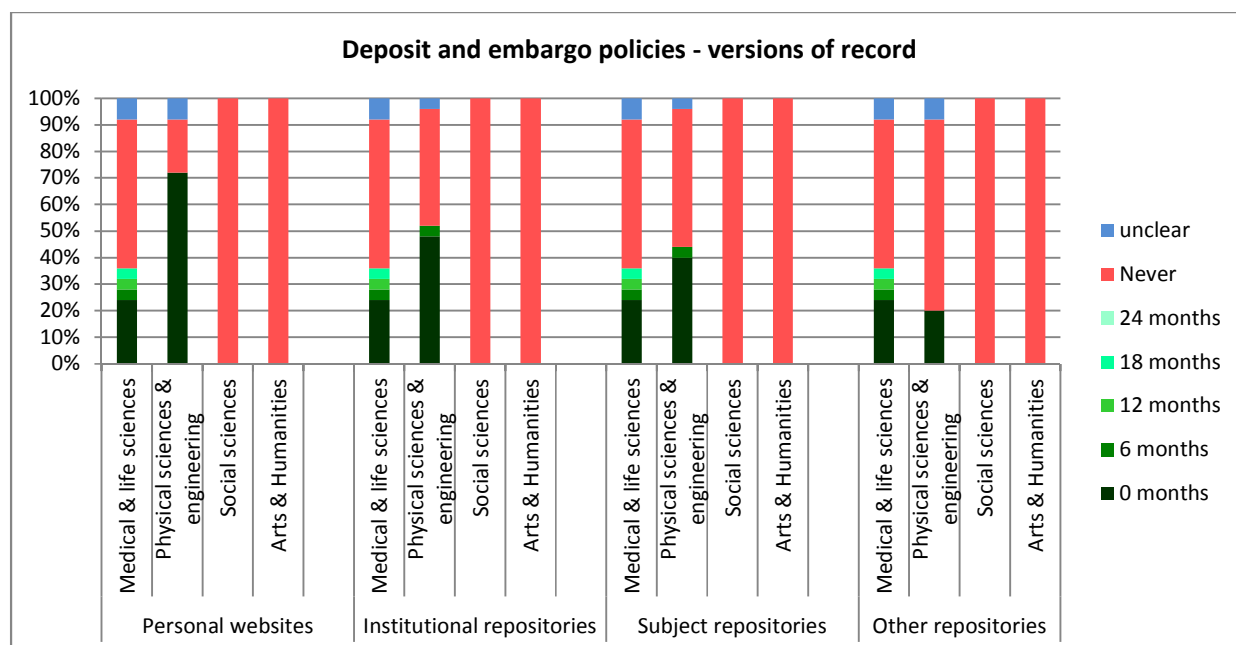
Figure 2. AAM deposit and embargo policies for journals in each of four subject areas



Versions of record (VoRs)

Only a small minority of journals in our sample allow posting of VoRs in any location. Most of those that do are in the physical sciences, and include journals published by the American Physical Society and the American Institute of Physics, which allow posting on personal and institutional websites, along with a statement including a citation and links to the version on the publisher's site. Data from the journals in which UK authors most frequently publish indicates – as shown in Figure 3 – that as we proceed through IRs and subject repositories to other locations, a few journals impose embargo periods, but an increasing number prohibit any posting at all. Relatively small numbers of journals in the health and life sciences also allow posting, and they do not appear to distinguish in their policies according to the location of the posted version. But in HSS subjects, no journals in our sample allow VoRs to be posted in any location.

Figure 3. VoR deposit and embargo policies for journals in each of four subject areas



Summary of findings

- ❑ OA publishing options are now widely available: two-thirds of the journals published globally, and more than three-quarters of the journals in which UK authors publish, offer an OA option of some kind.
- ❑ By far the largest group of journals have adopted the hybrid publishing model: just under half of all journals across the world operate in this way, and nearly two-thirds of those in which UK authors publish their work. Hybrid journals are particularly common among the journals used by UK authors in the HSS subject areas, where take-up of fully-OA journals is low.
- ❑ Fully-OA journals which charge an APC feature strongly among the journals in which UK authors publish in the STEM subject areas; but UK take-up of fully-OA journals with no APC is falling.
- ❑ Both the numbers and proportions of subscription-only journals have fallen in the past two years; and the fall has been particularly sharp among the journals in which UK authors publish.
- ❑ The publishers with whom UK authors most commonly publish show markedly varying profiles in their adoption of fully-OA, hybrid and subscription-only publishing models.

APCs and other charges

- ❑ Only a small minority of fully-OA journals charge APCs of more than £2,000; but below that level there is wide variation between both journals and publishers. The great majority of hybrid journals charge APCs between £1,000 and £2,000; only small minorities, concentrated in a few publishers, charge either less than £1,000 or more than £2,000.
- ❑ Among the journals in which UK authors most frequently publish, there are again wide variations in levels of APCs; but there is relatively little variation across subject areas.
- ❑ Submission or publication fees, and page and colour charges, can add significantly to authors' costs for publishing, particularly with American society publishers.

Licensing

Creative Commons licences have become widely, but not universally, accepted as a mechanism for promoting OA. Most journals - in social sciences and humanities (HSS) as well as in STEM subject areas - *allow* publishing under a Creative Commons CCBY licence, but there is widespread variation as to whether it is a default, or an option; and as to whether the CCBY licence or the more restrictive licences that preclude commercial use (CCBYNC) and/ or the creation of derivatives (CCBYND) are employed. Some journals allow Creative Commons licences, particularly CCBY, only when it is a funder requirement.

Posting policies and embargoes

- ❑ Posting policies are more permissive for pre-prints and for AAMs than for VoRs; and there is a similar gradation in moving from postings on author websites through IRs and subject repositories to sites seen as commercial operations. But the detail of policies is sometimes hard to find, and/or to interpret; and there are considerable differences between publishers. Policies are often modified in response to funder requirements.
- ❑ The great majority of journals allow the posting of pre-prints, although many require that once a paper has been accepted for publication, any pre-print should be accompanied by a statement that it has been accepted, and a citation link.
- ❑ Nearly four-fifths of journals allow the posting of AAMs on personal websites with embargoes between zero and six months. But policies are progressively more restrictive for posting in IRs, subject repositories or other services.
- ❑ Only a small minority of journals – mostly in the physical sciences – allow VoRs to be posted in any location.

3. Accessibility: authors' take-up of OA options

3.0 Research Objectives

Our aim in this part of the study was to determine the numbers - and the proportions of the overall population – of articles published and openly accessible:

(1) **Via different publishing models**

- ❑ in fully-OA journals that charge an APC;
- ❑ in fully-OA journals that do not charge an APC;
- ❑ on OA terms in hybrid journals;
- ❑ in subscription-based journals that nevertheless provide access free of charge after an embargo period where that period is no more than 24 months.

(2) **Posted versions** made accessible on authors' web pages, IRs, subject repositories, or other locations.

- ❑ 'pre-prints' which show no sign of peer review (PP);
- ❑ accepted author manuscripts (AAM);
- ❑ the published version of record (VoR), also known as 'published journal article' (PJA).

(3) To assess the extent to which posted versions comply or do not comply with the stated policies of journals.

We also sought to compare UK and world take-up, as well as subject areas as defined by the four main REF panels.

3.1 Terminology

Terminology relating to OA and publishing options can be ambiguous. To minimise confusion the labels used here to describe various publishing options (and their groupings) and versions of articles posted are those set out in Section 1.3.

3.2 Methodology

Two parallel approaches were used:

- ❑ A **census** of more than 22,000 journals covered in Scopus to determine the volume of articles published under each publishing model
- ❑ A **sample-based** approach to estimate the level of postings as well as *what* (PP, AAM, VoR) was posted, *where* it was posted (personal website, institutional repository, subject repository, etc.) and *when* (after publication) it was posted and accessible. The same factors were then compared with journal-level policies, where known, to estimate the proportion of postings that did or did not comply with journal policies. The same sample-based approach was also used to assess uptake of the hybrid option (Gold-Hybrid).

A brief overview of the methodology is given below (full details are in Annex E). It should also be noted that the figures we present in this section represent accessibility in the developed world. We do not deal in this study with the large corpus of literature that is freely-accessible to users in developing countries via Research4Life, INASP, EIFL and similar programmes (see Meadows, 2015 for further information about these initiatives).

Census:

- Data sourced from 22k+ journals covered in Scopus
- Only peer-reviewed document types counted
- Publishing models assigned based on DOAJ and extensive desk research
- Scopus journal-level classifications further grouped in the four REF panels (some in more than one panel)
- Measurements at Global and UK levels:
 - (1) number of articles published
 - (2) field-weighted citation impact (FWCI)

Sample:

- Random samples of articles from each of four periods post-publication (1-2, 7-8, 13-14, 25-26 months) to determine **when** a posted article was accessible
- Global sample of over 9400; UK sample of over 5100
- Oversampled REF panels C & D to get more robust breakdowns (totals across all panels were weighted to account for oversampling)
- Google searches and classification of genuine "hits" in terms of **what** (PP, AAM, VoR) as well as **where** article found

	Publishing Model	Census	Sample		Postings	Sample	
Immediate OA	Gold - APC	☑		}	PP	☑	}
	Gold - no APC	☑			AAM	☑	
	Gold - Hybrid		☑		VoR	☑	
Subscription-based	Hybrid Potential	☑		}	PP	☑	}
	Delayed OA	☑			AAM	☑	
	Subscription only	☑			VoR	☑	

Over-sampling for REF panel breakout

REF Panel	Actual distribution		Sample
	World	UK	
A	46%	49%	31%
B	42%	32%	38%
C	8%	14%	16%
D	3%	6%	15%

3.3 Publishing models

3.3.1. Publishing models – choices available to authors

As part of our study of accessibility, we developed a large-scale view on the publishing model choices available to authors. We discuss the results, and other matters relating to publishing choices, in Section 2. But the results (Tables 8a and 8b) and key summary points are repeated here to provide context for our discussion of the actual uptake of publishing models.

Table 8a. Journal publishing models: All journals

	Publishing model	Journals		
		Global % of journals	Global % of journals	Global Growth*
Potential Immediate OA	Gold-APC	7.8%	8.2%	4.4%
	Gold-no APC	8.7%	8.6%	0.8%
	Hybrid-Total	45.5%	49.0%	5.5%
Subscription based	Delayed OA	2.3%	2.4%	5.4%
	Subscription only	35.6%	31.7%	-4.0%
	Immediate OA	62.1%	65.8%	4.7%
	Subscription-based	37.9%	34.2%	-3.4%
	Total No. Journals	21,741	22,486	1.7%

Table 8b. Journal publishing models: Journals used by UK authors

	Publishing model	Journals		
		UK % of journals	UK % of journals	UK Growth*
Potential Immediate OA	Gold-APC	7.2%	7.3%	1.4%
	Gold-no APC	6.2%	5.6%	-4.5%
	Hybrid-Total	59.9%	63.8%	3.9%
Subscription based	Delayed OA	3.5%	3.8%	4.3%
	Subscription only	23.1%	19.5%	-7.7%
	Immediate OA	73.4%	76.8%	3.0%
	Subscription-based	26.6%	23.2%	-6.0%
	Total No. Journals	13,411	13,585	0.6%

*Growth 2012-2014 is for the number of journals (**not** change in share) within each publishing model. The journal counts are of those covered in Scopus and classified into publishing models as described in Appendix A2.1.1

Tables 8a and 8b show the proportions of *journals* offering different choices to authors. The proportions of *articles* published in each category of journal are of course different, since journals differ in size. But these tables show that

- ❑ Of the more than 22,000 peer-reviewed journals globally (Table 8a), UK authors published in a little over 13,500 of them in 2014 (Table 8b).
- ❑ In 2014, 66% of journals in the world (77% of journals in which UK authors published) had an immediate OA publishing model (Gold-APC or Gold-no APC) or offered an option for immediate OA (Gold-Hybrid).
- ❑ The vast majority of this immediate OA choice was accounted for by hybrid journals (49% globally and 64% for the UK). This, of course, represents a potential and not actual take up.

3.3.2. *Publishing models – uptake of models*

Our results for the take-up of the options available to authors (thus the number of articles published in the different categories of journals) are shown in Tables 9a (for the world) and 9b (for the UK). They show that:

- ❑ Numbers of research articles published grew at around 3.5% a year globally and at around 2.9% in the UK between 2012 and 2014. This level of growth is consistent with longer-term trends observed by others (Bornmann and Mutz, 2015; Mabe and Amin, 2001).
- ❑ Immediate OA models accounted for about 17% of global output of articles and 18% of UK articles in 2014. Of these:
 - Gold-APC accounted for a little under 10% of articles globally and a little over 9% of UK articles.
 - Gold-Hybrid (where an option to pay an APC was taken up for immediate OA) accounted for 2.4% of world articles but a much higher proportion (6.5%) of UK articles.
 - Gold-no APC journals accounted for a little under 5% of articles globally but only 2% of UK articles and, unlike APC and hybrid models, uptake is flat or declining.
- ❑ These results are broadly in line with previously-published estimates (Laakso and Bjork, 2012) which found that for articles published in 2011 and indexed in Scopus 11.0% were published in ‘full immediate OA journals’ (thus the Gold-APC + Gold-no APC models in this study, which sum to 12.8% globally in 2012), 0.7% as Gold -Hybrid (versus 0.8% for 2012 in the present study) and 5.2% in delayed OA journals (5.3% for 2012 in the present study). This high degree of agreement with independently-derived estimates from the same base data source for an earlier period suggests that our figures are robust.
- ❑ Over 5% (world) and over 11% (UK) of all published content from subscription-based models was freely-available at publisher sites in delayed OA journals after an embargo, more than three-quarters of them within 12 months. A further 3% were also available freely on publishers sites, we believe for promotional or public service purposes (e.g. several publishers opened their archives to content relevant to the recent Ebola crisis).

Table 9a Global potential and actual uptake of publishing models (% of articles published)

POTENTIAL ***		Articles			ACTUAL		Articles			FWCI**
		Global % of articles	Global % of articles	Global Growth*			Global % articles	Global % articles	Global Growth*	
	Publishing model	2012	2014	CAGR 12-14	Publishing model	2012*	2014	CAGR 12-14		
Potential Immediate OA	Gold-APC	7.9%	9.6%	13.9%	Gold-APC	7.9%	9.6%	13.9%	0.90	
	Gold-no APC	4.9%	4.6%	0.4%	Gold-no APC	4.9%	4.6%	0.4%	0.56	
	Hybrid-Total	45.7%	52.3%	10.8%	Gold-Hybrid	0.8%	2.4%	78.9%	1.32	
Subscription based	Delayed OA	5.3%	5.4%	4.7%	Delayed OA	5.3%	5.4%	4.7%	1.89	
	Subscription only	36.2%	28.1%	-8.9%	Subscription	81.1%	78.0%	1.6%	1.03	
	Immediate OA	58.5%	66.6%	10.4%	Immediate OA	13.6%	16.6%	14.2%	0.82	
	Subscription-based	41.5%	33.4%	-7.1%	Subscription-based	86.4%	83.4%	1.7%	1.07	
	Total No. Articles	2,351,119	2,519,824	3.5%	Total No. Articles	2,351,119	2,519,824	3.5%		

Table 9b UK potential and actual uptake of publishing models (% of articles published)

POTENTIAL ***		Articles			ACTUAL		Articles			FWCI**
		UK % of articles	UK % of articles	UK Growth*			UK % articles	UK % articles	UK Growth*	
	Publishing model	2012	2014	CAGR 12-14	Publishing model	2012*	2014	CAGR 12-14		
Potential Immediate OA	Gold-APC	7.4%	9.3%	15.4%	Gold-APC	7.4%	9.3%	15.4%	1.58	
	Gold-no APC	2.4%	2.1%	-3.1%	Gold-no APC	2.4%	2.1%	-3.1%	0.96	
	Hybrid-Total	64.0%	67.4%	5.6%	Gold-Hybrid	2.7%	6.5%	61.7%	1.65	
Subscription based	Delayed OA	11.0%	11.2%	3.5%	Delayed OA	11.0%	11.2%	3.5%	2.37	
	Subscription only	15.1%	9.9%	-16.6%	Subscription	76.4%	70.7%	-1.0%	1.57	
	Immediate OA	73.9%	78.9%	6.4%	Immediate OA	12.6%	18.0%	23.8%	1.50	
	Subscription-based	26.1%	21.1%	-7.6%	Subscription-based	87.4%	81.8%	-0.4%	1.66	
	Total No. Articles	148,466	157,240	2.9%	Total No. Articles	148,466	157,240	2.9%		

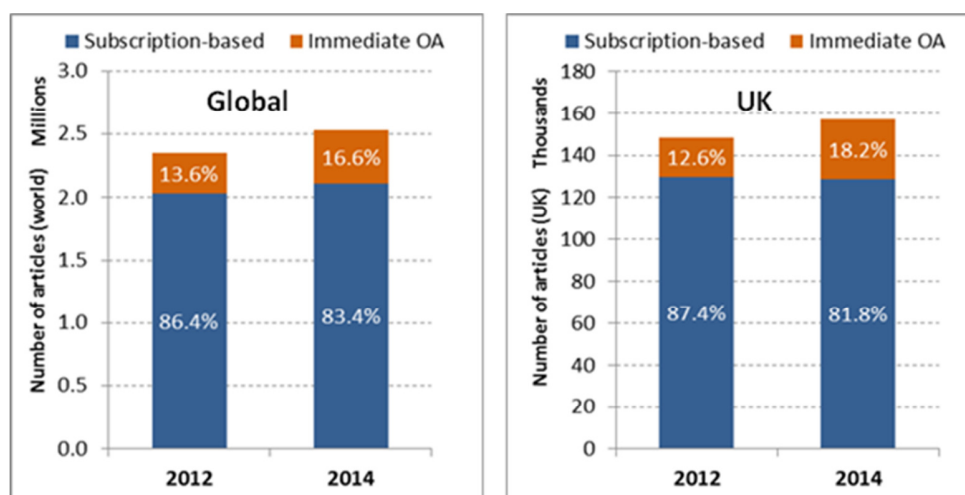
* Growth 2012-2014 is for number of articles and for most publishing models based on applying current journal publishing model to articles published in 2012. The 2012 value for actual hybrid uptake (Gold-Hybrid) was based on an earlier study for BIS: "International comparative performance of the UK Research base - 2013"

** FWCI = Field-weighted Citation Index

***The potential represents 100% take up of hybrid immediate-OA option. UK potential is based on journals in which UK authors have published. Of course, technically the UK potential is the same as that globally. However, UK authors may not publish in some journals for reasons of subject scope, regional focus or quality.

- ❑ Take-up across the world of immediate OA publishing models grew much faster than for subscription-based models. Nevertheless, take-up of subscription-based models grew in absolute terms globally at around 2% pa; but it was more or less flat (-0.4%) in the UK.
- ❑ Of the immediate OA models, uptake of the Gold-Hybrid option grew the fastest globally and in the UK, while for Gold-no APC journals it was flat (globally) or declining (UK). For Gold-APC journals, growth was more modest than for hybrid journals, although much faster than for subscription-based journals.
- ❑ Although uptake of the Gold-Hybrid grew faster, the Gold-APC journals still dominate take-up of the immediate OA publishing options: Gold-APC journals account for 9.6% out of the total 16.6% of immediate OA articles globally).
- ❑ While APC-supported immediate open access publishing is growing fast, subscription-based publishing still accounted for over 83% of all research outputs globally (and nearly 82% of UK outputs) in 2014.
- ❑ Of subscription-based models, publishing in journals that have a delayed-OA policy grew faster (4.7% a year globally; 3.5% in the UK) than subscription-only, the latter declining in the UK (1.6% a year globally; -1.0% in the UK). The higher uptake of delayed-OA journals may in part be because they have, on average, the highest citation rates as measured by field-weighted citation impact (FWCI), and this may in turn attract more papers.
- ❑ In 2012 UK uptake of immediate OA models (13%) was below the global average (14%), mainly because of low uptake of Gold-no APC OA model. But UK uptake of such models grew 24% a year between 2012 and 2014, faster than the global average of 14%. As a result, immediate OA, despite a further decline in Gold-no APC model, accounted for over 18% of UK articles in 2014, higher than the global average of under 17% (Fig. 4).

Figure 4. Journal publishing models employed by Global and UK authors

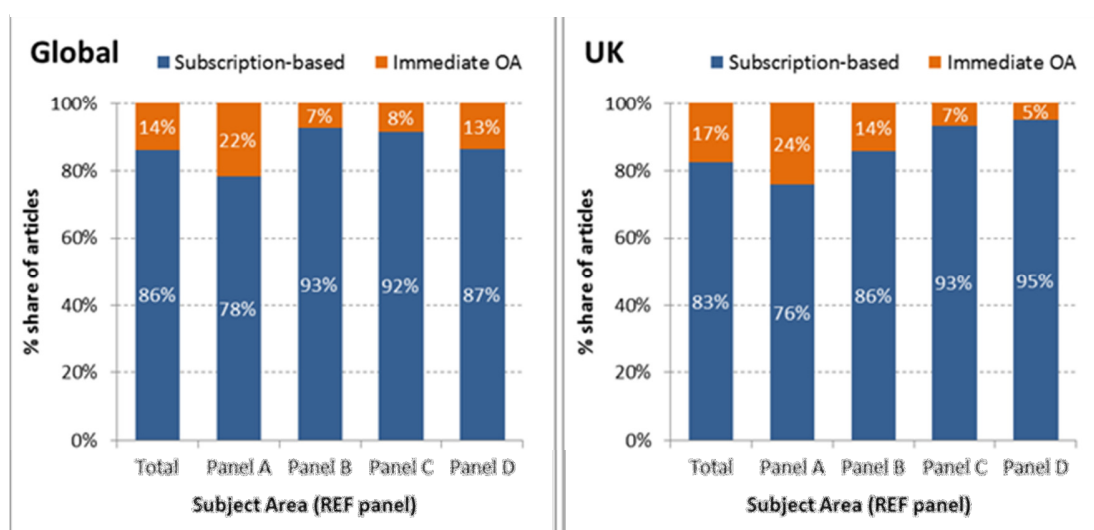


3.3.3. Publishing models – uptake by broad subject areas

We grouped journal titles into the four broad subject areas covered by the main panels employed in the REF exercise. The results summarised in Figure 5 show that:

- ❑ UK authors' uptake of immediate OA models matches or exceeds world levels in health and life sciences (Panel A) and physical sciences and engineering (Panel B) but is below world levels in social sciences (Panel C) and arts and humanities (Panel D).
- ❑ UK authors' lower-than-average take-up of immediate OA in social sciences and in arts and humanities may be explained in part by their preference to publish in high-impact journals, of which there are relatively fewer amongst the immediate OA options in these subject areas. This is discussed further in the next section (3.3.4).

Figure 5. Journal publishing models employed in 2014 by broad subject areas (REF panels)



3.3.4. Relative citation impact of choices available to authors

- ❑ Field-Weighted Citation Impact or FWCI (see Annex E for definition) is used here as a measure of the relative status and impact of publication options available to authors. The weighting employed in this measure allows for comparison across subject areas. It is important here to note that the FWCI measure here is for journals on average and not for individual articles that UK authors publish in these journals.
- ❑ As we have shown in Tables 9a and 9b, on average, subscription-based journals (including hybrid journals) are more highly cited than other immediate-OA journals.
- ❑ Those tables also show that UK authors tend to choose more highly-cited journals for their articles compared to the world average (the FWCI for the UK shown in Table 9b is higher than the world average shown in Table 9a for all publishing models). This may in part explain the higher take-up in the UK (6.7%) compared to the world average (2.4%) for hybrid journals. Similarly, a much higher proportion of UK articles (11.2%) appear in delayed OA journals (the most-highly-cited of all journal categories) as compared to the world average (5.4%).
- ❑ The proportion of global articles in the top 20% of most-cited journals broken down by publishing models is shown in Figure 6 for each REF panel area. Although hybrid and subscription-only journals dominate in the highly-cited spectrum, it is no surprise that health and life sciences (Panel A) - the area where OA has been established for the longest time - has a relatively higher proportion of articles in top cited journals in immediate OA (particularly fully OA Gold-APC) journals. Physical

sciences and engineering (Panel B) also have a high proportion of top journals that offer an immediate OA option, albeit more hybrid than fully-OA (Gold-APC or Gold-no APC).

- Social sciences (Panel C) and arts and humanities (Panel D) have relatively fewer high-impact fully-OA journals, and the hybrid option has been developed more recently in these areas. Together with the UK preference for high-impact journals, this may explain in part the lag in UK take-up of immediate OA options in these areas, relative to the world average. Relative differences in the availability of funding to enable immediate OA in these areas may also play a part. Further investigation of these issues is, however, beyond the scope of the current study.

Figure 6. Proportion of global articles in the top impact (top 20% most cited relative to their subject area) journals within each subject area (by REF panel) and business model

**Proportion of articles in top 20% most cited journals
(as measured by FWCI)**

	Gold - APC	Gold - No APC	Hybrid*	Subscription only
Panel A	15%	1%	72%	12%
Panel B	4%	3%	83%	10%
Panel C	2%	3%	81%	14%
Panel D	2%	3%	64%	31%

FWCI = field-weighted citation impact

** Hybrid-Total*

3.4. Online Posting

Our study aimed to replicate user behaviour by searching for and locating articles on publicly accessible sites. In the context of this study the term “posting” thus refers to articles that have been posted to websites or repositories and are publicly accessible in full but *excludes* any article posted but not publicly accessible, since it was awaiting the expiry of an embargo.

3.4.1. Overall findings

The results of our search for articles published in the 24 months up to March 2015 which were accessible on sites other than the publishers’ are shown in Tables 10a (global sample) and 10b (UK sample). They show that:

- Overall 19% of the UK articles had been posted and were publicly accessible online and in accordance with the policies of the journals concerned. This level of posting is higher than the world average of 15%. The great majority of these postings were of the VoR (10% for the global sample, 12 % for the UK sample).
- Those overall totals include, however, 62% of the UK papers (56% for global papers) that were accessible immediately on publisher sites (because they were published OA), but which were also accessible as versions posted in various repository or other sites. Effectively, around 37% of all posted global or UK articles that are accessible via such sites are already openly accessible via publisher sites. This level of duplication of access via immediate OA (Gold) and posted (Green) routes may partly account for the high estimates of total OA from an earlier study using a robot-

based approach to search selected websites and repositories, but which was “based on the addition of gold OA and of hybrid and green OA”¹².

- ❑ We estimate that a further 11% of UK articles, and 9% of global articles – mostly VoRs - have been posted ‘illicitly’ in the sense that they were not in accordance with the terms of relevant journal policies. Taking these into account, the total level of posting proportion of articles posted - across **all** journals (OA and subscription-based) and including those posted illicitly - is 30% for the UK and 24% for the world.
- ❑ However, the above totals include postings of articles that are already openly accessible via publisher sites. In order to gain a picture of the proportions of papers that are made OA solely as a result of posting, we must focus solely on **subscription-based journals** (including articles in hybrid journals where authors have not taken up the immediate OA option). We estimate that 10% of UK and 9% of world papers published in such journals are publicly accessible online in versions posted on various repository sites and in accordance with journal policies. Including illicit postings as well as those in accordance with the policies of such journals, a total of 19% of global papers, and 23% of papers with UK author, are freely accessible in repositories and other online sites.

¹² Archambault, E. *et al.* (2013) "Proportion of Open Access Peer-Reviewed Papers at the European and World Levels—2004-2011", Report for European Commission DG Research & Innovation, pg.ii. The study also unhelpfully combined Hybrid immediate OA and Delayed OA with posted content into an overall Green OA category. Furthermore that study also looked at access to much older posted papers (12 to 60 months after publication) compared to the current study which is focussed on access within 1-26 months after publication (effectively posted 0-24 months after publication allowing for any delays in posted items to get indexed and be visible online).

Table 10a Article postings within each publishing model (Global)

VERSIONS OF ARTICLES POSTED - GLOBAL*

Label	Definition	Immediate OA				Subscription-based			Total % articles	Total (excl illicit) % articles
		Gold- APC % articles	Gold- noAPC % articles	Gold- Hybrid % articles	immediate OA -total % articles	Delayed OA % articles	Subscription % articles	sub-total % articles		
PP	Preprints (pre-refereed versions)	2.7%	2.1%	3.4%	2.7%	2.5%	4.2%	4.2%	3.9%	3.9%
AAM	Accepted author manuscripts (post-refereed versions)	1.5%	2.2%	2.8%	1.8%	8.0%	3.2%	3.3%	3.1%	2.5%
VoR	Final published journal articles - the Version of Record	64.8%	32.1%	33.1%	52.5%	39.7%	11.1%	12.2%	17.8%	9.6%
Total	de-duplicated	67.8%	35.2%	38.4%	56.0%	47.2%	17.6%	18.8%	24.0%	15.1%
Total	excluding illicit postings	67.8%	35.2%	38.4%	56.0%	22.3%	7.9%	8.5%	15.1%	
<i>Proportion of total published articles</i>		9.6%	4.6%	2.4%	16.6%	5.4%	78.0%	83.4%	100%	

Table 10b Article postings within each publishing model (UK)

VERSIONS OF ARTICLES POSTED - UK*

Label	Definition	Immediate OA				Subscription-based			Total % articles	Total (excl illicit) % articles
		Gold- APC % articles	Gold- noAPC % articles	Gold- Hybrid % articles	Immediate OA -total % articles	Delayed OA % articles	Subscription % articles	sub-total % articles		
PP	Preprints (pre-refereed versions)	3.7%	6.4%	3.2%	3.7%	2.2%	6.1%	5.9%	5.5%	5.5%
AAM	Accepted author manuscripts (post-refereed versions)	0.9%	4.7%	3.9%	2.2%	5.9%	3.4%	3.6%	3.3%	2.6%
VoR	Final published journal articles - the Version of Record	66.3%	36.2%	48.4%	58.0%	39.8%	12.8%	14.6%	22.4%	12.4%
Total	de-duplicated	69.5%	45.7%	51.6%	61.6%	46.6%	21.2%	22.9%	29.8%	19.0%
Total	excluding illicit postings	69.5%	45.7%	51.6%	61.6%	17.0%	9.2%	9.7%	19.0%	
<i>Proportion of total published articles</i>		9.3%	2.1%	6.7%	18.2%	11.2%	70.7%	81.8%	100%	

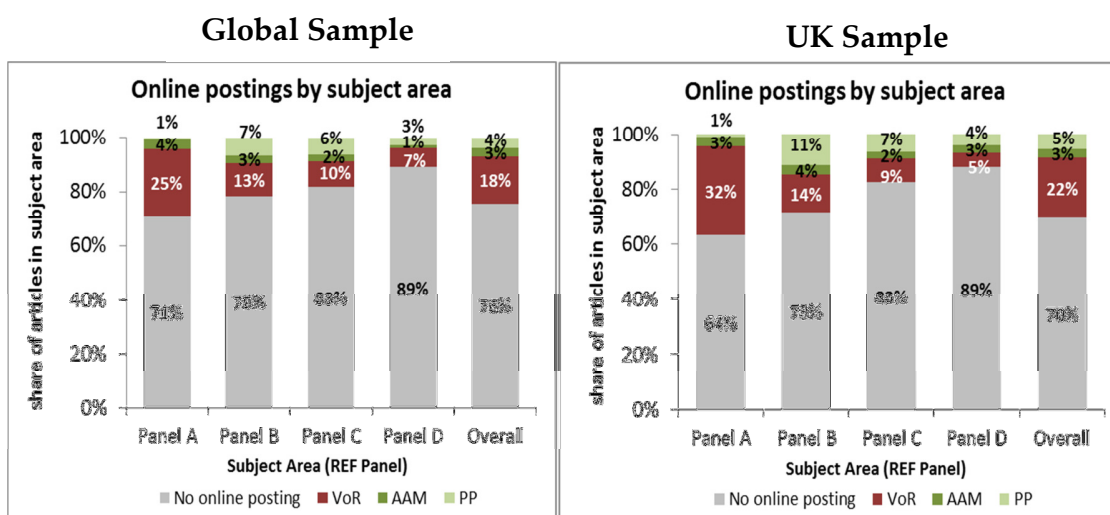
* measured in March 2015 for articles published between 1 to 26 months before

3.4.2. Variation in postings by subject area

Figure 7 shows variations in the proportions of articles found to be accessible on sites other than the publisher's, broken down by subject area. It shows that:

- Posting levels for the UK sample are particularly high - and higher than the world averages - in health and life sciences (Panel A) and physical sciences and engineering (Panel B); and on par with the world averages in social sciences (Panel C) and arts and humanities (Panel D).
- The posting of VoRs predominates, particularly in the health and life sciences (Panel A), while the posting of preprints remains strong in the physical sciences and engineering (Panel B) and social sciences (Panel C).
- Posting of AAMs is relatively low at around 3% across most subject areas. This may be in part because authors find VoRs easier than AAMs to locate for posting purposes.

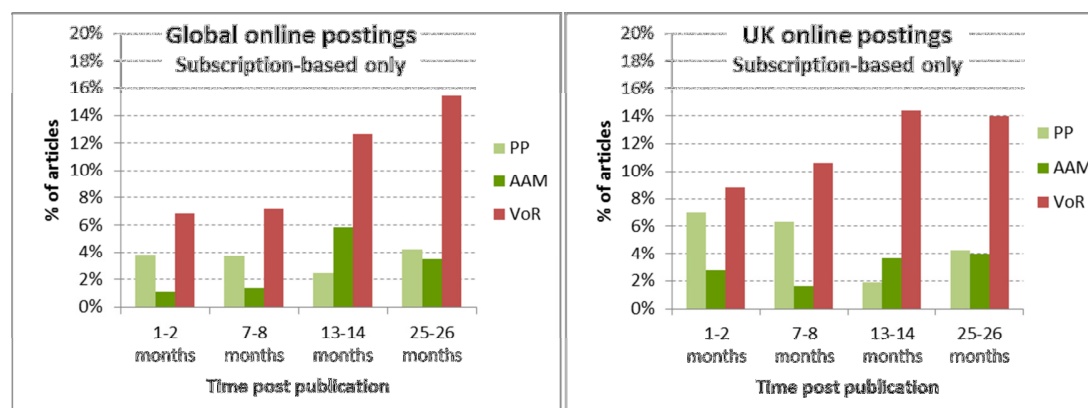
Figure 7. Variation in online posting by REF panel subject area (includes postings related to both immediate OA and subscription-based journals as well as illicit postings)



3.4.3. Online postings over time

The sampling of publications for four separate time periods (1-2, 7-8, 13-14 and 25-26 months post-publication) enabled an assessment of when posted articles became accessible. The results shown in Figure 8 cover postings for subscription-based articles only in each of those four time periods.

Figure 8. Variation in postings (including illicit postings) over time after publication (for subscription-based articles only)

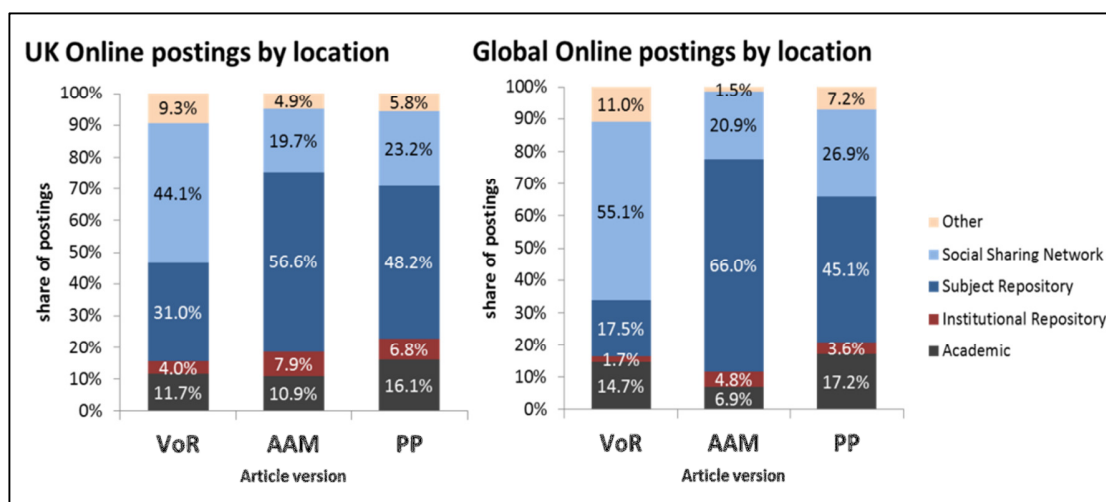


- The volume of accessible online postings of AAMs and VoRs for subscription-based journals globally increases with time post publication, peaking at 12 months for AAMs, as one might expect in line with embargo policies, and at 24 months for VoRs. Globally, we estimate that about 24% of AAM postings and 79% of VoR postings of subscription-based articles do not comply with journal policies. The vast majority (70%) of such postings are on social sharing or similar sites.
- For UK articles, however, VoR and AAM postings tend to become accessible earlier; and we estimate that 28% of AAM postings and 84% of VoR postings do not comply with journal policies. These figures are higher than the world averages, and it is possible that this is driven by attempts to meet funder requirements - including the new REF requirement for posting at the point of acceptance rather than publication - by posting early, and in the form of a VoR, without checking on journal policies.

3.4.4. Location of postings

- We found the vast majority of discoverable and accessible postings in subject repositories and in social sharing networks. Relatively few were in IRs. Since we aimed to replicate users' searching behaviour, we did not discover content on sites where it was *in theory* publicly accessible, but not amenable to discovery by search engines, and thus could not be found with reasonable effort.

Figure 9. Location of online postings (including illicit postings)



- Two-thirds of all UK AAM postings (71% globally) in subject repositories were on PubMed Central (PMC) and 83% of all UK social sharing network AAM postings were on ResearchGate (81% globally).
- For VoRs almost all subject repository postings are in PMC and almost all social sharing network postings are in ResearchGate.


- Nearly 70% of all illicit postings appear on social sharing sites, mostly as a result of VoRs posted on ResearchGate.

3.5 An estimate of the total proportion of OA articles


In order to estimate the total current extent of OA, we need to consider immediate and delayed OA modles, as well as posted content, de-duplicated where posted content is the same as that which is already open access via other means. The age of articles that are open access also needs to be taken into account, bearing in mind that we consider in this study only articles that are accessible within 24 months of publication. A summary total view is given in Tables 11a and 11b for Global and UK samples. They cover the *whole* of our global and UK samples, not simply those published in our four separate time periods, and present cumulative figures over time.

Tables 11a and 11b. Total proportions of OA content for the world and for the UK

(a) Total proportion of Open Access content - 2014

GLOBAL	Cumulative % 				
Months (after publication)	0	6	12	24	
Sampled months*	1-2	7-8	13-14	25-26	
Gold-APC	9.6%	9.6%	9.6%	9.6%	All articles immediately OA upon publication
Gold-noAPC	4.6%	4.6%	4.6%	4.6%	
Gold-Hybrid	2.4%	2.4%	2.4%	2.4%	
Delayed OA		0.8%	4.1%	4.5%	A futher 1% accessible later for embargoes >24 months
OA Postings (excl. illicit postings)** ⁺	1.3%	2.5%	4.2%	6.2%	Subscription articles only
TOTAL (excl. illicit postings)	17.9%	19.9%	24.9%	27.3%	An additional 3% accessible free at publisher sites***
OA Postings (incl illicit postings)** ⁺	2.7%	5.5%	8.9%	12.8%	Subscription articles only
TOTAL (incl. illicit postings)	19.3%	22.9%	29.6%	33.9%	An additional 3% accessible free at publisher sites***

(b) Total proportion of Open Access content - 2014

UK	Cumulative % 				
Months (after publication)	0	6	12	24	
Sampled months*	1-2	7-8	13-14	25-26	
Gold-APC	9.3%	9.3%	9.3%	9.3%	All articles immediately OA upon publication
Gold-noAPC	2.1%	2.1%	2.1%	2.1%	
Gold-Hybrid	6.5%	6.5%	6.5%	6.5%	
Delayed OA		1.8%	8.7%	9.6%	A futher 1.5% accessible later for embargoes >24 months
OA Postings (excl. illicit postings)** ⁺	2.0%	4.2%	5.5%	7.5%	Subscription articles only
TOTAL (excl. illicit postings)	19.9%	23.9%	32.1%	35.0%	An additional 3% accessible free at publisher sites***
OA Postings (incl illicit postings)** ⁺	4.3%	8.4%	11.9%	15.8%	Subscription articles only
TOTAL (incl. illicit postings)	22.2%	28.1%	38.5%	43.3%	An additional 3% accessible free at publisher sites***

* for postings and Gold-Hybrid only. Up to two months were sampled to allow for lead time for posting to be discoverable and indexed by search engines

** includes posting for subscription content only, deduplicated for multiple postings and for content also available via delayed OA

*** usually for promotional or public service purposes

⁺ illicit postings are those that do not comply with journal policies. The vast majority (>90%) of 'illicit' postings was the result of publishers' versions of record being posted, contrary to the policies of those journals (which allow posting of AAMs but not VoRs)

The tables indicate that:

- ❑ Globally 18% of articles published in the last two years (19% including illicit postings) were openly accessible immediately on publication, rising to 20% (23% including illicit postings) within 6 months, 25% (30%) within 12 months and 27% (34%) within 24 months.
- ❑ For UK articles, 20% (22% including illicit postings) were openly accessible immediately upon publication, rising to 24% (28%) within six months, 32% (38%) within 12 months and 35% (43%) within 24 months.
- ❑ A further 3% of articles (globally and for the UK) are also accessible freely at publisher sites for promotional or public service purposes (e.g. related to the recent Ebola epidemic).

Summary of key findings

Publishing Models

- ❑ Two-thirds of the world's journals offer the potential for immediate OA to the VoR on the publisher's site (nearly 77% for the journals in which UK authors typically publish).
- ❑ Global uptake of immediate OA models grew faster than subscription-based models between 2012 and 2014, although both increased in absolute terms. Over the same time the UK showed an even faster uptake of immediate OA models, but uptake of subscription-based models was more or less static (-0.4%).
- ❑ Immediate OA models (Gold-APC, Gold-no APC and Gold-Hybrid) accounted for just under 17% of global output in 2014 (14% in 2012) and over 18% of UK output of research papers (13% in 2012).
- ❑ Take-up of APC-based models, particularly hybrid, grew fastest between 2012 and 2014. Take-up of non-APC OA models was static (global) or declined (UK)
- ❑ UK take-up of hybrid and delayed OA models is considerably higher than the world average.
- ❑ UK authors show a preference for publishing in higher impact journals (as measured by FWCI)
- ❑ UK take-up by subject is similar to the world averages for health and life sciences (Panel A) and a little above the world average for physical sciences and engineering (Panel B). Take up of immediate OA models in HSS subjects (Panels C and D) is lower than the world averages.
- ❑ Lower uptake for immediate OA models in HSS subjects maybe explained in part by there being fewer high-impact immediate OA options in these areas, and to hybrid options being available more recently than in the STEM disciplines.
- ❑ Health and life sciences (Panel A) has highest uptake of immediate OA models.

Postings

- ❑ Versions of 19% of the papers published by UK authors in the past two years – and 15% of papers published globally - have been posted and are accessible online and in accordance with the relevant journal policies.
- ❑ UK postings are most strongly ahead of the world average in health and life sciences
- ❑ We estimate that a further 11% of UK articles and 9% of global articles – mostly VoRs on file sharing sites - have been posted 'illicitly' in the sense that they were not in conformity with the relevant journal policies.
- ❑ Overall, therefore, including illicit postings as well as those in conformity with journal policies, a total of 24% of global, and 30% UK papers are freely accessible in repositories and other online sites. The volume of postings increases over time post-publication, but peaks at 12 months.
- ❑ Some of these postings are of articles that are already openly accessible via publisher sites. If we focus solely on **subscription-based journals** (including articles in hybrid journals where authors have *not* taken up the immediate OA option), we estimate that **10%** of UK and **9%** of world papers published in such journals are publicly accessible online in versions posted on various repository sites and in conformity with journal policies (i.e. excluding illicit postings).

Overall (publishing models + postings)

- ❑ Globally 18% of articles published in the last two years (19% including illicit postings) were openly accessible immediately on publication, rising to 20% (23% incl illicit postings) within 6 months, 25% (30%) within 12 months and 27% (34%) within 24 months.
- ❑ For UK articles, 20% (22% including illicit postings) were openly accessible immediately upon publication, rising to 24% (28%) within six months, 32% (38%) within 12 months and 35% (43%) within 24 months.

4. Usage

4.1 Background

A key aim behind moves to promote OA is to make articles freely accessible, so that they can be read and used by more people, particularly those outside the higher education sector and the relatively small number of companies and other organisations that purchase subscriptions and thus make journal contents accessible to their members and staff. It is thus important to examine whether the policies now being implemented do indeed have that effect. The key questions therefore are:

- i. the extent to which the different versions of articles that are made accessible through one or other OA route are viewed and downloaded, and whether in sum they are viewed and downloaded more extensively than non-OA articles
- ii. whether there are any differences in the demographics of usage of OA as distinct from non-OA articles (in particular, whether OA articles are used more extensively by users from outside the HE and research sectors)

4.2. Context

There have been some efforts to answer the first question in terms of usage of OA and non-OA articles from individual journals and publishers, notably by Davis in two studies of the journals of the American Physiological Society (Davis 2010) and of a sample of journals from a range of publishers using the Highwire platform (Davis 2011); and by the RIN in a study of *Nature Communications* (RIN 2014). They have all shown that OA articles do indeed receive more downloads from publisher platforms than non-OA articles. More than 70 published studies have also sought to ascertain whether OA articles enjoy more citations than non-OA articles¹³. A majority show that there is some citation advantage gained by publishing on OA terms; but randomised trials conducted by Davis show no such effect. The scope of the current study does not allow us to explore this issue.

Moreover, the growth of repositories and of bibliographic sharing sites adds to a long-standing problem in assessing usage, let alone citations. Usage of any article (OA or non-OA) can take place on a number of platforms including personal and departmental websites, course packs, institutional and subject-based repositories, sharing services such as ResearchGate (<https://www.researchgate.net/>), Academia (<https://www.academia.edu/>), Mendeley (<https://www.mendeley.com/>), Zotero (<https://www.zotero.org/>), ReadCube (<https://www.readcube.com/>) etc., and on publishers' and aggregators' sites.

Standards for measuring usage are set by the codes of practice issued by COUNTER (Counting Online Usage of Networked Electronic Resources) (<http://www.projectcounter.org/about.html>), which define methodologies for reporting at the level of individual journals. The Publisher and Institutional Repository Usage Statistics (PIRUS) project (<http://www.cranfieldlibrary.cranfield.ac.uk/pirus2/tiki-index.php?page=pirus2>) developed technical and organisational models for gathering and aggregating statistics at individual article level; and COUNTER issued a PIRUS Code of Practice in 2014, including provision for a central clearing house. That code is used by the IRUS-UK service (see below) which aggregates data from UK institutional repositories. But as the report of the PIRUS 2 project noted (Shepherd and Needham 2011) most publishers have been unwilling to implement or to participate in an aggregating service.

¹³ A summary is maintained by SPARC Europe: http://sparceurope.org/oaca_table/

4.3. Data sources

We have therefore gathered data from four main sources: from publishers; from Institutional Repository Usage Statistics UK (IRUS-UK), the aggregator of download data from UK institutional repositories; from the Journal Usage Statistics Portal (JUSP), the aggregator of publishers' data on downloads from UK universities and colleges; and from Europe PubMed Central. We also sought data from sharing services, but without success.

4.3.1 Publishers

Only a small minority of publishers were willing or able to provide us with data on downloads of OA and non-OA articles via their sites; and some publishers cautioned us about the accuracy of their data (they were not confident, for example, that they were always able to detect robots and crawlers and that they were thus fully COUNTER-compliant). And no publisher was able to provide data that might be used to analyse the demographics of usage in terms, for example, of users from within and outside academia.

Those publishers that did provide data confirmed earlier findings that OA articles are downloaded more than non-OA articles; but they showed a complex pattern of levels of usage across different journals. For the ten journals (two of them fully OA) of a relatively small but high-status publisher, as shown in Table 12, the ratios between downloads in the years 2013-2014 of articles published OA in hybrid journals in those years ranged from just over two-and-a-half to more than seven times in favour of OA; and the highest level of average downloads per article was for OA articles in a hybrid, rather than a fully-OA journal.

Table 12. Downloads of OA and non-OA articles in 2013-14 from a small publisher

Journals	All publications		Open Access		Non Open Access		Ratio of downloads of OA/non-OA
	Total number of articles)	HTML/PDF downloads	No. articles	Av. downloads per article	No. articles	Av. downloads per article	
1	678	285,922	58	1,463	620	324	452%
2	815	887,130	204	1,957	611	799	245%
3	443	376,065	40	3,579	403	578	619%
4	1,208	1,709,396	223	2,758	985	1,111	248%
5	452	587,593	65	2,001	387	1,182	169%
6	654	1,340,695	151	4,874	503	1,202	405%
7	136	78,014	24	1,383	112	400	346%
8	141	23,538	8	557	133	143	390%
9	136	622,370	136	4,576	0	n/a	n/a
10	52	63,606	52	1,223	0	n/a	n/a
Total	4,715	5,974,329	961	2,947	3,754	837	352%

Journals 9 and 10 are fully-OA

For a second slightly larger medium-sized publisher, a key distinction arises in patterns of downloads of OA and non-OA articles as between users at subscribing institutions and those whose location is unknown. As shown in Table 13, users at subscribing institutions showed in 2014 (and for articles

published that year) the highest average levels of downloads for OA articles in hybrid journals, but the lowest average levels for articles in fully-OA journals. Users at unknown locations, on the other hand, showed average levels of downloads of articles in hybrid journals nearly twice as high as users in subscribing institutions, but for articles in fully-OA journals usage is nearly three times as high. Not surprisingly, downloads for non-OA content were much lower than shown by users in subscribing institutions, and related in the main to delayed OA journals, and to articles made freely-accessible for promotional purposes. As with the smaller publisher, there were also significant differences between OA and non-OA downloads for particular journals: among the hybrid journals, the ratios ranged from equality to more than 10 times higher for OA articles.

Table 13. Downloads of OA and non-OA articles in 2014 from a medium-sized publisher

User category (known or unknown institution)	Journal and article type	No. Downloads to Current Year Articles	No. Articles Downloaded	Ave Downloads / Article
Known	Non-Open Access Articles	2,400,916	22,211	108
Known	Open Access Articles - Fully Open Access Journal	449,956	8,359	54
Known	Open Access Articles - Hybrid Journal	129,354	809	160
Known	Subtotal	2,980,226	31,379	95
Unknown	Non-Open Access Articles	646,182	8,419	77
Unknown	Open Access Articles - Fully Open Access Journal	1,204,456	8,406	143
Unknown	Open Access Articles - Hybrid Journal	246,761	798	309
Unknown	Subtotal	2,097,399	17,623	119
All Users	Non-Open Access Articles	3,047,098	22,215	137
All Users	Open Access Articles - Fully Open Access Journal	1,654,412	8,407	197
All Users	Open Access Articles - Hybrid Journal	376,115	809	465
All Users	Total	5,077,625	31,431	162

4.3.2 Journal Usage Statistics Portal (JUSP)

JUSP gathers download data on behalf of nearly 180 higher education libraries in the UK, using COUNTER reports, from 78 publishers and intermediaries¹⁴. The data supplied to us is aggregated at publisher level, and as Table 14 shows, just under 174 million successful requests for full-text articles were recorded in 2014 (the JR1 report). Among the top 25 publishers, the requests recorded ranged from nearly 59 million to over half a million. The table also shows the number of requests for a sub-set of articles that came from publishers' archive or back-file collections which institutions purchase separately (JR1A); and the number of full-text requests for Gold OA articles (JR1 GOA)¹⁵. In both cases, the figures shown in these columns are included in those shown in the JR1 column.

The final column shows the proportion of Gold OA article requests to all requests, amounting to four percent for all publishers. But it is clear from the table that there is no obvious pattern by overall number of downloads per publisher platform; nor could we find any pattern when we sorted in any other way.

¹⁴ Since these publishers and intermediaries often use more than one platform JUSP records COUNTER data from 148 locations

¹⁵ It should be noted, however, that the JR1 GOA report does not cover fully-OA publishers such as PLOS.

Access to the data for individual journals rather than aggregated at publisher level would enable a deeper level of analysis which would be valuable; but we did not have access to such data. And the data available via JUSP at present does not include the JR5 COUNTER reports which distinguish requests by year of publication. The JR1 reports include requests for all articles, whatever the year of publication, and since we know that OA is growing faster than subscription-based publishing the figures presented here may understate the proportion of requests for OA as distinct from non-OA publications. We understand that JUSP has identified issues that will need to be resolved before data from the JR5 report can be included with confidence; and we hope that these issues can be dealt with satisfactorily as soon as possible. Again, however, the obvious point needs to be stressed, that JUSP data relates only to usage in UK universities, and not in any other location.

Table 14. Download data for 2014 from JUSP

Top 25 publishers	JR1	JR1a	JR1 GOA	JR1 GOA as % if JR1
1	58,999,520	0	4,346,332	7.40%
18	22,053,079	1,067,834	364,462	1.70%
99	21,410,318	20,320,682	0	0.00%
66	10,279,480	533,826	78,485	0.80%
8	8,335,528	849,265	511,720	6.10%
12	7,300,215	452,818	57,575	0.80%
3	6,908,741	0	408,861	5.90%
2	5,635,371	502,055	276,372	4.90%
111	3,831,928	0	0	0.00%
98	3,122,825	502,387	44,342	1.40%
229	2,762,589	269,143	0	0.00%
30	2,399,502	729,164	0	0.00%
17	2,363,183	632,239	326,655	13.80%
124	1,755,099	0	0	0.00%
10	1,456,982	118,722	5,793	0.40%
96	1,395,808	0	23,738	1.70%
7	1,345,411	0	117	0.00%
13	806,833	0	97,226	12.10%
129	785,173	0	170,379	21.70%
14	672,183	349,881	0	0.00%
155	663,229	0	31,854	4.80%
118	663,007	0	16,068	2.40%
6	643,968	128,753	20,295	3.20%
152	589,666	70,287	2,173	0.40%
117	561,862	0	2,730	0.50%
Total for all publishers	173,917,681	26,678,228	7,040,035	4.00%

4.3.3 Institutional repositories

IRUS-UK aggregates COUNTER-compliant usage statistics at individual article level for some 80 institutional repositories in the UK. It recorded six million downloads of articles in 2014, just over half of which (3.2 million) could be ascribed to a specific journal title. In other cases the metadata was not of sufficient quality to link the downloaded article with sufficient certainty to a particular journal. In total, over 7,000 journals were identified with downloads.

It is notable, however, that just twenty of those journals accounted for a tenth of all downloads. And Table 15 shows that within those journals, one or two articles accounted in some cases for up to 90% and more of all downloads. Indeed, a single methodological article published in 2006 in the journal *Qualitative Research in Psychology*¹⁶ accounted for 2.5% of all downloads recorded by IRUS-UK in 2014. In other cases, it is not easy to discern the reasons for sudden spikes of many thousands of downloads over a short period for articles that were published several years previously.

Five further points are worth making in relation to the IRUS data.

- ❑ First, the numbers of downloads are small in comparison, as we shall see, to those of a major subject repository such as PMC, or to those of articles on the publisher website published in a single year by a small or medium-sized publisher.
- ❑ Second, social science – and especially management - journals figure prominently in the top twenty journals by downloads; this may reflect a higher preponderance of IRs as distinct from subject repositories in the social sciences as compared with the life or the physical sciences.
- ❑ Third, the top twenty and indeed the top hundred journals by downloads comprise a mix of major journals in their field, alongside journals with a narrower focus which nevertheless have published articles that authors have posted and which have proved popular among users.
- ❑ Fourth, we have analysed the IRUS data for the 25 journals most popular with UK authors in our four subject areas, and have been unable to detect any relationship between numbers of articles published by UK authors on the one hand, and numbers of articles downloaded from UK IRs on the other. Indeed, six of the 25 most popular journals in the physical sciences and two of those in the life sciences do not feature in IRUS records at all. And in the humanities in particular - and to a slightly lesser extent in the social sciences - numbers of downloads are low, with one or two exceptions that are not easy to explain.
- ❑ Fifth, it is worth noting that fully-OA journals including PLOS ONE– and others slightly lower down the list and thus not shown in Table 15 - show high numbers of downloads from IRs in the UK. Articles from such journals are thus being posted in IRs and used by readers alongside their free availability on the journal platform.

Again, as with the data from publishers, no data is available that might be used to analyse the demographics of usage. The IRUS-UK data thus gives rise to some interesting questions; but it would

¹⁶ The article ‘Using thematic analysis in psychology’ has been cited more than 13,000 times according to Google Scholar; and it is notable that the results of Google searches for the term ‘thematic analysis’ tend to put the PDF available in the University of the West of England repository higher in listings on the results page than the version available from the journal website. Downloads from the repository at the University of Auckland, where the lead author still works, currently (August 2015) amount to 523.

require a very considerable level of detailed analysis – beyond the scope of this study – to reach any firm conclusions about the extent to which, and how, usage via UK repositories is contributing to meeting the aims underlying the moves towards OA.

Table 15. The top 20 journals downloaded from UK institutional repositories in 2014

20 journals with greatest number of article downloads, 2014	Total	Accounted for by single most downloaded article across all IRs		Accounted for by two most downloaded articles across all IRs		No. of articles downloaded
		Nr	%	Nr	%	
Qualitative Research in Psychology	83,496	78,047	93.5%	82,787	99.2%	16
Corporate Governance: An International Review	34,554	33,068	95.7%	33,614	97.3%	8
PLOS ONE	32,892	607	1.8%	840	2.6%	2,844
Coaching: An International Journal of Theory, Research and Practice	18,705	13,523	72.3%	16,573	88.6%	9
Annals of Tourism Research	17,158	3,091	18.0%	5,482	32.0%	45
British Journal of Management	16,322	5,464	33.5%	9,477	58.1%	33
Journal of Business Ethics	14,948	5,151	34.5%	9,082	60.8%	49
Nature	14,929	4,766	31.9%	5,884	39.4%	104
Journal of Social Policy	14,714	3,183	21.6%	4,565	31.0%	89
International Journal of Human Resource Management	14,634	4,816	32.9%	6,925	47.3%	34
European Journal of Marketing	14,338	6,796	47.4%	8,845	61.7%	24
New Media & Society	14,289	6,783	47.5%	8,695	60.9%	24
International Journal of Operations & Production Management	14,137	5,577	39.4%	8,167	57.8%	46
Ageing & Society	13,998	3,501	25.0%	4,813	34.4%	82
Journal of Applied Physics	13,228	2,496	18.9%	3,069	23.2%	302
Magnetics, IEEE Transactions on	12,581	1,669	13.3%	3,033	24.1%	93
Applied Physics Letters	12,492	296	2.4%	551	4.4%	572
World Development	12,337	6,767	54.9%	8,016	65.0%	33
Pattern Analysis and Machine Intelligence, IEEE Transactions on	12,329	7,136	57.9%	8,694	70.5%	46
Industrial Marketing Management	12,147	5,818	47.9%	6,980	57.5%	33

4.3.4 Subject repositories

The major repository from which we sought usage data was PubMed Central (PMC), the principal such repository in the health and life sciences. Many publishers deposit articles direct in PMC where that is a requirement of major funders in the US, the UK and elsewhere. As Table 16 indicates, the number of articles deposited is growing at 3-400,000 a year, and now totals over 3.5 million. The number of articles retrieved is also growing fast, both absolutely and in terms of average per article. In 2014, articles were retrieved nearly 659 million times, an average of 188 retrievals per article. It is also notable that the number of retrievals via PMC of the full text HTML version is growing much faster than retrievals of the PDF. Yet again, however, no data is available that might be used to analyse the demographics of usage.

Table 16. Retrievals of articles from PubMed Central

Year	Total No. Articles Available	HTML Full Text Retrieval	Article PDF Retrieval	Av retrieval per article
2012	2,790,219	251,363,758	104,153,931	127
2013	3,119,643	389,623,123	106,260,140	159
2014	3,506,234	513,545,220	145,227,684	188

4.3.5 Sharing sites

As noted above, none of the sharing services we contacted were prepared to provide any usage statistics.

Summary of key findings

- ❑ Any rigorous attempt to answer the question whether OA leads to higher levels of usage must take account both of the different varieties of OA – articles in fully-OA, hybrid and ‘delayed OA’ journals, and those made accessible in different versions on websites, repositories and sharing services - and of the many different locations via which they might be used. The data we have been able to gather does not enable us to address these issues in any kind of rigorous or consistent way.
- ❑ **Publishers** can in principle tell us about usage on their platforms, and the great majority of the larger publishers provide such data in relation to usage from UK institutions at the journal level to the JUSP service. But until JUSP includes data from the JR5 report which provides the year of publication, any conclusions about levels of usage of OA as distinct from non-OA articles must be treated with caution. And widespread implementation by publishers of the PIRUS code of practice would be required in order to analyse usage at the level of individual articles.
- ❑ The **IRUS-UK** service enables us to analyse usage of individual articles via the main IRs in the UK; and usage patterns appear to raise significant questions - for example about the subject distribution of usage, the age of the articles that are being used, and their availability on other sites and platforms - that should be addressed in future work. At present, however, the data suggests that both deposit (see Section 3) and usage of articles via UK IRs is dwarfed by use via subject repositories and sharing sites, and indeed via publisher platforms.
- ❑ The key issue of **consolidating or aggregating usage data** from the various locations on which usage of an individual article can take place is thus unresolved; and the absence of usage data from the sharing sites means that a significant element in the overall pattern of usage is currently missing.
- ❑ Our key conclusion, therefore, is that unless and until article-level download data is made openly available – or at least the PIRUS code and the proposed central clearing house are widely adopted -any overall analysis of usage data for OA and non-OA articles will be problematic and partial at best. In seeking any wider adoption of PIRUS, both the major subject repositories and the sharing services need to be brought into the picture. Even then, however, it appears unlikely that we shall be able to answer key questions about the demographics of usage.

5. Financial sustainability: universities

5.1 Research objectives

The research in this part of our study had the following objectives:

1. To analyse expenditure on APCs during 2014 by a sample of UK HEIs; and to compare this with previous APC expenditure (as reported in Pinfield, Salter, & Bath, 2015)
2. To model APC expenditure and new administration costs with existing subscription expenditure (a measure previously labelled “total cost of publication”) (Willettts, 2014); and to compare this with previous such calculations (Pinfield et al., 2015)
3. To make recommendations about future approaches to data collection

5.2 Method

Data on expenditure on APCs was collected in partnership with Jisc during the first quarter of 2015 from a sample of UK institutions. Jisc compiled the data into a single dataset in a standard template (with the dataset in this form also being made available on Figshare). This included:

- ❑ Detailed APC data (including a record of all centrally-managed individual APCs paid) from 24 HEIs in non-anonymised form as reported by: Bangor, Bath, Birmingham, Bristol, Cranfield, Durham, Glasgow, Imperial College London, Lancaster, Leicester, Liverpool, Loughborough, London School of Hygiene and Tropical Medicine (LSHTM), Newcastle, Plymouth, Portsmouth, Queen Mary University of London (QMUL), Royal Holloway London (RHUL), Salford, Sheffield, Sussex, Swansea, University College London (UCL), Warwick.
- ❑ Headline APC data (including only total APC expenditure) from 23 HEIs (the same HEIs covered in previous work and therefore useful for longitudinal analysis) reported in anonymised form only

The data received was in a considerably better shape than that which had been previously analysed. 2013 data had required extensive checking, normalisation and augmentation (Pinfield et al., 2015; Woodward & Henderson, 2014). However, the 2014 data still required considerable work, including:

- ❑ Adding missing publication dates
- ❑ Disambiguating journal titles
- ❑ De-duplicating records
- ❑ Checking apparently anomalous figures
- ❑ Adding missing APC prices
- ❑ Carrying out currency conversions

Missing publication dates were added by manually searching for each article based on DOI or title. The journal titles were manually checked to remove misspellings and abbreviations, making them consistent throughout the dataset. Duplicate records were removed through checking of matching DOIs or article titles. Anomalous APC prices were checked with the institutions themselves and changed where appropriate. Missing APC prices were supplied at list price based on data on publisher web sites. Currency conversions were carried out at 0.65 US dollars (\$) and 0.75 euros (€) to the pound (£) respectively. Figures provided include Value Added Tax (at 20%) where it was paid.

A number of issues arose in processing the data, arising from important aspects of the current APC market:

- ❑ Very low APC prices for some items: these were usually explained by discounts, often linked to pre-payment deals. For example, one institution recorded 40 APC payments made to a single publisher averaging less than £40 each, and confirmed that this was the result of a one-off deal with the publisher. There was also widespread use of schemes such as the Royal Society of Chemistry's Gold4Gold scheme which resulted in some £0 payments being recorded (since subscribers were given vouchers enabling some APCs to be free). Such 'free' or highly-discounted APCs were usually parts of wider deals with publishers (including some early offsetting arrangements) and so therefore need to be considered in this overall context (hence the importance of considering total costs to universities, below, rather than APC expenditure in isolation). These values have been checked where possible and corrected (if an error was identified) or accepted (where a low or zero APC payment was verified). Since this study aimed to analyse what institutions were actually paying not simply list prices, APCs were recorded at the discounted rate.
- ❑ Splitting of APC payments, normally between two funders: although still rare, sometimes two or more funders were listed under a single APC payment, presumably where multiple funders had jointly contributed to a research project from which outputs resulted. For analysis, these payments were merged and the agency listed as paying the greater amount was recorded as the funder. For the very small number of payments where there was an even split between funders, the first named funder was recorded.
- ❑ Inclusion of additional charges with APCs: these included colour and page charges being recorded in the same payment as APCs and were often apparent by anomalously high APC prices. Wherever possible these were identified and excluded from the APC figures used for analysis.
- ❑ Inconsistency in the definition of 'publication date': this was noticeable even within single institutions' records, with common definitions of 'publication date' apparently being either when the VoR was made public on the journal site or when it was made part of an issue of a journal. This inconsistency is potentially significant since there can sometimes be a considerable length of time between these two dates. However, it was impossible to correct this inconsistency without wholesale checking and it therefore had to be accepted as a feature of the data.

These features mean that the dataset, although more accurate than previous similar datasets, still comes with caveats. Efforts were made to check and correct obvious anomalies but such efforts did not extend to checking every single payment. It is likely, therefore, that the dataset still includes some inaccuracies and inconsistencies. Recommendations arising from this are discussed in Annex M.

There was some overlap between the 24 institutions covered in the 2014 APC data and the 23 covered in the previous study. They are reported separately here because of the agreement made with the 23 institutions 18 months ago that their anonymity would be preserved. One institution submitted 2014 data (having also taken part in the previous study) and still requested anonymity. Data from this institution has not been included in the 2014 analysis since the other 24 institutions have been named and it would have been inconsistent to maintain anonymity for just institution.

It should be noted that this study included only centrally-managed APC expenditure within institutions. HEIs are currently unable to report reliably on expenditure made elsewhere and it is difficult to estimate levels of such expenditure. It is unlikely that payments of APCs outside the centre would occur at *significant* levels for research funded by the members of RCUK or the Charity Open Access Fund

(COAF) or where institutional pre-payment schemes with publishers are in place, but they may occur for other research outputs, depending on institutional arrangements for funding of APCs (see below).

Subscription data for 2014 used in the study was already in the public domain (Lawson & Meghreblian, 2014), covering seven publishers only: Cambridge University Press (CUP), Elsevier, Oxford University Press (OUP), Sage, Springer, Taylor & Francis, and Wiley. This was considered to be a reasonable sample covering a large proportion of overall subscriptions; but it does not, of course, provide complete coverage of institutional subscriptions.

Administrative cost data was based on averages from Johnson, Pinfield, & Fosci (n.d.). Their study, based on data reported by 29 UK institutions, identified an average administrative cost £88 per APC. This includes only the direct costs of administering each APC payment and excludes more substantial indirect costs such as advocacy, compliance reporting and policy development.

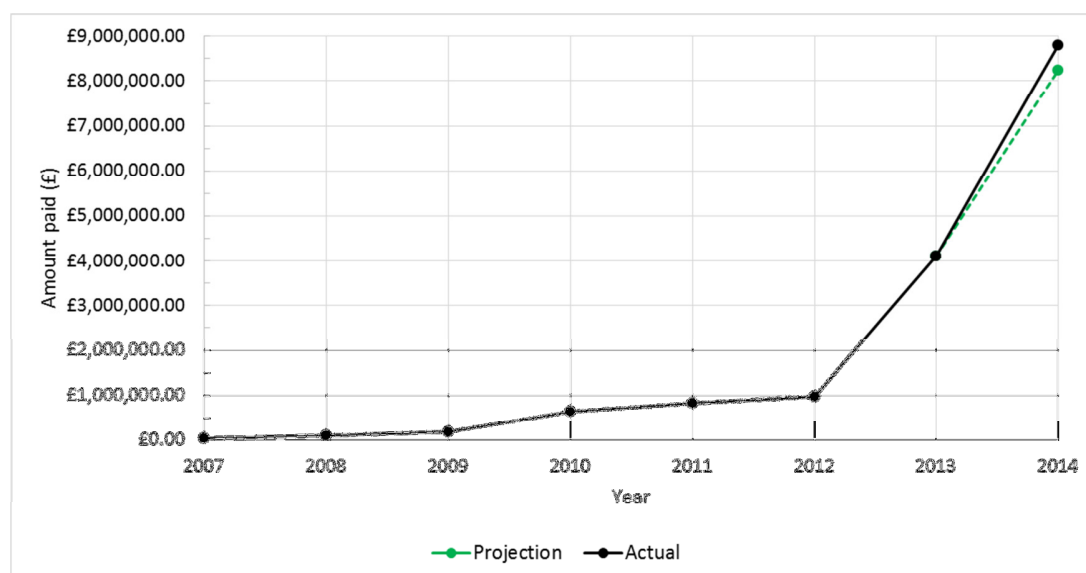
Analysis of the data was based on publication year as the most-easily-publicly-verifiable date, accepting the caveats outlined above. An alternative would have been to carry out analysis by date of payment, but there was insufficient data for both APCs and subscriptions to allow this.

5.3 Results

5.3.1 APC expenditure growth

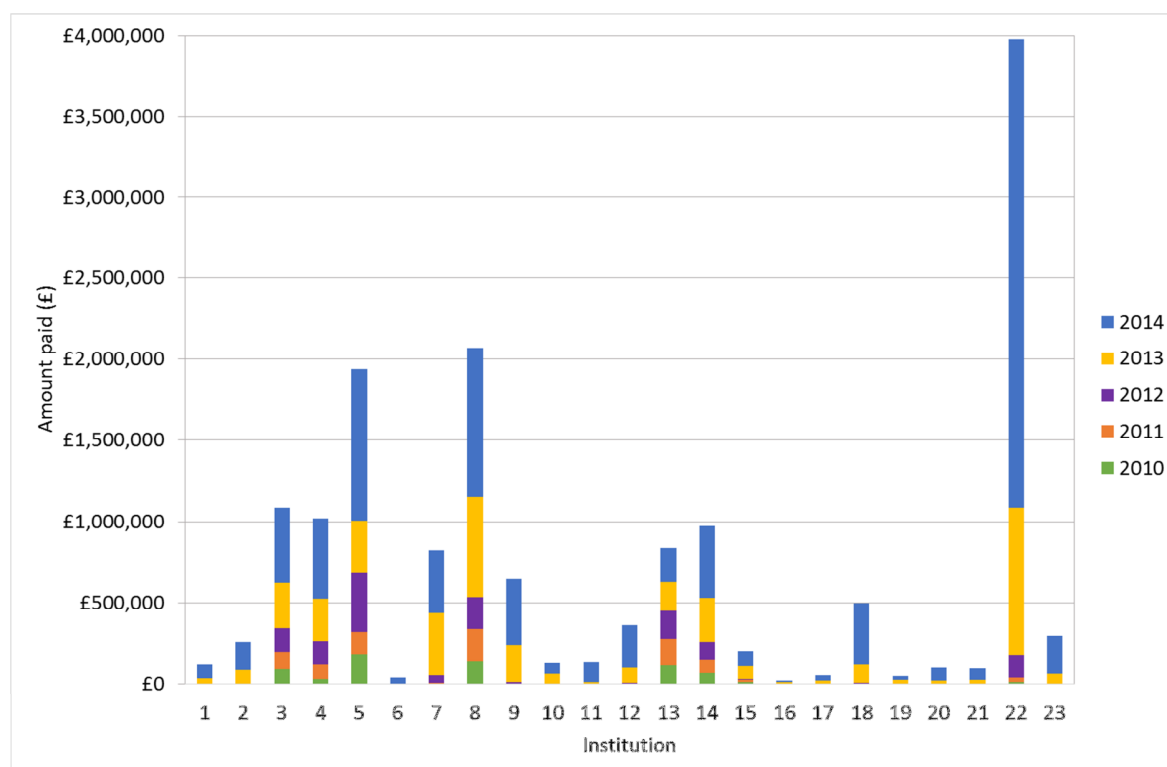
The previous study by Pinfield et al. (2015) found there had been a marked rise in centrally-managed APC payments from 2012 onwards in the 23 sample HEIs. The latest data (Figure 10) from the same institutions show this rise continued in 2014 approximately in line with the total as projected in the previous study. In 2014, the 23 HEIs spent a total of £8,806,723 on centrally-managed APC payments. This amounts to a 550% rise in expenditure since 2012, flowing from an increase in the number of payments. It is reasonable to assume that large-scale increases will continue in the next three years as compliance rates for RCUK and COAF-funded research outputs increase.

Figure 10: Centrally-managed APC payments for 23 institutions for items published 2010-2014



The growth trends of APC expenditure by institution for the 23 HEIs are illustrated in Figure 11 in anonymised form. It is apparent that the APC expenditure was spread very unevenly across institutions, with research-intensive HEIs (e.g. 5, 8, 22) having much higher levels of expenditure. Institution 22 alone was responsible for nearly a third of all expenditure. However, 21 of the institutions experienced a rise in payments between 2013 and 2014, and 12 of these increases were by more than 100%. Two institutions' expenditure showed a very slight decrease, although the very small number of APCs involved meant that no firm conclusions could be drawn from this.

Figure 11: Centrally-managed APC expenditure by institution, 2010-2014



5.3.2 APC expenditure by institution in 2014

The detailed figures for APCs paid for articles published in 2014 gathered from the new sample of 24 HEIs (who agreed to its being reported in non-anonymised form) provide an interesting insight into the current APC market as experienced by HEIs. Direct comparisons cannot be drawn between this dataset and the data from the earlier study since they are from a different set of institutions (albeit with some overlap). There were 4,853 payments in 2014 totalling £7,695,341 (as compared with £8,806,723 for the 23 institutions followed up from the earlier study). Payments ranged from zero (waived payments as part of deals with publishers) to £4,536, with a mean of £1,586. When zero payments were excluded ($n=40$), the mean was £1,599 ($n=4,813$). Payments by institution are shown in Table 17 by institutional 'mission group': Russell Group (large research-intensive institutions), 'Pre-92' institutions (other research institutions), 'Post-92' institutions (teaching-led institutions) and 'Specialist' HEI. Payments show marked differences in numbers of payments made from less than 10 (three institutions) to approaching 2,000 (UCL).

Table 17: Centrally-managed APC expenditure by institution for articles published in 2014

Group	Institution	Mean	N	Minimum	Maximum	Sum	Median
Russell Group	Birmingham	£1,387	334	£0	£3,780	£463,221	£1,481
	Bristol	£1,792	277	£115	£3,780	£496,467	£1,800
	Durham	£1,492	99	£500	£2,797	£147,660	£1,560
	Glasgow	£1,638	237	£200	£3,600	£388,180	£1,500
	Imperial	£1,844	495	£205	£3,958	£913,017	£1,800
	Liverpool	£1,783	145	£210	£3,780	£258,466	£1,656
	Newcastle	£1,892	236	£240	£4,248	£446,503	£1,800
	QMUL	£1,322	70	£0	£3,780	£92,549	£1,394
	Sheffield	£1,556	243	£0	£3,780	£378,153	£1,500
	UCL	£1,451	1995	£0	£4,536	£2,893,864	£1,500
Warwick	£1,823	127	£356	£3,884	£231,461	£1,753	
'Pre-92' Universities	Bangor	£1,939	42	£431	£3,360	£81,424	£1,924
	Bath	£1,529	112	£0	£3,900	£171,243	£1,500
	Cranfield	£1,857	19	£842	£2,340	£35,274	£2,084
	Lancaster	£1,465	45	£480	£3,780	£65,945	£1,500
	Leicester	£1,743	70	£552	£3,810	£122,030	£1,644
	Loughborough	£1,413	57	£0	£3,331	£80,567	£1,462
	RHUL	£1,379	7	£785	£2,026	£9,654	£1,243
	Salford	£1,894	18	£600	£2,407	£34,088	£2,146
	Sussex	£1,926	41	£293	£3,780	£78,952	£1,907
Swansea	£1,647	45	£817	£3,780	£74,129	£1,500	
'Post-92' Universities	Plymouth	£1,641	8	£514	£2,934	£13,131	£1,754
	Portsmouth	£1,599	9	£962	£2,245	£14,390	£1,590
Specialist HEI	LSHTM	£1,680	122	£789	£3,808	£204,972	£1,721
	Overall	£1,586	4853	£0	£4,536	£7,695,341	£1,502

Figure 12 shows the 2014 APC data for the same institutions normalised by number of research-active staff. The mean APC expenditure was calculated per member of 'Category A' staff submitted by each institution for the 2014 Research Excellence Framework (REF). Any normalisation, of course, comes with caveats about the extent to which the factor used for normalisation is a reasonable. In this case, it should be observed that different institutions used somewhat different criteria in identifying 'Category A' staff and so some variation might be expected between institutions. Nevertheless, it is a useful 'yardstick'. Interestingly, the results show that large research-intensive institutions like UCL and Imperial, with highest total expenditure levels, also have a higher mean expenditure per member of research-active staff. In addition, LSHTM, a smaller more specialised institution, has relatively high mean expenditure. There is some variability amongst other research-intensive institutions (e.g. Newcastle and Warwick), with post-92 institutions with lower levels. There is likely to be a combination of explanatory factors for this, including varying institutional policies and practices (where certain institutions may actively promote and

support Gold OA compared with others), different disciplinary make-up of institutions (particularly where institutions have large medical schools use of Gold OA may be higher, see below), and possible differences in REF inclusion criteria.

Figure 12: 2014 APC expenditure per member of research-active staff (REF2014 ‘Category A’ staff)

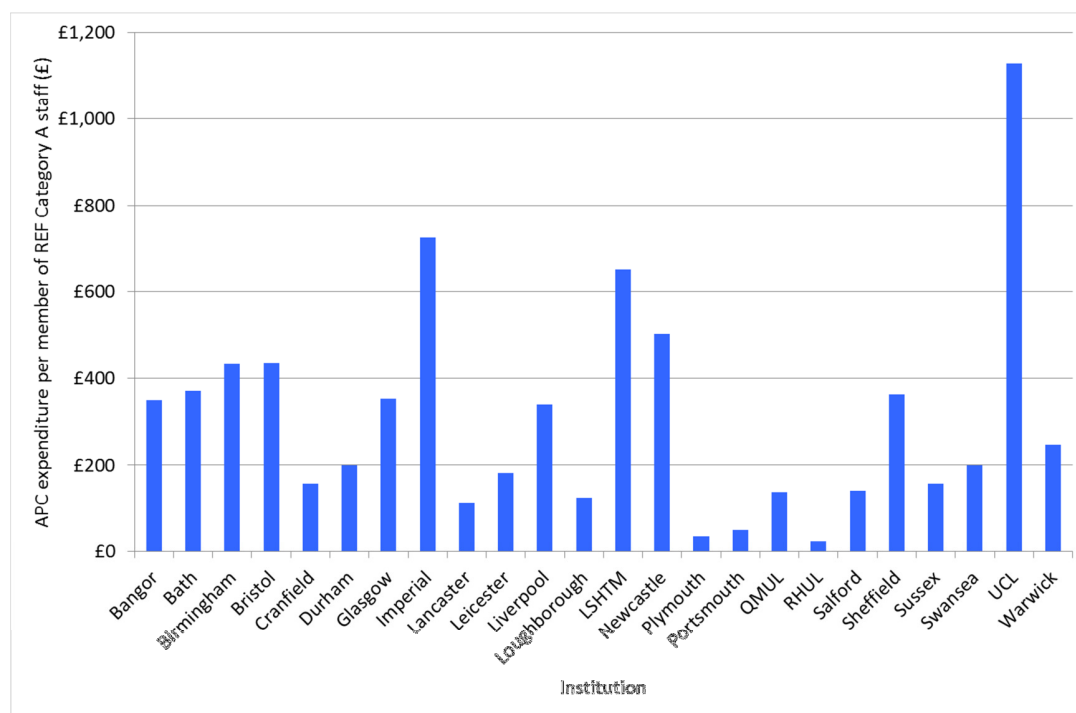
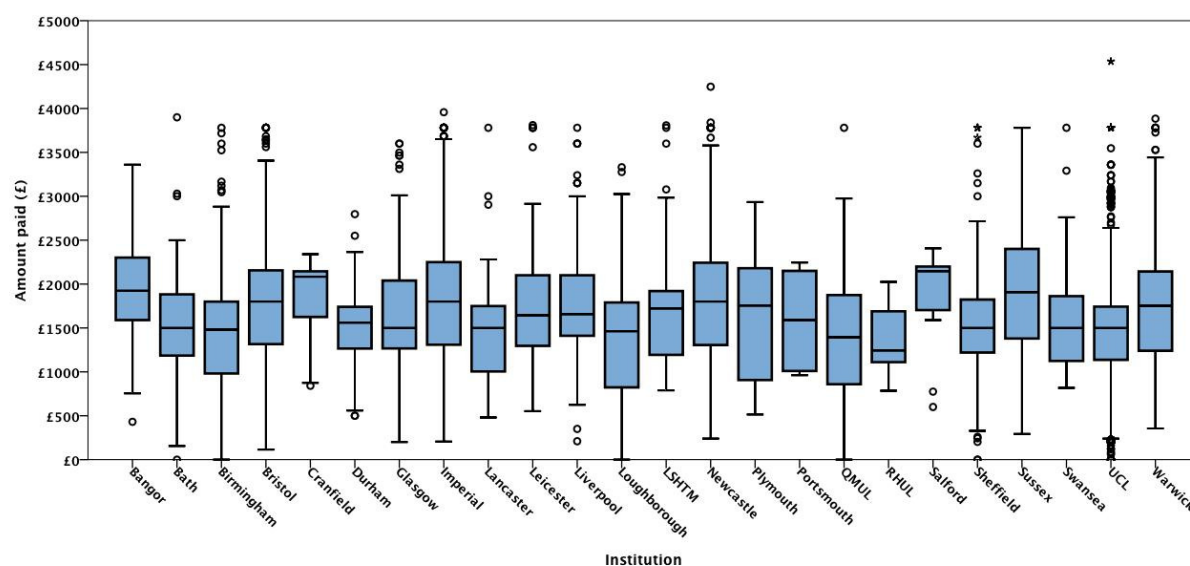


Figure 13 shows the data in a box plot illustrating the general pattern of APC prices paid across institutions. There is a wide range of APC prices paid. The ‘Tukey’ box plot distinguishes the majority of payments from outliers and extreme values. It shows the interquartile range as a boxed area, with the first quartile (25th percentile) as the bottom and the third quartile (75th percentile) as the top of each box. The line in each box is the median (second quartile or 50th percentile). The box extensions (or ‘whiskers’) extend to the furthest data point within the range of 1.5IR, with data points outside these shown as outliers (o) and extreme values (*). The highest payment for a single APC was £4,536, whilst several institutions recorded £0 APC payments.

Figure 13: The range of APC prices paid by institution for articles published in 2014

Only 3,285 of the 4,853 records included a funder. Of these, 2,152 (65% of those recorded) cited the funder as RCUK; 500 (15%) Wellcome; 249 (8%) COAF; and 288 (9%) institutional. The remaining 3% were smaller amounts for a variety of funders. Centrally-managed payments are therefore largely being funded by block grants. With funder preference for licences that allow for liberal reuse (including commercial exploitation), it is unsurprising that 89% (1,909) of APC records in the dataset with the licence field completed (2,146 (44%) of the 4,853 total) were listed as having a CC-BY licence.

The journals for which APC payments were made were mapped against REF panels using subject classifications from Scopus in order to assess their broad disciplinary coverage (Table 18). For the 4,710 of the 4,853 payments that could be matched and verified (97% of the records), there is a clear predominance evident for health and life sciences (just over 60% of the articles and spend). This is higher than the proportions of all papers by UK authors in Scopus (including all organisation types, HE and others), which in 2014 was 49% for Panel A, 32% for Panel B, 14% for Panel C, and 6% for Panel D. APC payments for health and life sciences were, therefore, disproportionately high, and social sciences and arts and humanities, disproportionately low; a similar trend is evident in previous studies of Gold OA.

Table 18: APC payments matched to broad subject area from Scopus, 2014 (N=4,710)

(* sum of the panels is more than the total as some journals are classified into more than one REF panel)

Data for the 24 UK HEIs	Panel A: Health and Life Sciences	Panel B: Physical Sciences and Engineering	Panel C: Social Sciences	Panel D: Arts and Humanities	Total (de-duplicated)
Total spend*	£5,526,217	£2,757,244	£620,368	£115,216	£7,596,649
No of articles*	3337	1701	428	88	4710
Mean	£1,656	£1,621	£1,449	£1,309	£1,611
Min		£0	£71	£71	£0
% spend	61.3%	30.6%	6.9%	1.3%	100%
% articles	60.1%	30.6%	7.7%	1.6%	100%
% of all papers by UK authors	49%	32%	14%	6%	100%

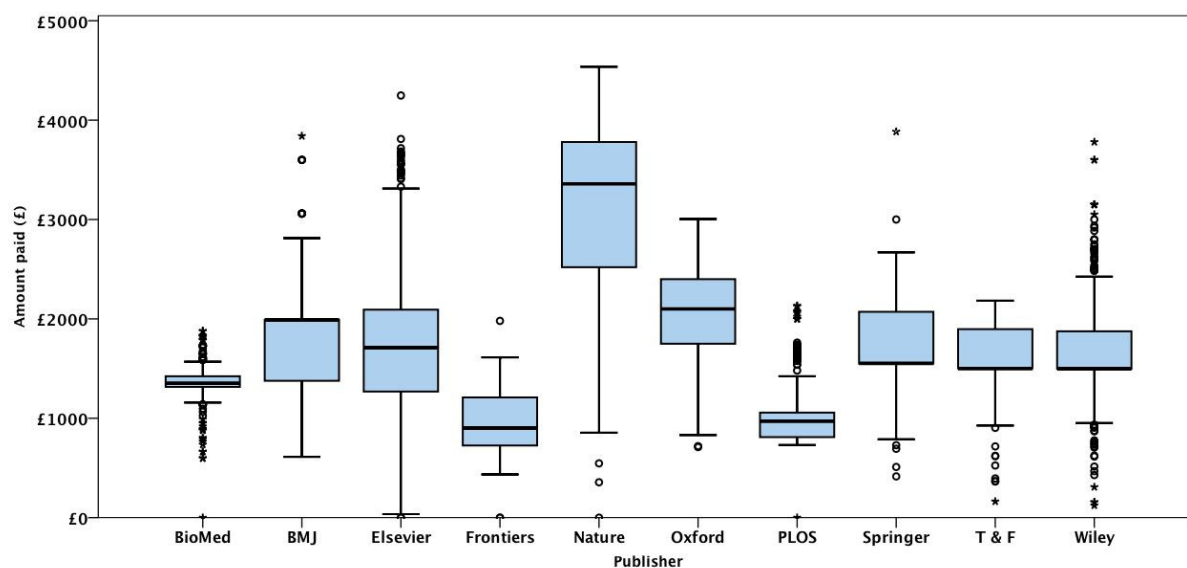
5.3.3 APC payments in 2014 by publisher

Centrally-managed APC payments were made to a total of 128 publishers. Over 70% of the payments were made to the top 10 publishers (Table 19), with Elsevier and Wiley receiving 19% and 15% of payments respectively – very similar proportions to the data covering the period up until Q1 of 2014 reported in Pinfield et al. (2015). More than three-quarters of these payments (76%) were made to hybrid journals. Of the top-10 publishers, three were fully-OA publishers: PLoS, BMC and Frontiers, compared with two (PLoS and BMC) in the previous study. BMC has been treated as a separate ‘fully-OA publisher’ since various factors, not least price, justify a distinction from its parent company, Springer; but it is moot how long such a classification will remain valid. Payments were made to a wide range of journals, with only three titles accounting for more than 1% of all the payments by number: *PLOS ONE* (5.3%), *BMJ Open* (1.5%) and *Nature Communications* (1.4%).

Table 19: Frequency of articles in OA and subscriptions journals among top-10 publishers, 2014 based on APC payments made, with OA breakdown

Publisher	Articles in Fully-OA Journals	Articles in Hybrid Journals	Total (%)
Elsevier	20	906	926 (19.1)
Wiley	25	709	734 (15.1)
Springer	8	329	337 (6.9)
PLOS	322	-	322 (6.6)
BioMed Central	290	-	290 (6.0)
Oxford University Press	28	202	230 (4.7)
BMJ	80	138	218 (4.5)
Taylor & Francis	1	167	168 (3.5)
Frontiers	140	-	140 (2.9)
Nature Publishing Group	34	106	140 (2.9)
Others	232	1116	1348 (27.8)
Total	1180 (24.3)	3673 (75.7)	4853 (100)

Figure 14 illustrates the price range of APC payments for the top-10 publishers. Most publishers charged a relatively wide range of different APC prices. It is noticeable that payments to Nature and Elsevier cover a particularly wide range, with some at very low levels for Elsevier being consistent with information provided by institutions of one-off discounts on APCs provided by Elsevier as part of deals with HEIs. There is also a marked difference in the median price among the different publishers. Two publishers had median APCs below £1,000: Frontiers (£902) and PLoS (£972). Two publishers had median APC levels above £2,000: OUP (£2,100) and Nature (£3,360).

Figure 14: Range of APC payments for the top-10 publishers measured by receipt of APC payments

Analysis of the APC expenditure by journal type shows a marked difference between the mean APC charged by hybrid and OA journals, with hybrids considerably more expensive (Table 20). This is consistent with previous studies (Björk & Solomon, 2014; Pinfield et al., 2015). There is also a difference between fully-OA journals from publishers who also publish subscription titles and those who publish only fully-OA journals (previously observed by Björk & Solomon, 2014). The hybrid mean is 58% higher than the mean of fully-OA journals from ‘non-subscription’ publishers. Of course, comparisons of price only tell some of the story. Journals may offer different levels of service and also deliver different products (most hybrids, for example, provide both paper and electronic formats whereas fully-OA journals do not). These points (as well as price) need to be considered in any full comparison.

Table 20: APC payments by journal types, 2014

(* FWCI: Field-Weighted Citation Index derived from Scopus)

Publisher Type	Mean	Number of journals	Number of articles	Sum	Min	Max	Median	Ave. FWCI*
Hybrid journals – published by ‘subscription publishers’	£1,725	1613	3673	£6,337,723	£0	£4,536	£1,680	1.78
Fully-OA journals – published by ‘subscription publishers’	£1,311	74	306	£401,149	£0	£3,810	£1,229	1.49
Fully-OA journals – published by ‘non-subscription publishers’	£1,094	181	874	£956,469	£0	£2,960	£1,071	1.29

An interesting question arising from these price differences is the relationship between price and quality. To address this, the APC price data were matched against field weighted citation impact (FWCI) scores to test whether there was a correlation between APC price and citation impact, using citation impact as a proxy measure of quality. Initial analysis of the journal types in Table 4 does show a correlation between

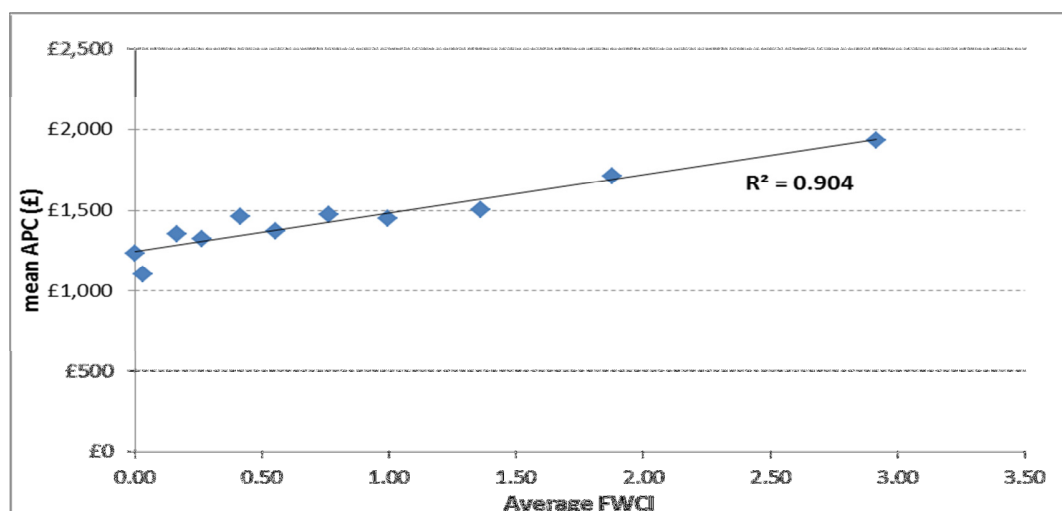
price and citation impact (“Ave. FWCI” column). More detailed analysis is shown in Table 21. Journals were grouped in 10 different FWCI categories for analysis, with all journals covered in Scopus being ordered according to their FWCI and then ranked into tiers, each tier accounting for 10% of the total number of journals, the top tier rated 1 and the bottom tier rated 10. Because the top 10% tier accounted for a large proportion of articles (38%), to provide more granularity this top level was further divided in two, with the top 5% rated 1 and second 5% rated 1.5, making a total of 11 tiers. For each tier, Table 21 shows the numbers of journals and of articles for which APCs were paid from the sample. The proportions of those journals and articles for the whole sample are also given. For example, for Tier 1, APCs were paid for 954 articles in 266 different journals, which constitute 15% of the journals and 20% of the articles covered in the sample. For each tier the weighted average and unweighted average FWCI are also shown.

Table 21: APC prices paid and Field-Weighted Citation Index values (based on Scopus data)

Based on all journals		Based on journals in which 24 UK universities published APC articles in 2014						
Distribution of all journals	Quality Tier (by FWCI)	No of journals with APC articles (from 24 UK HEIs)	No of articles with APCs (from 24 UK HEIs)	Proportion of journals	Proportion of articles	Weighted Ave FWCI	Ave FWCI	Ave APC paid (£) including VAT if charged
5%	1.0	266	954	15%	20%	2.92	3.11	£1,936
5%	1.5	288	864	16%	18%	1.88	1.90	£1,713
10%	2.0	475	1603	27%	34%	1.36	1.37	£1,503
10%	3.0	321	663	18%	14%	0.99	0.99	£1,449
10%	4.0	182	322	10%	7%	0.76	0.76	£1,472
10%	5.0	125	169	7%	4%	0.55	0.56	£1,371
10%	6.0	47	68	3%	1%	0.41	0.40	£1,459
10%	7.0	24	34	1%	1%	0.26	0.25	£1,325
10%	8.0	14	17	1%	0%	0.16	0.15	£1,352
10%	9.0	12	13	1%	0%	0.03	0.04	£1,102
10%	10.0	3	3	0%	0%	0.00	0.00	£1,237

The analysis shows a strong correlation between APC price and FWCI (illustrated in Figure 15). This is consistent with a recent study of APC list prices using different citation indexes and based on list prices (Björk & Solomon, 2015) which also found that highly-cited journals charge higher APCs. Highly-cited journals charging higher APCs may, of course, be explained in different ways. Higher prices may reflect higher costs, and/or authors’ being prepared to pay higher APCs for publication in highly-cited titles. Further work could usefully be carried out to explain this correlation, including more detailed comparisons of hybrid journals and fully-OA journals with similar FWCI scores, and more work on value and cost (not just price).

Figure 15: Mean APC against average Field Weighted Citation Index score for journals, 2014



5.3.5 'Total cost' to universities

Pinfield et al. (2015) sought to gauge the additional costs being experienced by institutions arising from the shift to OA, taking the cost of existing subscriptions plus APCs for hybrid journals and new administration costs. Following David Willetts, the UK Minister for Science and Universities (Willetts, 2014), they referred to this as the 'total cost of publication' (TCP), which has been used as a term in trying to develop an evidence base relating to perceptions of 'double dipping' and also to concerns about additional administrative costs (Pinfield et al., 2015). The measure as framed does not include fully-OA journals since these are seen as an alternative to subscriptions costs rather than an additional cost.

There is no evidence that subscriptions have declined compared with previous years (with most HEIs in the short term, at least, participating in multi-year subscription deals which normally build in agreed year-on-year subscription price rises). And since APCs have been included here at discounted rates, if applicable (therefore taking into account any offsetting which is occurring), there is every reason to assume that, currently, the APCs are examples of additional cost. This is likely to continue for the foreseeable future until institution-level offsetting agreements are more widely adopted and have had time to take effect.

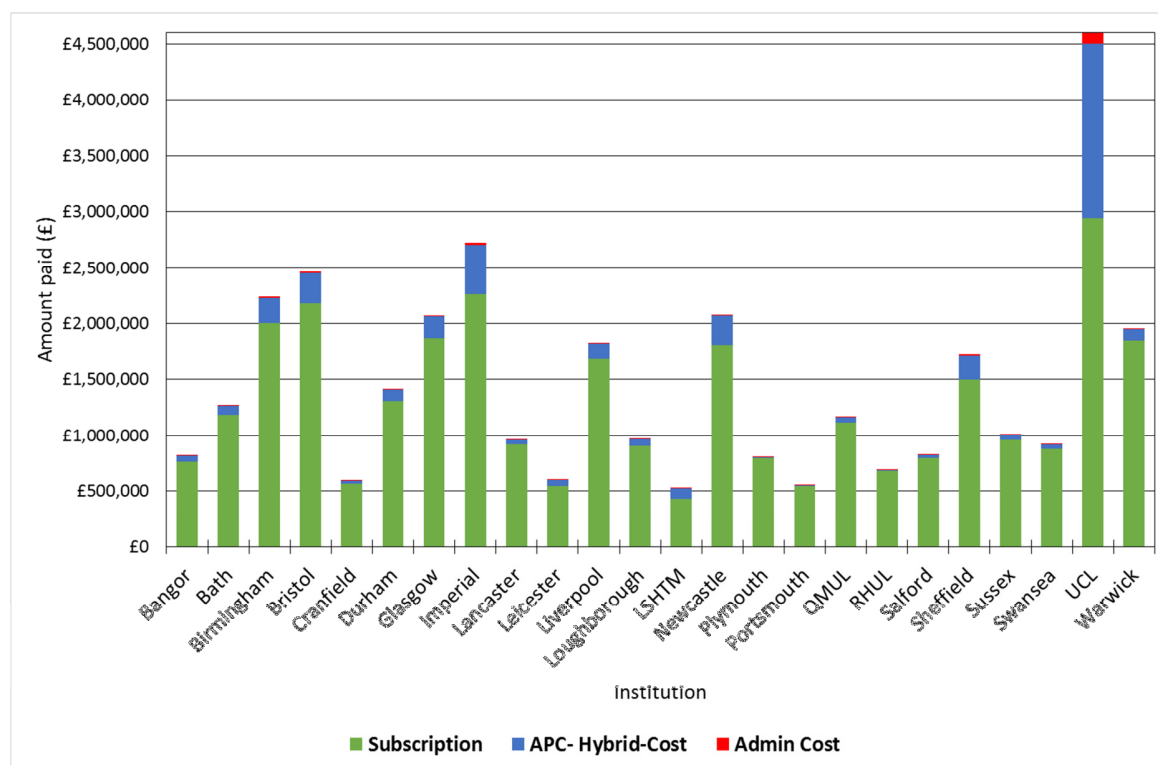
The earlier study can be updated using the current APC data in combination with 2014 subscription data available for seven major publishers (Lawson & Meghreblian, 2014) and administrative cost data from Johnson et al. (n.d.) of £88 per APC. The calculations are shown in detail in Table 22 and illustrated in Figure 16. Across the 24 institutions and seven publishers in the sample, subscriptions constituted in 2014 87% of the cost to universities, APCs 12%, and administrative costs less than 1%. These figures show a slightly higher proportion of costs in the APC category than in the previous study (Johnson et al., n.d.; Pinfield et al., 2015) but are not directly comparable since the earlier study covered all subscriptions paid to publishers to which APCs had also been paid, and was, therefore somewhat wider in its coverage.

Table 22: Costs to universities for seven publishers*, 2014 (excluding fully-OA titles)

(*CUP, Elsevier, OUP, Sage, Springer, Taylor & Francis, and Wiley)

Institution	Subscriptions (%)		APC (%)		Admin cost (%)		Total
	£	%	£	%	£	%	
Bangor	£765,872	93.2%	£53,607	6.5%	£2,200	0.3%	£821,679
Bath	£1,186,086	93.4%	£78,992	6.2%	£4,488	0.4%	£1,269,566
Birmingham	£2,004,295	89.5%	£222,069	9.9%	£14,168	0.6%	£2,240,532
Bristol	£2,181,422	88.5%	£271,226	11.0%	£12,408	0.5%	£2,465,056
Cranfield	£567,832	94.9%	£29,467	4.9%	£1,320	0.2%	£598,620
Durham	£1,308,700	92.7%	£97,268	6.9%	£5,456	0.4%	£1,411,424
Glasgow	£1,871,363	90.3%	£192,080	9.3%	£10,032	0.5%	£2,073,474
Imperial	£2,262,852	83.0%	£443,124	16.3%	£18,744	0.7%	£2,724,720
Lancaster	£919,913	95.6%	£40,053	4.2%	£2,200	0.2%	£962,166
Leicester	£545,000	90.4%	£55,058	9.1%	£2,552	0.4%	£602,610
Liverpool	£1,678,451	91.6%	£146,634	8.0%	£6,864	0.4%	£1,831,950
Loughborough	£903,882	92.9%	£66,003	6.8%	£3,432	0.4%	£973,317
LSHTM	£431,170	80.8%	£98,051	18.4%	£4,576	0.9%	£533,798
Newcastle	£1,806,955	86.7%	£264,885	12.7%	£11,616	0.6%	£2,083,456
Plymouth	£797,744	98.8%	£9,076	1.1%	£352	0.0%	£807,172
Portsmouth	£547,687	98.4%	£8,763	1.6%	£352	0.1%	£556,802
QMUL	£1,117,813	95.8%	£47,055	4.0%	£2,200	0.2%	£1,167,068
RHUL	£683,004	99.0%	£6,425	0.9%	£352	0.1%	£689,782
Salford	£798,763	96.5%	£27,583	3.3%	£1,144	0.1%	£827,490
Sheffield	£1,498,839	87.1%	£211,113	12.3%	£10,208	0.6%	£1,720,160
Sussex	£958,613	94.7%	£51,844	5.1%	£2,288	0.2%	£1,012,745
Swansea	£879,687	95.3%	£41,167	4.5%	£2,200	0.2%	£923,055
UCL	£2,940,492	64.0%	£1,565,022	34.0%	£91,080	2.0%	£4,596,594
Warwick	£1,849,466	94.6%	£100,762	5.2%	£4,312	0.2%	£1,954,540
Total	£30,505,902	87.5%	£4,127,329	11.8%	£214,544	0.6%	£34,847,775

Figure 16: Costs to universities for seven publishers, 2014 (excluding fully-OA titles)



Interestingly, only five HEIs have proportions for APCs above the mean (i.e. 11.8% of the total cost): Imperial (16.3%), LSHTM (18.4%), Newcastle (12.7%), Sheffield (12.3%) and UCL (34%). These research-intensive institutions, particularly UCL, therefore, have a major impact on the overall average. With UCL removed from the calculations, the calculations change somewhat to subscriptions 91%, APCs 8% and administrative costs, less than 1%.

We also analysed the data including APCs for fully-OA as well as hybrid titles. If the payments for fully-OA journals from the same seven publishers are included, the difference to the APCs paid by institutions is marginal. They add only £121,924 expenditure to the total for all of the institutions and therefore have little impact on the overall proportions. However, if the top three fully-OA publishers (PLoS, BMC and Frontiers) are included in the calculation (for which, of course, there are no subscription payments), there is, unsurprisingly, a more marked rise of the proportion of APC payments to the overall costs. The table in Annex F shows this, with subscriptions now 85% of the overall cost, APCs 14% and administration costs 1%.

The value of this kind of approach in calculating institutional costs is limited, of course, when we can use data relating to only a sample of publishers. The value of subsequent work would be considerably increased were data to be available from institutions covering all publishers in receipt of any subscriptions and/or APCs, so that it could be included in the modelling.

5.3.6 Non-centrally-managed APC expenditure

As already observed, the APC data reported by institutions and used in this study include centrally-managed payments only. Whilst this can be reasonably assumed to encompass most RCUK and COAF-funded APCs, other APC payments may in some institutions occur outside the centre. Reliable data on this, however, is not available. During this project, some data was obtained from publishers based on their records of APC payments received from institutions and it was hoped that these records could be compared with data from HEIs of payments made to determine whether there was a substantial difference between the two. However, it proved impossible to reconcile these datasets reliably, particularly since the publishers did not know who had paid an APC (or their institution), only that the invoice was sent to the corresponding author. Very sketchy evidence would, however, seem to indicate that well-established fully-OA publishers which do not have pre-payment deals in place with institutions may receive a much greater number of non-centrally-managed payments than hybrid publishers. It seems that fully-OA journals may have established relationships with authors, and the comparatively low APCs mean that authors are willing to pay from local funding sources. It seems that hybrids, on the other hand, tend to be paid centrally to a greater degree. Such a conclusion is compatible with our findings in Section 3 which shows that fully-OA journals still account for a larger proportion of OA articles than hybrids, while the data in this section indicates a preponderance of payments for hybrids (Table 20). Such observations are, however, highly impressionistic and need further testing.

Summary of key findings

- ❑ The APC market is currently complex (e.g. variable pricing, discounts, other additional charges etc.) and institutional data reflect this
- ❑ Centrally-managed APC expenditure has continued to rise steeply (555% since 2012 for the original sample of 23 HEIs), flowing from an increase in the number of APC payments
- ❑ APC payments in 2014 varied from £0 to £4,536 with the mean £1,586 (from our new sample of 24 HEIs)
- ❑ There was considerable variation in the levels of payments across different institutions, reflecting different levels of research activity as well as policy differences
- ❑ The largest number of APC payments were made for articles in the health and life sciences
- ❑ Commercial publishers are responsible for the largest proportion of the centrally-managed APC market in UK HEIs
- ❑ APCs for hybrid journals are on average more expensive than those for fully-OA titles
- ❑ There is a correlation between APC price and the citation rates of journals, with hybrid journals showing on average the highest citation rates of journals offering immediate OA
- ❑ For a sample of seven subscription-based publishers, APCs in the 24 institutions constitute in 2014 12% of the total costs for institutions compared with 1% for APC administrative costs, and 87% for subscriptions. If the three major fully-OA publishers are added to the sample, APC costs represent 14% of total costs, and subscriptions 85%.

6. Financial sustainability: learned societies

6.1 Background

The potential impact on UK learned societies of the transition to OA was a matter of significant concern to the Finch Group and to societies themselves. The concerns include the extent to which increases in APC revenues might match potential falls in subscription income (as well as other sources of income from rights, advertising and so on), the effect OA might have on membership subscriptions, and administrative costs. Our work here was therefore designed to gather data on the overall income and expenditure – as well as the volumes of journal-related income and expenditure – of UK learned societies, and to develop metrics to allow changes in their financial health to be monitored over time. We hope this work will enable individual societies to contextualise their own positions, and help other stakeholders to understand the potential implications for learned societies of the move to OA.

A summary of our findings is at the end of this section.

6.2. Methodology

Our methodology involved two key stages:

First we identified the number and characteristics of UK learned societies which may be at risk from a move to OA. Thus we:

- ❑ Developed a comprehensive list of UK learned societies which publish scholarly journals or conference proceedings with an ISSN, drawing on a wide range of existing sources
- ❑ Established for each society the number of journals/conference proceedings published, the value of their incoming resources/turnover for the most-recently-reported financial year, and whether they self-publish or contract out to a third party
- ❑ Allocated the societies to a subject area according to the UK REF panels, to allow further analysis by broad subject area.
- ❑ For those societies with a turnover in excess of £10m (which collectively account for 80% of all UK societies' total income), sought to establish the income and the surpluses generated from publishing for the most-recently-reported financial year.

Second we undertook a financial analysis of a sample of 30 learned societies. Thus we:

- ❑ Selected a stratified sample of 25 societies reflecting the characteristics of the whole population, supplemented by a further judgemental sample of five UK societies with high levels of publishing activity.
- ❑ Analysed the financial statements of the selected 30 societies, based on their published financial statements for periods ending in the 2011, 2012 and 2013 calendar years (the latter being the most recent for which data is consistently available).
- ❑ Developed a set of metrics to allow societies' levels of publishing income and expenditure, and overall financial health, to be monitored over time.
- ❑ Used the value of revenues generated from publishing for the sampled societies to extrapolate across the remaining population of societies with <£10m in revenues, and determine an overall figure for UK learned societies' revenues from publishing.

Full details on our methodology can be found in Annex G.

6.3. Context and scope of work

Previous studies have considered such issues as the costs of scholarly publishing for societies (Morris, 2005), their use of publishing surpluses (Baldwin, 2004), and the impact of OA on specific disciplinary areas, particularly in the humanities and social sciences (Dingwall et al, 2014; Darley et al, 2014). However, we are not aware of any previous attempt to develop aggregate measures of UK learned societies' financial health, or to evaluate the overall level of revenues generated from publishing.

Learned societies do not enjoy a specific legal status, and they are highly heterogeneous. Our aim has been to provide a comprehensive picture of these organisations, whose crucial role as 'intermediaries in the process of knowledge exchange' (Dingwall et al, 2014) could be put at risk as a result of the transition to OA. We therefore adopted a broad definition of 'learned society', covering any organisation whose mission includes the development and dissemination of knowledge in a specific discipline or field. As a result, our findings include organisations that might equally be classified as professional bodies or subject associations (e.g. the British Medical Association, the Institute and Faculty of Actuaries, or the Oral History Society), as well as more 'traditional' discipline-based societies (e.g. the Royal Society of Chemistry or the Royal Geographical Society). Our findings are necessarily broad in scope, but where possible we have also sought to identify particular groups of societies that may be disproportionately affected by the move to OA, whether due to size or to disciplinary focus.

Our work is subject to a number of important caveats, which stem from the variable quality of information on societies' publishing revenues, expenditure and surpluses that is disclosed in their Annual Reports. Representatives of learned societies have cautioned that publicly-available information may understate the importance of publishing to societies, since neither the costs nor the income relating to their publishing activities are always clearly delineated in their published Annual Accounts; and our findings should be understood with that in mind. Further information on these limitations to our work can be found in Annex G.

6.4 . Key findings: UK societies that publish peer-reviewed journals and conference proceedings

We have identified a total of 279 UK learned societies that publish academic journals and conference proceedings, out of a total of nearly 600 societies reviewed in the course of our work. In many cases the journals published by these societies are among the leading journals in their field internationally. The societies that publish journals generate total revenues across all their activities of over £1.2 billion per annum.

Based on publishing revenues for the largest 30 societies, and extrapolation from our sample for the remainder, we estimate the total revenues generated from publishing by these 279 societies to be £318m, or 26% of their total income, much of it generated from overseas. It should be noted that many societies do not disclose what proportion of their publishing revenues derive from scholarly journals and conference proceedings, as distinct from trade journals, magazines and monographs. However, review of a sample of 18 societies where this information is disclosed showed that some 90% of total publishing revenues related to scholarly journals in the STEM subjects and the social sciences (Panels A-C), and over 80% in the arts and humanities (Panel D).

Although the number of societies associated with the subjects covered by each of the four main REF panels is roughly comparable, societies in the life and physical sciences (Panels A and B) are typically much larger than those in the social sciences and humanities (Panels C and D), and generate vastly greater revenues, as shown in Table 23. Their published accounts suggest that societies in the physical sciences collectively derive the highest proportion of their income from publishing, which represents over one third

of their total income, compared with one fifth for those in the life sciences and social sciences (Panels A and C). Societies in the arts and humanities (Panel D) generate less than 10% of their revenues directly from publishing, but many consider access to the society journal to be a key driver of membership subscriptions. The relationship between publishing and income from membership subscriptions is complex and it was not within the scope of our work to assess this in detail. Nevertheless, it is important to note that some societies' income from membership subscriptions could also be jeopardised were their journals to be made available in OA form.

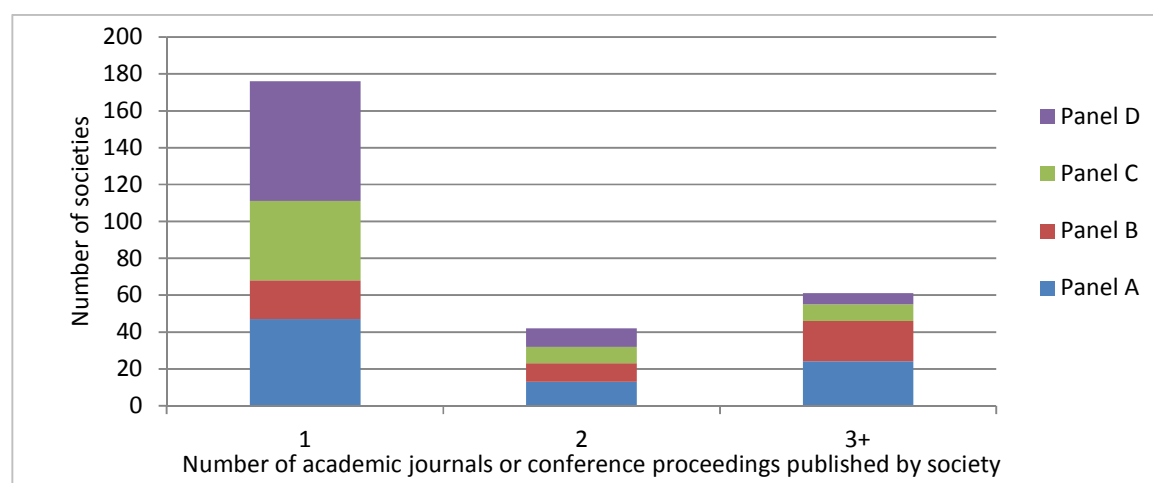
Table 23: UK learned societies that publish peer-reviewed journals/conference proceedings

REF Panel	Number of societies (as % of total)	Total society revenues (as % of total)	Estimated revenues from publishing	Publishing as % of total revenues
Panel A	84 (30%)	£653m (53%)	£137m	21%
Panel B	53 (19%)	£440m (36%)	£157m	36%
Panel C	61 (22%)	£113m (9%)	£22m	19%
Panel D	81 (29%)	£26m (2%)	£2m	9%
Total	279 (100%)	£1,232m (100%)	£318m	26%

6.5. Key findings: Number of publications per society

The challenge faced by societies in responding to market changes in the field of publishing is effectively illustrated by the fact that 176 societies (63%) only publish a single peer-reviewed journal, which typically accounts for most or all of their publishing revenue. A further 42 societies (15%) publish two journals, while 61 (22%) publish three or more journals. Figure 17 shows the number of journals published by learned societies across disciplines, illustrating that those in the social sciences and the arts and humanities are most likely to publish a single journal.

Figure 17: Learned Societies' publishing activity



6.6. Key Findings: Society publishing partners

Only 67 of the 279 societies considered in our work publish in-house, and outsourced publications are dominated by a small group of major publishers, shown in Table 24. Anecdotal evidence suggests societies have increasingly sought to partner with third party publishers in recent years. We have not been able to verify this within the scope of our work, but it would be useful to monitor this trend in future years.

Table 24 – Leading publishing partners for UK learned societies

Publisher	REF Panel				Grand Total
	A	B	C	D	
In-house	20	19	10	18	67
Wiley	20	11	15	11	57
Taylor & Francis	7	3	10	12	32
Cambridge University Press	5	4	4	18	31
Oxford University Press	9	4	2	9	24
Maney Publishing	2	2	7	6	17
Elsevier	8	5	1	0	14
Sage	4	3	5	1	13
BMJ	5	0	0	0	5
Other publishers	4	2	7	6	19
Grand Total	84	53	61	81	279

6.7 Key Findings: Financial sustainability metrics

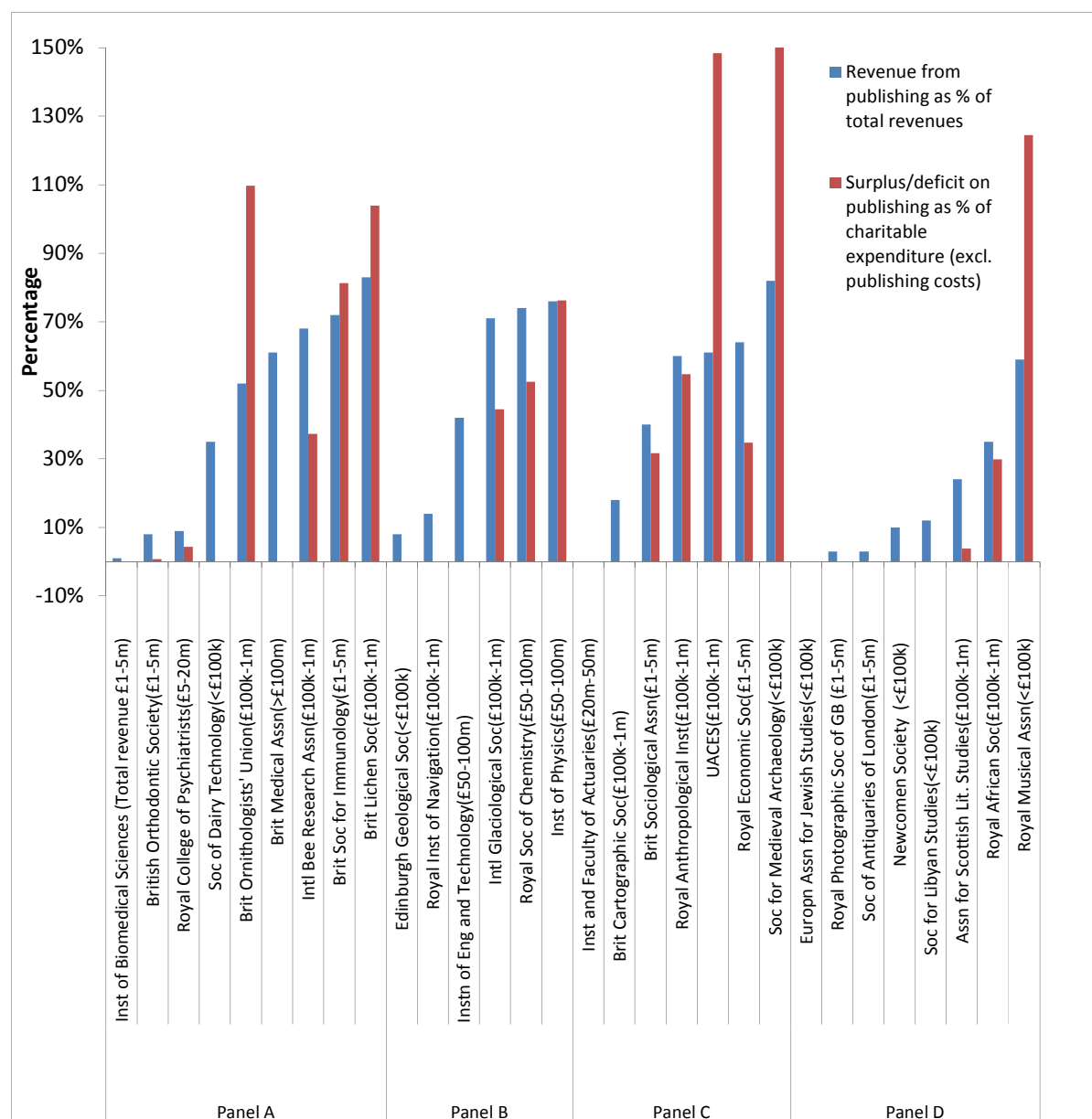
The overall population considered in our work includes many large, diversified membership organisations that derive only a small proportion of their revenues from publishing, while others do not separately disclose any revenues from this source. The summary figures presented above, therefore tend to obscure the much higher level of dependence on publishing for some societies.

For this reason, we considered the published accounts of 30 societies in more detail, including several of the largest society publishers such as the Royal Society of Chemistry, the Institute of Physics, the Institution of Engineering and Technology, and the British Medical Association/BMJ. These very large societies were deliberately selected because of their significant publishing operations. In addition, we selected a sample of 25 societies at random, which collectively derive 27% of their combined revenues

from publishing, consistent with the figure of 26% for the population as a whole. The full list of societies included in our sample can be found in Annex I.

Figure 18 summarises the revenues and surplus/deficit from publishing for each of the 30 societies, as reported in their accounts for financial years ending in 2013. The total revenues for these societies vary widely, and so the table provides an indication of total revenue after each society's name. For comparative purposes, publishing revenues are expressed as a percentage of each society's total revenues; they are of far more significance to some societies than others. There is no simple correlation between the proportion of revenues derived from publishing and a society's size or disciplinary focus. Surplus from publishing is presented as a proportion of charitable expenditure excluding publishing, in order to illustrate the extent to which societies' charitable activity might be put at risk by a reduction in publishing surpluses. Of the 30 societies reviewed, one does not disclose the level of income and expenditure on publishing, nine societies report a deficit on their publishing activities, and a further two report a surplus but, as companies, have no charitable activities. In consequence, results are shown only for the 17 societies which both make a surplus from publishing and report expenditure on charitable activities. For nine of these societies, the surplus from publishing represents more than 50% of their charitable expenditure, indicating that this activity is heavily dependent on the success of the society's publishing operations. In a minority of cases, surpluses from publishing represent more than 100% of a society's charitable expenditure, suggesting that these surpluses are also used to support other activities such as membership services. Based on our sample, societies in the physical sciences (Panel B) and in the arts and humanities (Panel D) are more likely to report a deficit on their publishing activities than those in the life sciences (Panel A) and the social sciences (Panel C), with the latter reporting particularly high levels of surplus on publishing.

Figure 18: Revenue and surplus/deficit from publishing (30 societies, 2013)



We also developed a set of metrics that can be used to monitor changes in levels of publishing revenue, expenditure and overall financial health for these 30 societies over time. The key findings from this exercise are provided below, with the metrics themselves presented in Table 25. Further analysis of this data by REF panel is presented in Annex I. In all cases, the financial values and metrics are presented both as a sum of the sampled societies (where the metrics primarily reflect the position of the largest societies), and as a simple average across the sampled population (which ignores societies' relative sizes, and provides a more balanced representation of the population as whole).

Income and expenditure

- Societies have been successful at growing their revenues from both publishing and other sources over the three years to 2013, meaning publishing revenues as a percentage of total revenues have

remained remarkably consistent for the sampled societies, at 54% in each of the 3 years reviewed. This is higher than for the population as a whole, due to the deliberate inclusion in the sample of several large societies with high levels of publishing such as the Royal Society of Chemistry and Institute of Physics.

- ❑ It is notable, however, that the average net income (surplus/deficit) generated from publishing has not grown in line with revenues, and indeed fell between 2012 and 2013, from 16% to 12% of publishing revenues. This reflects a rapid rise in the level of expenditure on publishing in 2013. It is too early to tell whether this a long-term trend, but it raises the prospect of societies generating a reduced contribution from publishing, whether as a result of OA or other factors, with a consequent reduction in levels of charitable expenditure.
- ❑ Surpluses from publishing equated to 29% of the societies' aggregate charitable expenditure in 2013, down from 44% in 2012. Surpluses as a proportion of charitable expenditure are notably higher for societies in the social sciences (Panel C), at around 50% throughout the period reviewed.

Financial health

- ❑ Most societies, as charities, are obliged by law to take a prudent approach to managing their reserves, setting aside sufficient funds to cover many months' expenditure and ensuring they are able to meet a number of other obligations, such as the upkeep of heritage buildings. Hence it is to be expected that a well-managed society would maintain a high level of reserves and cash holdings relative to its income and expenditure.
- ❑ This is indeed what our figures show, and collectively it is clear that the vast majority of the societies are not only in good financial health, but have steadily improved their position in recent years, with net assets and discretionary funds/reserves rising consistently over the three years reviewed.
- ❑ Cash holding and liquidity measures both indicate that the societies are well-placed to meet their obligations as they fall due, with year-on-year fluctuations remaining within reasonable boundaries, and giving no cause for concern.
- ❑ There is however wide variation within the sample, most notably between the different subject areas. While societies in the life sciences and physical sciences (Panels A and B) are in the strongest financial position, those in the social sciences and in the arts and humanities (Panels C and D) typically have lower financial reserves and operate on fine financial margins, with several reporting losses in recent years.

Table 25 - Financial Analysis – All Disciplines	2013	2012	2011		2013	2012	2011
	Sample of 30 societies (Sum)				Sample of 30 societies (Average)		
	£000s	£000s	£000s		£000s	£000s	£000s
<i>Income and expenditure</i>							
Total income	361,202	342,323	320,777		12,040	11,411	11,061
Net income	17,391	21,268	17,263		580	709	595
Total income from publishing	193,290	184,970	174,030		6,443	6,166	6,001
Net income from publishing	22,384	29,976	27,180		796	1,065	1,004
Total charitable expenditure (excl. publishing)	76,020	68,721	65,065		2,813	2,543	2,501
Publishing income as % of total income	54%	54%	54%		38%	35%	38%
Net income from publishing as % of total publishing income	12%	16%	16%		17%	21%	21%
Net income from publishing as % of charitable expenditure (excl. publishing costs)	29%	44%	42%		43%	31%	32%
<i>Financial health</i>							
Net assets	486,201	406,885	354,051		16,207	13,563	12,209
Discretionary funds/reserves	345,842	307,593	283,599		11,528	10,253	9,779
Cash at bank and in hand	99,965	94,453	125,520		3,332	3,148	4,328
Discretionary funds/reserves as % of total income	96%	90%	88%		138%	139%	130%
Current ratio	1.2	1.1	1.4		1.2	1.1	1.4
Liquidity (Net current assets as no. of days' expenditure)	34	20	57		311	367	336

6. 8 Conclusions

To date, there is no evidence that OA has had any adverse impact on societies' publishing revenues and overall financial health. This is not unexpected, since the data presented in this report derives from published accounts relating to financial years ending in 2013 (in some cases as early as January or March that year). The Research Councils UK OA policy did not come into effect until April 2013, and initially required only partial compliance for the minority of UK publications supported by RCUK funding. For most societies the impact of UK policies has been further diluted by the high proportion of revenues they generate overseas, and the long-term nature of agreements with commercial publishers. As a result, the impact on societies' finances of the transition to OA may not be evident for a number of years.

Nevertheless, our results indicate a clear link between publishing surpluses and levels of charitable expenditure, which includes support for activities such as grant-making, member services, development of communities of practice, and public education. It is too early to tell whether the fall in surpluses seen in 2013 will continue, but it will be important to monitor this, and the other measures presented in Table 25, in future years.

Summary of key findings

- ❑ Nearly 280 learned societies in the UK publish scholarly journals, and we estimate that out of their total revenues of c£1.2bn, some £318m, or 26% of the total, derives from publishing. In many cases the journals published by these societies are among the leading journals in their field internationally.
- ❑ Societies are evenly distributed across the four subject areas defined by the main REF panels; but both overall and publishing revenues are highly concentrated in the health and life sciences, and physical sciences and engineering.
- ❑ Most societies (63%) publish a single journal, but a sizeable minority (22%) publish three or more. Just under a quarter of societies (24%) publish on their own account, but the majority make use of the services of commercial publishers and university presses.
- ❑ The proportion of revenues derived from publishing varies widely, and there is no simple correlation between the proportion of revenues derived from publishing and a society's size or disciplinary focus. Levels of surplus and deficit from publishing also vary widely, with some societies showing a deficit while societies in the social sciences show on average a strong surplus.
- ❑ Published accounts provide no evidence that up to the end of 2013 OA has had any adverse impact on societies' publishing revenues and overall financial health. This is not unexpected, since RCUK's OA policy came into effect only in April 2013g. Moreover, many societies generate a high proportion of their revenues overseas, and have long-term agreements with commercial publishers. Hence the impact of the transition to OA may not be evident in societies' finances for a number of years. Further work should include more detailed analysis with a sample of societies of their finances, including management and other sets of accounts as well as published annual accounts.

Annex A

Publishers responsible for the most popular journals in which UK authors publish in each of the subject areas covered by the four main REF panels

American Association for the Advancement of Science

American Society for Biochemistry and Molecular Biology

American Chemical Society

American Society of Hematology

American Society of Immunologists

American Institute of Physics

American Physical Society

BMJ

Cambridge University Press

EDP Sciences

Edinburgh University Press

Elsevier

Emerald

Institute of Physics Publishing

MA Healthcare

Modern Humanities Research Association

National Academy of Sciences

Nature Publishing Group

Optical Society of America

Oxford University Press

PLOS

Portland Press

Royal Society of Chemistry

Royal College of Surgeons

Royal Society

Sage

Society for Neuroscience

SPIE

Springer

Taylor & Francis

Wiley

Wolters Kluwer

Annex B - Journals with the highest number of articles published in 2010-2014 by authors with a UK affiliation

A: Health and Life Sciences	FWCI	B: Physical Sciences and Engineering	FWCI	C: Social Sciences	FWCI	D: Arts and Humanities	FWCI
Plos One	1.37	Monthly Notices of the Royal Astronomical Society	1.22	British Journal of Social Work	0.63	Journal of Archaeological Science	5.54
PNAS	3.14	Proceedings of SPIE, the International Society for Optical Engineering	1.07	Sociological Research Online	1.49	Notes and Queries	0.21
BMJ Case Reports	0.07	PNAS	3.14	Criminal Justice Matters	0.35	Philosophical Studies	1.78
Journal of Biological Chemistry	1.22	Physical Review Letters	2.67	Political Quarterly	0.66	Antiquity	4.03
BMJ Online	n/a	Astrophysical Journal	1.58	Geoforum	2.56	Synthese	1.54
Journal of the American Chemical Society	2.76	Astronomy and Astrophysics	1.08	Sociology	2.11	Historical Journal	2.33
Journal of Neuroscience	1.42	Physical Review B Condensed Matter and Materials Physics	1.62	Social Science and Medicine	1.14	Historical Research	1.45
BMJ Open	1.42	Physical Review D Particles Fields Gravitation and Cosmology	1.40	Public Money and Management	0.95	Expository Times	0.16
Nature	7.85	AIP Conference Proceedings	0.45	Royal United Services Institute Journal	0.90	Textual Practice	1.88
Methods in Molecular Biology	0.28	Applied Physics Letters	1.28	World Development	2.33	International Journal of the history of sport	0.83
Neuroimage	1.74	Journal of High Energy Physics	1.41	International Affairs	2.70	Modern Language Review	0.21
Science	2.64	Chemical Communications	2.60	British Educational Research Journal	1.40	Analysis	1.16
BMC Public Health	1.16	Physical Chemistry Chemical Physics	1.48	Third World Quarterly	1.14	Women's History Review	0.68
Nature Communications	3.52	Physical Review A:Atomic Molecular and Optical Physics	1.17	Area	1.46	Shakespeare	0.27
British Journal of Nursing	0.94	Journal of the American Chemical Society	2.76	Work Employment and Society	1.28	Journal of British Cinema and Television	1.31
British Journal of Cancer	1.32	Optics Express	2.28	Parliamentary Affairs	1.70	Journal of American Studies	0.33
Lancet	15.36	Physical Review E: Statistical Nonlinear and Soft Matter Physics	1.39	Oxford Review of Education	0.98	English Historical Review	1.43
Annals of the Royal College of Surgeons of England	0.43	Journal of Applied Physics	1.32	Sport in Society	1.26	Contemporary British History	0.28
Nucleic Acids Research	2.30	Dalton Transactions	1.54	Economic and Political Weekly	0.75	Reflective Practice	1.73
Proceedings of the Royal Society B	1.30	Ceur Workshop Proceedings	n/a	Ethnic and Racial Studies	1.99	Classical Quarterly	1.21
Blood	2.12	Nature Communications	3.52	Sociological Review	1.07	Notes and Records of the Royal Society	0.04
Current Biology	2.13	Proceedings of the Royal Society B: Biological Sciences	1.30	Theory Culture and Society	2.24	Journal of Beliefs and Value	0.30
Biochemical Society Transactions	1.03	Geophysical Research Letters	1.80	Journal of Integrated Care	0.52	Review of English Studies	0.92
British Journal of Hospital Medicine	0.20	Journal of Physical Chemistry C	1.66	Journal of Development Studies	1.42	Journal of Imperial and Commonwealth History	0.89
Journal of Immunology	1.15	Langmuir	1.60	Modern Law Review	1.05	Journal of Postcolonial Writing	1.92

Annex C

APCs and Licensing for OA and hybrid journals

The tables here show data on levels of APCs and on licensing from sub-sets of our sample of 32 publishers:

- a) those that publish fully-OA journals; and
- b) those that publish hybrid journals

Table C1. APCs by band for fully-OA journals from 20 publishers

Publisher	Fully-OA journals	APC band £0-£500	APC band £501-£1,000	APC <£1001 %	APC band £1,001-£1,500	APC band £1,501-£2,000	APC <£2001 %	APC band £2,001-£2,500	APC band £2,501-£3,000	APC band over £3,000	APC > £2000 %
1	516	114	101	42%	255	46	58%	0	0	0	0%
2	392	272	60	85%	60	0	15%		0	0	0%
3	44	0	6	14%	19	16	80%	2	1	0	7%
4	42	22	15	88%	5	0	12%	0	0	0	0%
5	36	3	1	11%	1	1	6%	3	26	1	83%
6	36	7	20	75%	4	5	25%	0	0	0	0%
7	29	20	3	79%	6	0	21%	0	0	0	0%
8	20	20	0	100%	0	0	0%	0	0	0	0%
9	12	5	5	83%	2	0	17%	0	0	0	0%
10	10	0	0	0%	3	6	90%	0	1	0	10%
11	7	0	1	14%	4	2	86%	0	0	0	0%
12	5	1	4	100%	0	0	0%	0	0	0	0%
13	4	0	0	0%	4	0	100%	0	0	0	0%
14	3	0	0	0%	0	3	100%	0	0	0	0%
15	3	1	1	67%	1	0	33%	0	0	0	0%
16	2	0	1	50%	1	0	50%	0	0	0	0%
17	2	0	1	50%	1	0	50%	0	0	0	0%
18	1	0	0	0%	0	0	0%	0	1	0	100%
19	1	0	0	0%	1	0	100%	0	0	0	0%
20	1	0	0	0%	1	0	100%	0	0	0	0%
Total	1166	465	219	59%	368	79	38%	5	29	1	3%

Table C2. APCs by band for hybrid journals from 28 publishers

Publisher	Hybrid	APC band £0-£500	APC band £501- £1,000	% <£1k	APC band £1,001- £1,500	APC band £1,501- £2,000	% <£2k	APC band £2,001- £2,500	APC band £2,501- £3,000	APC band over £3,000	% >£2k
1	1866	4	49	3%	0	1813	100%				0%
2	1638	0	0	0%	449	1023	90%	153	0	13	10%
3	1608	0	0	0%	0	1608	100%				0%
4	1257	0	0	0%	0	1207	96%	50			4%
5	745	2	465	63%	3	275	100%				0%
6	300	0	0	0%	200	100	100%				0%
7	267	0	12	4%	21	233	100%	1			0%
8	230	2	2	2%	7	219	100%				0%
9	212	0	0	0%	0	139	66%	62	11		34%
10	47	0	0	0%	0	0	0%	47			100%
11	44	0	0	0%	0	44	100%				0%
12	43	0	0	0%	0	1	2%	14	28		98%
13	40	1	2	8%	1	36	100%				0%
14	40	0	0	0%	0	40	100%				0%
15	39	0	39	100%			100%				0%
16	10	10	0	100%			100%				0%
17	9	0	0	0%	0	9	100%				0%
18	8	0	0	0%	8		100%				0%
19	8	0	0	0%	0	8	100%				0%
20	5	5		100%			100%				0%
21	4	0	0	0%	4		100%				0%
22	3	0	0	0%	3		100%				0%
23	3	0	0	0%	0	0	0%	3			100%
24	3	0	0	0%	0	3	100%				0%
25	1	0	0	0%	0	0	0%	1			100%
26	1	0	1	100%			100%				0%
27	1	0	1	100%			100%				0%
28	1			0%			0%				0%
Total	8433	24	571	7%	696	6758	95%	331	39	13	5%

Table C3. Licensing for fully-OA journals from 21 publishers

Publisher	OA journals	CCBY as default or requirement	%	Allow publication under CCBY	%
1	516	516	100%		0%
2	392	0	0%	392	100%
3	44	29	66%	15	34%
4	42		0%	42	100%
5	36	15	42%	20	56%
6	36	5	14%	31	86%
7	29	7	24%	19	66%
8	20	20	100%		0%
9	12	0	0%	12	100%
10	10	0	0%	10	100%
11	7	7	100%		0%
12	5	0	0%	0	0%
13	4	4	100%		0%
14	3	3	100%		0%
15	3	0	0%	3	100%
16	2	2	100%		0%
18	2	2	100%		0%
19	1	0	0%	1	100%
20	1	0	0%	1	100%
21	1	1	100%		0%
Totals	1166	609	52%	546	47%

Table C4. Licensing for hybrid journals from 28 publishers

Publisher	Hybrid journals	CCBY as default	%	CCBY offered	%
1	1866	0	0%	1866	100%
2	1638	0	0%	1638	100%
3	1608	1608	100%		0%
4	1257	0	0%		0%
5	745	0	0%	745	100%
6	300	0	0%	300	100%
7	267	32	12%	232	87%
8	230	0	0%	230	100%
9	212	0	0%	212	100%
10	47	0	0%	47	100%
11	44	44	100%		0%
12	43	0	0%	32	74%
13	40	0	0%	40	100%
14	40	0	0%	40	100%
15	39	0	0%	39	100%
16	10	10	100%		0%
17	9	9	100%		0%
18	8	0	0%	8	100%
19	8	0	0%		0%
20	5	0	0%	5	100%
21	4	0	0%		0%
22	3	3	100%		0%
24	3	0	0%	3	100%
25	1	1	100%		0%
26	1	0	0%	0	0%
27	1	1	100%		0%
28	1	0	0%	0	0%
Totals	8430	1708	20%	5437	64%

Annex D

Posting and embargo policies from 31 publishers

Publisher	Subscription and hybrid journals	AAMs 0-6 months						AAMs 7-12 months						AAMs 13-24 months					
		website	IR	Subject repository	% website	% IR	% sub repos	website	IR	Subject repository	% website	% IR	% sub repos	website	IR	Subject repository	% website	% IR	% sub repos
1	2164	1913	0	0	88%	0%	0%		1913	1913	0%	88%	88%				0%	0%	0%
2	2162	2162	292	292	100%	14%	14%		1894	1894	0%	88%	88%		248	248	0%	11%	11%
3	1966	1966	49	49	100%	2%	2%		871	871	0%	44%	44%		1046	1046	0%	53%	53%
4	1597	6	6	6	0%	0%	0%	889	889	889	56%	56%	56%	438	438	438	27%	27%	27%
5	790	790	790	0	100%	100%	0%			790	0%	0%	100%				0%	0%	0%
6	344	344	208	208	100%	60%	60%		136	136	0%	40%	40%				0%	0%	0%
7	309	15	15	15	5%	5%	5%	135	135	135	44%	44%	44%	147	147	147	48%	48%	48%
8	300	300	300	0	100%	100%	0%			0	0%	0%	0%			300	0%	0%	100%
9	227	227	0	0	100%	0%	0%		227	227	0%	100%	100%				0%	0%	0%
10	77	77	77	77	100%	100%	100%				0%	0%	0%				0%	0%	0%
11	66	8	8	8	12%	12%	12%	51	51	51	77%	77%	77%				0%	0%	0%
12	50	50	50	0	100%	100%	0%			0	0%	0%	0%			0	0%	0%	0%
13	47	0	0	0	0%	0%	0%	47	47	47	100%	100%	100%				0%	0%	0%
14	44	0	0	0	0%	0%	0%	44	44	44	100%	100%	100%				0%	0%	0%
15	40	40	0	0	100%	0%	0%		40	40	0%	100%	100%				0%	0%	0%
16	39	39	39	39	100%	100%	100%				0%	0%	0%				0%	0%	0%
17	31	0	0	0	0%	0%	0%	0	0	0	0%	0%	0%	0	0	0	0%	0%	0%
18	16	16	16	16	100%	100%	100%				0%	0%	0%				0%	0%	0%
19	13	13	13	0	100%	100%	0%			0	0%	0%	0%			0	0%	0%	0%
20	10	10	10	0	100%	100%	0%			0	0%	0%	0%			0	0%	0%	0%
21	9	9	9	9	100%	100%	100%				0%	0%	0%				0%	0%	0%
22	8	8	8	8	100%	100%	100%				0%	0%	0%				0%	0%	0%
23	5	0	0	0	0%	0%	0%	0	0	0	0%	0%	0%	0	0	0	0%	0%	0%
24	4	0	0	0	0%	0%	0%	4	4	0	100%	100%	0%			0	0%	0%	0%
25	3	3	3	0	100%	100%	0%			3	0%	0%	100%				0%	0%	0%
26	3	0	0	0	0%	0%	0%	0	0	0	0%	0%	0%	0	0	0	0%	0%	0%
27	1	0	0	0	0%	0%	0%	0	0		0%	0%	0%	0	0	0	0%	0%	0%
28	1	1	0	1	100%	0%	100%				0%	0%	0%				0%	0%	0%
29	1	1	0	0	100%	0%	0%		1	1	0%	100%	100%				0%	0%	0%
30	1	0	0	0	0%	0%	0%	0	0	0	0%	0%	0%	0	0	0	0%	0%	0%
31	1	0	0	0	0%	0%	0%	1	1	1	100%	100%	100%				0%	0%	0%
Total	10329	7998	1893	728	77%	18%	7%	1171	6253	7042	11%	61%	68%	585	1879	2179	6%	18%	21%

Annex E

1. Methodology and quality assurance for the assessment of accessibility

Two approaches were used for this assessment:

- A census of all sources covered in Scopus to determine the publishing models used. This enabled counts of journals and articles by the main publishing options: Gold (APC-only); Subsidised; Hybrid (APC optional); Subscription; Delayed open access (subscription with delayed open access)
- A sample-based approach to determine the level of Hybrid uptake (where an APC has been paid for immediate open access) as well the level and type of publicly-accessible online postings to various repositories (Green OA)

1.1 Census-based approach

- Scopus publication data were extracted from the SciVal Analytical Services Scopus database, a database snapshot of Scopus data created in November 2014.
- Only publications corresponding to the four major peer-reviewed document types in Scopus were included: ‘Article’, ‘Review’, ‘Conference Paper’ and ‘Short Survey’.
- For each journal covered in the dataset, aggregate counts of publications and citations in appropriate time windows were compiled, and advanced citation indicators (such as field-weighted citation impact) were calculated. This was repeated for publications where at least one of the authors has an affiliation to a UK institution.
- Counts were made for 2012 and 2014. As the snapshot was taken before full 2014 data was processed in Scopus, the full year 2014 values were extrapolated based on Scopus coverage profiles.
- While journals are the predominant “source” for peer-reviewed articles, some are also found in book- or conference proceedings- series. These were also included where covered by Scopus

The **field-weighted citation impact (FWCI)** is the average ratio of actual citation count to expected citation count for any grouping of articles, such as those published in a single source or under a given journal business model. It takes into account the differences in publication and citation behaviour across disciplines, and for the prevalence and citation rates of different document types. A value of exactly 1.00 means that the source is cited at the expected rate, while a value greater than 1.00 means that the output is cited more than expected. FWCI uses a single publication year and up to five years of citation thereafter (or as data currency allows); for example, the FWCI reported for 2009 includes publications in 2009 and citations received in 2009-13 inclusive, while the value reported for 2014 includes publications in 2014 and citations received in 2014-to date; the relative nature of the FWCI means that such shifting windows necessitated by the currency of data and the lagging nature of citation accrual do not alter the validity of the measure.

1.1.2 Classification into REF Panels

- The All Science Journal Classification (ASJC), a higher level of subject classification in Scopus, that indicates the one or more of the 27 subject areas to which each source (journal) is classified, was also extracted from the database. These were then used to assign each source to one or more relevant REF Panels (A-D).

1.1.3. Publishing (or Business) models

- Publishing models were assigned to each journal in Scopus using a combination of the Directory of Open Access Journals (DOAJ; which is comprised mostly of Gold-APC and Gold-noAPC journals) and desk research on publisher price lists and catalogues and individual journal website information on business model (Gold-APC, Gold-no APC, Hybrid or Subscription only) and for the latter two classes, if there is a journal-specific delayed access policy (and if so were classified as Delayed-OA). Also captured at this time was journal-specific manuscript posting policies and embargo periods (see sample-based approach below).
- Further manual checks were done to improve the classification of publishing models as follows:
 - All sources without as assigned publishing model after the first round described above and with article counts greater than 500 in any year in the period 2005-2014 (some 250 different sources) were assigned using individual journal website information.
 - The 50 largest journals assigned as Gold-APC journals were also spot-checked to ensure the assignment was correct.
 - All remaining titles were assumed to be Subscription.

1.2. Sample-based approach

- Scopus publication data were extracted from the SciVal Analytical Services Scopus data feed, a syndicated version of Scopus data which permits extraction of article-level metadata with a weekly refresh.
- Only publications corresponding to the four major peer-reviewed document types in Scopus were included: 'Article', 'Review', 'Conference Paper' and 'Short Survey'.

1.2.1. Sampling plan

- For each period to be analysed (Jan-Feb 2015, Jul-Aug 2014, Jan-Feb 2014, and Jan-Feb 2013) all documents with publication dates in the period were extracted from the data feed and were randomised as follows: each document was assigned a random number and then sorted from smallest to largest, then assigned a fresh random number and so on until the process had been repeated 3 times. The required number of documents for each sampling period (according to a pre-prepared sampling plan designed to ensure good coverage of all four REF Panel main subject areas) was then taken from the top of the final sorted list (see sampling plan in Table A2.1). This was repeated for publications where at least one of the authors has an affiliation to a UK institution; for reporting, publications where at least one of the authors has an affiliation to a UK institution in the global sample were grouped with this UK sample to increase robustness. Note that both the global and UK samples were over-sampled for REF Panels C and D relative to their actual representations in Scopus to allow robust analysis of all four REF Panels; in reporting totals, these were reweighted to the appropriate actual global and UK proportion in the population in Scopus.
- The REF Panel, business model and journal-specific manuscript posting policies and embargo periods assigned to each source in the Scopus database for the census-based approach (see above) was applied to the sample data to ensure consistency and comparability across all analyses.
- Since the sample was purposefully created in such a way to oversample on some panels to allow assessments at Panel level, the results were weighted to account for this when in aggregated totals.

Table A2.1. Sampling plan for global and UK samples.

Time post-publication	REF Panel	Global sample	UK sample
1-2m	A	722	401
	B	885	492
	C	382	212
	D	356	170
7-8m	A	722	401
	B	885	492
	C	382	212
	D	355	186
13-14m	A	722	401
	B	885	492
	C	382	212
	D	384	203
25-26m	A	722	401
	B	885	492
	C	382	212
	D	359	205
TOTAL		9,410	5,184
<i>Subtotals by time post-publication</i>	1-2m	2345	1275
	13-14m	2344	1291
	25-26m	2373	1308
	7-8m	2348	1310
<i>Subtotals by REF Panel</i>	A	2888	1604
	B	3540	1968
	C	1528	848
	D	1454	764
<i>Sampled population by REF Panel</i>	A	31%	31%
	B	38%	38%
	C	16%	16%
	D	15%	15%
<i>Actual population (in Scopus) by REF Panel</i>	A	46%	49%
	B	42%	32%
	C	8%	14%
	D	3%	6%

1.2.2. Search and coding

- An algorithm was created to derive from each sampled document key metadata elements (including the document title) and an automated Google query was designed to replicate human search behaviour: from only the top 10 links returned by the Google search, unique links were stored. Of these links, those indicating HTTP response status codes (e.g. 404 Not Found and 502 Bad Gateway) and domains shown through manual verification to never contain full-text copies of sample documents were removed.
- The remaining links were stored in a database and marked by a temporary workforce of 5 individuals trained to code each link as representing a full-text version of the document in question (marked 'TARGET') or not (marked 'NOT TARGET'). A link was marked as 'TARGET' were if two criteria were fulfilled: the document there found was (a) recognisably the same as the published article being searched for, typically indicated by the article title (with caution exercised for very generic titles or with very similar titles), and (b) a full-text copy of the document (not an abstract or just the first page).
- Those links identified as 'TARGETs' were then coded by experienced full-time staff trained to differentiate between:

- (a) **Preprint (PP)**, defined as author's versions of publications prior to submission to a journal for peer review.
 - (b) **Accepted author manuscript (AAM)**, which have undergone peer review and incorporate any revisions required for acceptance by a journal.
 - (c) **Versions of Record (VoR)**, also known as published journal article (PJA), typically easily recognised by having the journal and/or publisher's logo or running head, professional typesetting, and journal-specific markers such as volume and issue numbers and pagination.
- However, the differentiation between PP and AAM versions is notoriously difficult and depends on often subtle markers in the text of a document; the guiding principle used was that versions lacking any indication that they have been accepted for publication in a journal were classed as pre-prints (and includes working papers in fields where these are used), while those showing some indication that they have been accepted for publication in a journal were classed as accepted author manuscripts. Often, the latter have watermarks or text on the title page making their status obvious, but if the acknowledgements section (where present) mentioned the contributions of anonymous peer reviewers to the improvement of the manuscript, this was deemed to constitute evidence that the paper had passed peer review and so should be considered as an accepted author manuscript. The same staff also assigned each link's root domain to a website class (e.g. publisher website, social sharing network, etc.) and verified for Hybrid journals where publication availability represented Hybrid Uptake. Any publication availability representing Promotional access, under which a publication is made available (often temporarily) for promotional reasons, was coded as Subscription/Free. For each 'TARGET' document, all versions and locations in which it appears were recorded: for example, a publication for which a PP version appears on an author's homepage and an AAM is deposited at a subject repository will both be recorded, but of course de-duplicated in aggregated counts where necessary in subsequent analysis.

1.2.3. Criteria for determining legitimacy of posted content

- Adherence to online posting policy was assessed on the basis of **what** was posted, **when** it was posted and publicly accessible, and **where** it was posted compared to the journal-level policy:
 - **What** - The version of the article that was posted and accessible: (i) the preprint (PP), accepted author manuscript (AAM), or (iii) Version of Record (VoR).
 - **When** - postings were considered to be in policy if posted and accessible no earlier than 25% before the expiry of the relevant embargo period; for example, an AAM from a journal with a 12 month embargo found online earlier than 9 months was deemed to be in contravention of policy, while those posted after this point were deemed to be in compliance.
 - **Where** - Posting locations were checked against policy on the basis of the assignment of the website class (e.g. publisher website, social sharing network, Institutional or Subject repository, institutional or departmental or personal websites, etc.).

Annex F

Costs to universities for seven ‘subscription’ and three ‘fully-OA’ publishers*, 2014 (including fully-OA titles)

(* ‘Subscription’ publishers: CUP, Elsevier, OUP, Sage, Springer, Taylor & Francis, and Wiley; ‘Fully-OA’ publishers: BMC, Frontiers and PLoS)

Institution	Subscriptions (%)		APC (%)		Admin cost (%)		Total
	£	%	£	%	£	%	
Bangor	£765,872	91.8%	£65,248	7.8%	£2,992	0.4%	£834,112
Bath	£1,186,086	92.3%	£93,492	7.3%	£5,368	0.4%	£1,284,946
Birmingham	£2,004,295	85.4%	£321,589	13.7%	£21,296	0.9%	£2,347,180
Bristol	£2,181,422	86.3%	£329,420	13.0%	£16,720	0.7%	£2,527,562
Cranfield	£567,832	94.9%	£29,467	4.9%	£1,320	0.2%	£598,620
Durham	£1,308,700	92.0%	£107,990	7.6%	£6,248	0.4%	£1,422,938
Glasgow	£1,871,363	88.9%	£221,726	10.5%	£12,144	0.6%	£2,105,233
Imperial	£2,262,852	79.7%	£549,145	19.3%	£26,400	0.9%	£2,838,397
Lancaster	£919,913	95.0%	£46,116	4.8%	£2,728	0.3%	£968,757
Leicester	£545,000	86.4%	£81,229	12.9%	£4,312	0.7%	£630,541
Liverpool	£1,678,451	89.4%	£190,224	10.1%	£9,592	0.5%	£1,878,267
Loughborough	£903,882	92.8%	£66,793	6.9%	£3,520	0.4%	£974,195
LSHTM	£431,170	73.9%	£144,473	24.8%	£7,920	1.4%	£583,563
Newcastle	£1,806,955	84.3%	£320,804	15.0%	£15,488	0.7%	£2,143,247
Plymouth	£797,744	98.5%	£11,403	1.4%	£616	0.1%	£809,763
Portsmouth	£547,687	97.5%	£13,299	2.4%	£704	0.1%	£561,690
QMUL	£1,117,813	94.5%	£60,943	5.2%	£3,872	0.3%	£1,182,628
RHUL	£683,004	98.8%	£7,668	1.1%	£440	0.1%	£691,113
Salford	£798,763	96.3%	£29,773	3.6%	£1,320	0.2%	£829,856
Sheffield	£1,498,839	85.1%	£249,496	14.2%	£13,288	0.8%	£1,761,623
Sussex	£958,613	93.3%	£65,378	6.4%	£3,080	0.3%	£1,027,071
Swansea	£879,687	94.9%	£44,431	4.8%	£2,464	0.3%	£926,583
UCL	£2,940,492	58.8%	£1,938,816	38.8%	£121,440	2.4%	£5,000,749
Warwick	£1,849,466	93.5%	£122,903	6.2%	£5,808	0.3%	£1,978,177
Total	£30,505,902	85.0%	£5,111,827	14.2%	£289,080	0.8%	£35,906,809

Annex G

Methodology for the assessment of financial sustainability for Learned Societies

The methodology followed in our work on learned societies can be summarised as follows:

Step 1: For the purposes of our study, we developed a comprehensive list of potential organisations for inclusion from the following sources:

- ❑ UK learned societies listed by Europa World of Learning;
- ❑ A list of learned society members supplied by the ALPSP;
- ❑ The list of UK learned societies found on Wikipedia;
- ❑ The British Academy Directory of Subject Associations and Learned Societies in the Humanities and Social Sciences;
- ❑ The approved list of professional organisations and learned societies identified by HM Revenue & Customs (specifically those in the list that were identified as allowing reclaimed tax on journal subscriptions and other publications);
- ❑ The list of members of the Academy of Social Sciences.

Step 2: From a consolidated list of nearly 600 societies, we identified those societies with their primary, registered address in the UK

Step 3: Among the identified UK learned societies, we selected only those that publish academic journals or conference proceedings (i.e. peer-reviewed publications with an ISSN).

Step 4: We then identified the number of journals/proceedings published by each society (societies publishing only one journal, those publishing two journals and those publishing 3 or more journals), and the value of their incoming resources/turnover for the most recent available financial year. For those with a turnover >£10m we also recorded the value of their income from publishing.

Step 5: We established how many self-publish journals and how many are contracted out, and recorded the identity of any third party publishing partner.

Step 6: We categorised the societies by discipline using the classifications adopted by the UK's Research Excellence Framework (REF)¹⁷ (an indicative classification only, given that several societies have a multi-disciplinary focus).

Step 7: From the list of UK learned societies producing academic publications, we selected a stratified sample of 25 organisations reflecting the characteristics of the broader population of LS¹⁸, supplemented by a further judgemental sample of 5 UK societies with high levels of publishing activity.

Step 8: We analysed the financial statements of the selected 30 LS, based on the published financial statements from the 2013 calendar year (which is the most recent year for which data is consistently available).

¹⁷ REF classification: disciplines falling under panel A (medicine and biological sciences), panel B (maths, physics, natural sciences and engineering), panel C (social sciences) and panel D (arts and humanities)

¹⁸ Characteristics reflected in the sample: (a) the different levels of publishing activity by LS; (b) the overall number of active LS publishers across the four groups of academic discipline used in the REF; (c) the proportion of large and small LS; (d) representative proportion of societies operating with (22) and without (8) a publishing partner.

Step 9: Combining the results for the total population with extrapolated results from our sample, we estimated the proportion of societies' total income which is derived from publishing.

Limitations in the availability and reliability of financial data

We chose to draw on published financial information to complete our work since virtually all societies, whether registered charities or companies, are required to provide this information annually, and make it publicly available. Statutory financial statements must be prepared in accordance with generally accepted accounting practice (GAAP) and, in the case of those societies which are charities, the Charity Commission's revised Statement of Recommended Practice (SORP 2005). Nevertheless, it is important to acknowledge that the information disclosed on societies' publishing revenues varies in scope and quality, and is often not directly comparable between societies. The level of publishing revenues disclosed depends not only on the total income generated by a journal, but also on the precise terms of the agreement between the society and any third-party publisher. For example, in some cases a third-party publisher may only pass the net revenues generated by a journal on to a society, meaning the total value of subscriptions revenue is not reflected in the society's accounts. Many publishers also operate websites and provide other services to learned societies which may be invisible from an accounting perspective, but can be of vital importance in practice. Finally, practices in cost and overhead allocation are also highly variable, and these could have a significant bearing on the figures reported for the surpluses generated from publishing

In a small number of cases, some relevant information on the sampled societies was unavailable, particularly in the case of measures such as expenditure on publishing, and income/expenditure on peer-reviewed journals. In such cases we either used the best available data, or excluded the society in question from some elements of the analysis.

Annex H

Financial Analysis of Sampled Learned Societies

This appendix presents the results of our in-depth analysis of the finances of 30 learned societies over the period 2011 to 2013. The data used in this analysis has been derived from published financial statements for periods ending in each of the three years reviewed, and has been summarised both in aggregate and by REF panel.

A set of key financial values and metrics were identified to allow the large amount of data gathered to be analysed effectively, and have been categorised as either 'income and expenditure metrics' or 'financial health metrics'. The chosen financial values and metrics for income and expenditure are as follows:

- ❑ **Total income** – defined as total incoming resources or total revenue
- ❑ **Net income** – defined as net incoming resources, equal to total incoming resources less total resources expended; or operating deficit/surplus, equal to total income less total expenditure. Accordingly, net income does not include other recognised gains/losses such as gains/losses on investment assets, finance income, taxation and staff pension scheme, etc.
- ❑ **Total income from publishing** – defined as total income from sales of peer-reviewed journals, monographs and other publications; journal royalties and online journal subscriptions. Income from member subscriptions is only included in this amount in a small number of cases where access to the peer-reviewed journals is deemed to be the primary benefit of membership.
- ❑ **Total charitable expenditure (excl. publishing)** – defined as resources expended on charitable activities, less total publishing expenditure.
- ❑ **Publishing income as % of total income** – defined as the ratio (expressed as a percentage) of total income from publishing to total income.
- ❑ **Net income from publishing as % of total publishing income** – defined as the ratio (expressed as a percentage) of net income from publishing to total income from publishing, where net income from publishing is defined as total income from publishing less total expenditure on publishing. Total expenditure on publishing is calculated as the sum of journal expenditure, other publication costs and costs associated with online journal subscriptions.
- ❑ **Net income from publishing as % of charitable expenditure (excl. publishing costs)** – defined as the ratio (expressed as a percentage) of net income from publishing to total charitable expenditure (excluding any publishing costs included under this heading)

Those for financial health are:

- ❑ **Net assets** – defined as total assets (fixed assets and current assets) less total liabilities/creditors
- ❑ **Discretionary funds/reserves** – defined as total unrestricted funds excluding any designated funds at the financial year end
- ❑ **Cash at bank and in hand** – defined as sum of cash at bank and in hand, cash held in liquidity funds and short-term deposits
- ❑ **Discretionary funds/reserves as % of total income** – defined as the ratio (expressed as a percentage) of discretionary funds/reserves to total income
- ❑ **Current ratio** – defined as the ratio (expressed a number) of total current assets to total current liabilities/creditors (amounts falling due within one year) at the financial year end
- ❑ **Liquidity** - Net current assets expressed as number of days' expenditure

Financial analysis – Sample of 9 learned societies (REF Panel A)

Income and expenditure

- ❑ The aggregate (sum) results for REF Panel A are dominated by the British Medical Association which, with a turnover of £126m, accounts for over 80% of the total revenues for the 9 sampled societies in this panel. Though primarily a professional association and trade union rather than a learned society, the BMA's stated aims include the promotion of the medical and allied sciences, and it was selected for inclusion in our sample given the BMJ's importance as a society publisher.
- ❑ The BMJ accounted 61% of the BMA's total revenues in 2013, and this contributed to the high overall percentage of publishing as a % of total revenues in our sample (54% by aggregate value). When the BMA is excluded, the figure is only 21% by aggregate value, but 43% as a simple average.
- ❑ In common with the wider sample, societies in Panel A have seen steady growth in their publishing revenues over the period 2011-2013, but experienced a sharp drop in the surpluses generated in 2013.
- ❑ While the medical societies included with Panel A are typically relatively large and have diversified income streams, the sample does include a number of smaller societies in the biological and veterinary sciences, which are highly dependent on publishing revenues. For example, the British Lichen Society derived 83% of its revenue from a single journal, *The Lichenologist*, in 2013.
- ❑ Net income from publishing accounts for a high proportion of charitable expenditure for Panel A societies, at between 40% and 60%.

Financial health

- ❑ The Panel A societies are collectively in the strongest financial position of those sampled, with most holding discretionary reserves that are twice their annual income, high levels of liquid assets, and a strong balance sheet position.
- ❑ In common with the sample as a whole, the financial health of the societies has strengthened over the period 2011-2013, with almost all measures showing a stable or upward trend.
- ❑ In general, societies in panel A are not at any immediate risk from a decline in their publishing revenues, and even the smallest societies sampled maintain significant reserves which could be used to cushion the impact of changes in the short-term.

Financial Analysis – REF Panel A	2013	2012	2011		2013	2012	2011
(Societies in the fields of medicine, health, biological sciences, agriculture, veterinary and food sciences)	Sample of 9 societies (Sum)				Sample of 9 societies (Average)		
	£000s	£000s	£000s		£000s	£000s	£000s
<i>Income and expenditure</i>							
Total income	150,087	144,478	138,051		16,676	16,053	15,339
Net income	2,383	6,624	4,883		265	736	543
Total income from publishing	80,279	75,822	72,235		8,920	8,425	8,026
Net income from publishing	7,892	11,972	11,628		984	1,494	1,451
Total charitable expenditure (excl. publishing)	19,850	19,428	19,580		2,827	2,768	2,790
Publishing income as % of total income	53%	52%	52%		43%	38%	39%
Net income from publishing as % of total publishing income	10%	16%	16%		17%	27%	39%
Net income from publishing as % of charitable expenditure (excl. publishing costs)	40%	62%	59%		48%	44%	60%
<i>Financial health</i>							
Net assets	157,352	129,680	116,807		17,484	14,409	12,979
Discretionary funds/reserves	112,948	108,761	103,399		12,550	12,085	11,489
Cash at bank and in hand	34,823	31,535	57,056		3,869	3,504	6,340
Discretionary funds/reserves as % of total income	75%	75%	75%		190%	199%	178%
Current ratio	1.1	1.0	1.6		1.2	1.1	1.4
Liquidity (Net current assets expressed as number of days' expenditure)	12	6	78		470	497	483

Financial analysis – Sample of 6 learned societies (REF Panel B)

Income and expenditure

- ❑ The Panel B sample includes three very large societies, the Institution of Engineering and Technology, the Royal Society of Chemistry and the Institute of Physics, which account for 99% of the total revenues across the 6 sampled societies. These three societies are each highly dependent on publishing revenues, and were selected for inclusion in the sample on this basis.
- ❑ Due to the presence of these three societies, publishing accounts for almost two-thirds of the total revenues generated within the sample (compared with only 26% for Panel B societies nationwide). The sampled societies have also seen steady growth in their publishing revenues over the last 3 years, but publishing as a proportion of total revenues in Panel B has fallen slightly since 2011, both in sum and on average.
- ❑ The margins generated from publishing also fell substantially between 2012 and 2013.
- ❑ The three smaller societies sampled (Edinburgh Geological Society, Royal Institute of Navigation, and the International Glaciological Society) generate highly variable levels of income from their journal publications, and all three reported a loss from their publishing activities in the 2011 and 2012 years.
- ❑ Net income from publishing accounts for around a third of charitable expenditure for large societies (though it has fluctuated year on year), but makes a much smaller contribution, if any, to the smaller societies.

Financial health

- ❑ The Panel B societies are generally in good financial health. All six have improved their balance sheet position over the three years reviewed, and most hold discretionary reserves that are roughly equivalent to their annual income.
- ❑ Considered both collectively and individually, there is no cause for concern over these societies' financial sustainability at the present time.

Financial Analysis – REF Panel B	2013	2012	2011		2013	2012	2011
(Societies in the natural sciences and engineering and related fields)	Sample of 6 societies				Sample of 6 societies		
	£000s	£000s	£000s		£000s	£000s	£000s
<i>Income and expenditure</i>							
Total income	173,188	160,124	150,960		28,865	26,687	25,160
Net income	13,526	12,438	11,532		2,254	2,073	1,922
Total income from publishing	110,426	106,446	99,358		18,404	17,741	16,560
Net income from publishing	13,568	17,046	14,745		2,261	2,841	2,458
Total charitable expenditure (excl. publishing)	50,543	43,625	41,974		8,424	7,271	6,996
Publishing income as % of total income	64%	66%	66%		47%	46%	50%
Net income from publishing as % of total publishing income	12%	16%	15%		-17%	-4%	-5%
Net income from publishing as % of charitable expenditure (excl. publishing costs)	27%	39%	35%		20%	17%	13%
<i>Financial health</i>							
Net assets	272,840	224,961	204,184		45,473	37,494	34,031
Discretionary funds/reserves	193,647	161,561	153,578		32,275	26,927	25,596
Cash at bank and in hand	38,932	35,701	42,110		6,489	5,950	7,018
Discretionary funds/reserves as % of total income	112%	101%	102%		121%	115%	99%
Current ratio	1.2	1.0	1.0		1.2	1.1	1.4
Liquidity (Net current assets expressed as number of days' expenditure)	25	-6	7		125	87	85

Financial analysis – Sample of 7 learned societies (REF Panel C)

Income and expenditure

- ❑ The aggregate (sum) results for REF Panel C are highly skewed by the inclusion in our sample of the Institute and Faculty of Actuaries, whose turnover of £29.7m is substantially larger than that of the remaining 6 societies reviewed.
- ❑ The Institute and Faculty of Actuaries' combines an objective 'to advance all matters relevant to actuarial science and its application' with its role in regulating and promoting the actuarial profession', but in practice it acts primarily as a professional body, and derives only a negligible proportion of its revenues from the publication of peer-reviewed journals. Our work identified a number of other large societies in the social sciences (such as the Institute of Fiscal Studies, Royal Institute of International Affairs, and Royal Institute of International Affairs) which derive a similarly low proportion of their revenues from their publishing activities.
- ❑ However, the remaining, much smaller, Panel C societies in our sample derive a very high proportion of their revenues from publishing (over 50% on average if the Institute and Faculty of Actuaries is excluded), and subscriptions to peer-reviewed journal, where disclosed, account for the vast majority (over 80%) of this income.
- ❑ Revenues from publishing have remained stable over the period reviewed, and in contrast to panels A and B, societies in panel C have also seen an increase in the net income generated from publishing over the period.
- ❑ The net income generated from publishing in Panel C, at around 50% of revenues, represents a much higher rate of return on this activity than is seen in other Panels. In consequence, publishing also makes a correspondingly greater contribution to charitable expenditure for societies in the social sciences.

Financial health

- ❑ The Panel C societies are collectively in reasonable financial health, but levels of discretionary reserves have fluctuated in recent years. This reflects the fact that three of the seven societies sampled made a loss in one or more of the years reviewed, and most others made only modest surpluses.
- ❑ Given their high level of dependence on publishing revenues, societies in panel C would appear to be the most at risk from a reduction in levels of income from publishing as a result of OA.

Financial Analysis – REF Panel C	2013	2012	2011		2013	2012	2011
	Sample of 7 societies (Sum)				Sample of 7 societies (Average)		
(Societies in the social sciences and related fields)	£000s	£000s	£000s		£000s	£000s	£000s
<i>Income and expenditure</i>							
Total income	33,650	32,048	29,366		4,807	4,578	4,195
Net income	1,655	1,513	862		236	216	123
Total income from publishing	2,194	2,214	2,116		313	316	302
Net income from publishing	1,140	1,126	980		180	166	154
Total charitable expenditure (excl. publishing)	2,458	2,081	1,887		410	347	315
Publishing income as % of total income	7%	7%	7%		46%	46%	45%
Net income from publishing as % of total publishing income	52%	51%	46%		46%	28%	36%
Net income from publishing as % of charitable expenditure (excl. publishing costs)	46%	54%	52%		66%	38%	40%
<i>Financial health</i>							
Net assets	34,697	32,508	27,527		4,957	4,644	3,932
Discretionary funds/reserves	27,402	26,232	21,570		3,915	3,747	3,081
Cash at bank and in hand	24,853	25,486	25,188		3,550	3,641	3,598
Discretionary funds/reserves as % of total income	81%	82%	73%		103%	123%	116%
Current ratio	2.2	2.2	2.2		1.2	1.1	1.4
Liquidity (Net current assets expressed as number of days' expenditure)	175	191	202		337	440	456

Financial analysis – Sample of 8 learned societies (REF Panel D)

Income and expenditure

- ❑ Societies in panel D are generally the smallest in size, with half of the eight societies sampled generating incoming resources of only £100,000 or less.
- ❑ Publishing generally accounts for a low proportion of total revenues for societies in panel D, reflecting the reduced importance of the journal as a means of scholarly communication in most disciplines included within this panel. The exception to this in our sample is the Royal Musical Association, which derived 59% of its total incoming resources from peer-reviewed journals in 2013.
- ❑ These societies are also unusual in reporting a loss on their publishing activities, both collectively and on average. Instead membership subscriptions (which are generally not treated as publications income in our analysis) are often used to support publication of the society journal(s), and access to the journal represents a more significant incentive to membership than in other fields.
- ❑ In the case of these societies, publishing therefore does not cross-subsidise the societies' other charitable activities, but it is itself subsidized from membership subscriptions, grants and other sources.
- ❑ **Note:** 2011 figures for Panel D exclude the results of one of the societies, the Society of Antiquaries of London, which did not report results in that year. This accounts for the apparent large increase in income and net asset figures between 2011 and 2012.

Financial health

- ❑ In aggregate, the analysis suggests societies in Panel D are in reasonably good financial health. In practice, though, this obscures the fact a small number of the sampled societies possess significant assets, while the others typically have very low levels of financial reserves – four of the eight have discretionary reserves which represent less than 4 months' income.
- ❑ While publishing is not a significant revenue generator for these societies, many are operating on fine financial margins, and thus may find it difficult to absorb even a moderate change in their revenue streams.
- ❑ It was not within the scope of this study to consider the extent to which membership subscriptions for these societies are driven by access to the journal. Nevertheless, in the longer term a move to OA publishing could reduce the perceived value of society membership, and thus have a more disruptive impact on Panel D societies' income than our figures would suggest.

Financial Analysis – REF Panel D	2013	2012	2011		2013	2012	2011
	Sample of 8 societies (Sum)				Sample of 8 societies (Average)		
(Societies in the arts and humanities)							
	£000s	£000s	£000s		£000s	£000s	£000s
<i>Income and expenditure</i>							
Total income	4,277	5,673	2,400		535	709	343
Net income	(173)	694	(14)		(22)	87	(2)
Total income from publishing	391	488	321		49	61	46
Net loss from publishing	(216)	(168)	(173)		(27)	(21)	(25)
Total charitable expenditure (excl. publishing)	3,169	3,587	1,624		396	448	232
Publishing income as % of total income	9%	9%	13%		18%	16%	19%
Net loss from publishing as % of total publishing income	-55%	-34%	-54%		26%	32%	5%
Net loss from publishing as % of charitable expenditure (excl. publishing costs)	-7%	-5%	-11%		32%	21%	1%
<i>Financial health</i>							
Net assets	21,311	19,735	5,533		2,664	2,467	790
Discretionary funds/reserves	11,844	11,039	5,053		1,481	1,380	722
Cash at bank and in hand	1,358	1,730	1,166		170	216	167
Discretionary funds/reserves as % of total income	277%	195%	211%		124%	103%	107%
Current ratio	1.9	2.5	2.9		1.2	1.1	1.4
Liquidity (Net current assets expressed as number of days' expenditure)	76	109	135		250	367	242

Annex I

Sample of 30 Learned Societies

The 30 societies selected for inclusion in our sample are listed below.

<i>Name of Society</i>	<i>REF Panel</i>	<i>Type</i>	<i>Total revenue (£, financial year ended in 2013)</i>	<i>Proportion of total revenues from publishing (% , financial year ended in 2013)</i>
British Lichen Society	A	Academic	120,780	83%
British Medical Association	A	Professional association	125,870,000	61%
British Ornithologists' Union	A	Academic	163,584	52%
British Orthodontic Society	A	Professional association	1,392,383	8%
British Society for Immunology	A	Academic	1,447,897	72%
Institute of Biomedical Sciences	A	Academic	4,023,477	1%
International Bee Research Association	A	Academic	203,373	68%
Royal College of Psychiatrists	A	Professional association	16,804,000	9%
Society of Dairy Technology	A	Academic	61,806	35%
Edinburgh Geological Society	B	Academic	46,793	8%
Institute of Physics	B	Academic	63,053,000	76%
Institution of Engineering and Technology	B	Academic	57,596,000	42%
International Glaciological Society	B	Academic	499,386	71%
Royal Institute of Navigation	B	Academic	390,648	14%
Royal Society of Chemistry	B	Academic	51,602,000	74%
British Cartographic Society	C	Academic	139,768	18%
British Sociological Association	C	Professional association	1,246,366	40%
Institute and Faculty of Actuaries	C	Professional association	29,692,000	0%
Royal Anthropological Institute	C	Academic	964,932	60%
Royal Economic Society	C	Academic	1,191,201	64%
University Association for Contemporary European Studies	C	Academic	353,184	61%
Association for Scottish Literary Studies	D	Academic	138,442	24%
European Association for Jewish Studies	D	Academic	69,071	0%
Newcomen Society	D	Academic	63,416	8%
Royal African Society	D	Academic	580,358	35%
Royal Musical Association	D	Academic	58,362	59%
Royal Photographic Society of Great Britain	D	Professional association	1,653,042	95%
Society for Libyan Studies	D	Academic	96,850	12%
Society for Medieval Archaeology	D	Academic	62,475	82%
Society of Antiquaries of London	D	Academic	1,617,634	3%

Annex J

References

- Archambault, E. et al. (2013) *Proportion of Open Access Peer-Reviewed Papers at the European and World Levels—2004-2011, Report for European Commission DG Research & Innovation*, http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf)
- Baldwin C. (2004) *What Do Societies Do with Their Publishing Surpluses?* (http://alpsp.org/Ebusiness/Libraries/LP_Stuff/C_Baldwin_NFPsurpluses.sflb.ashx)
- Björk, B-C 2012, 'The hybrid model for open access publication of scholarly articles – a failed experiment?' *Jasist*, vol 63, no. 8, pp. 1496-1504.
- Björk, B-C (2014), 'Open Access Subject Repositories - An Overview' *Jasist*, vol 65, no. 4, pp. 698-706
- Björk, B-C & Solomon, D (2012a), 'Pricing Principles used by Scholarly Open Access Publishers' *Learned Publishing*, vol 25, no. 2, pp. 132-137.
- Björk, B-C & Solomon, D (2012b), 'Open access versus subscription journals: a comparison of scientific impact' *BMC Medicine*, vol 10, no. 73
- Björk, B-C & Solomon, D (2014a), *Developing an Effective Market for Open Access Article Processing Charges*. Wellcome Trust, London.
- Björk, B.-C., & Solomon, D. (2014b). How research funders can finance APCs in full OA and hybrid journals. *Learned Publishing*, 27(2), 93–103.
- Björk, B-C, Laakso, M, Welling, P & Paetau, P (2014), 'Anatomy of Green Open Access' *Jasist*, vol 65, no. 2, pp. 237-250
- Björk, B-C & Solomon, D (2015), 'Article processing charges in OA journals –relationship between price and quality' *Scientometrics*, vol 103, no. 2, pp. 373-385.
- Bornmann, L. & Mutz, R (2015) Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references *Jasist*, (in press)
- Darley R, Reynolds D, Wickham C. *Open Access Journals in Humanities and Social Science*. London: The British Academy; 2014.(<http://www.britac.ac.uk/openaccess/index.cfm>)
- Davis P. (2010). 'Does Open Access Lead to Increased Readership and Citations? A Randomized Controlled Trial of Articles Published in APS Journals'. *The Physiologist* 53: 197-201
- Davis P (2011). Open access, readership, citations: a randomized controlled trial of scientific journal publishing *The FASEB Journal* 25: 2129-2134
- Dingwall R, Hewitt M, Turkmendag I. (2014) Evaluation of Learned Societies Project 2014. (<https://acss.org.uk/wp-content/uploads/2014/10/Evaluation-of-Learned-Societies-Project-2014.pdf>)
- Finch, Janet (2012) *Accessibility, sustainability, excellence: how to expand access to research publications: Report of the Working Group on Expanding Access to Published Research Findings* (<http://www.researchinfonet.org/wp-content/uploads/2012/06/Finch-Group-report-FINAL-VERSION.pdf>)
- Finch, Janet (2013) *Accessibility, Sustainability, Excellence: How to Expand Access to Research Publications: A Review of Progress in Implementing the Recommendations of the Finch Report* (<http://www.researchinfonet.org/wp-content/uploads/2013/02/Final-version.pdf>)

- Inger, S. and Gardner, T. (2013) *Scholarly Journals Publishing Practice. Academic journal publishers' policies and practices in online publishing. Fourth survey 2013.*
- Johnson, R., Pinfield, S., & Fosci, M. (n.d.). Business process costs of implementing “gold” and “green” open access in institutional and national contexts. *Jasist*, (In press).
- Kingsley, D (2014) ‘Paying for publication: issues and challenges for research support services’ *Australian Academic & Research Libraries*, vol45 no.4 pp262-281
- Laakso, M (2014) ‘Green open access policies of scholarly journal publishers: a study of what, when, and where self-archiving is allowed’ *Scientometrics*, vol 99 no2 pp 475-494
- Laakso, M & Björk, B-C (2012), 'Anatomy of open access publishing - a study of longitudinal development and internal structure' *BMC Medicine*, vol 10, no. 124.
- Laakso, M & Björk, B-C (2013), 'Delayed Open Access – an overlooked high-impact category of openly available scientific literature' *Jasist*, vol 64, no. 7, pp. 1323-1329.
- Lawson, S, ‘Total cost of ownership of scholarly communication: subscription and APC payments together’ *Learned Publishing* vol 28,no 1pp 9-13
- Lawson, S., & Meghreblian, B. (2014). Journal subscription expenditure of UK higher education institutions. *F1000Research*, 3.
- Mabe, M & Amin, M (2001) Growth dynamics of scholarly and scientific journals. *Scientometrics*, 51: 147-162
- Meadows, A. (2015) ‘Beyond Open: Expanding Access to Scholarly Content’, *Journal of Electronic Publishing*. Volume 18, Issue 3
- Morris S.(2005) ‘The true costs of scholarly journal publishing’. *Learned Publishing*.vol 18 no.2 pp115-126.
- Morrison, H.; Salhab, J.; Calvé-Genest, A.; Horava, T. (2015) Open Access Article Processing Charges: DOAJ Survey May 2014. *Publications* 2015, 3, 1-16.
- Pinfield, S., Salter, J., & Bath, P. A. (2015). The “total cost of publication” in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions. *Jasist*, (Early view)
- Redhead, C (2015) Growth of OA-only Journals Using a CC-BY License, OASPA Blogpost 19 May 2015 <http://oaspa.org/growth-of-oa-only-journals-using-a-cc-by-license/>
- Research Information Network (2014a) *Monitoring Progress in the Transition to Open Access: Report of a Working Group* (<http://www.researchinfonet.org/wp-content/uploads/2013/02/Report-final.pdf>)
- Research Information Network (2014b) *Nature Communications: Citation Analysis* (http://www.nature.com/press_releases/ncomms-report2014.pdf)
- Shepherd, P & Needham P ((2011) *Publisher and Institutional Repository usage Statistics: The PIRUS2 Project: Final Report* (http://www.cranfieldlibrary.cranfield.ac.uk/pirus2/tiki-download_wiki_attachment.php?attId=170&download=y)
- Solomon, D & Björk, B-C (2012a), 'Publication Fees in Open Access Publishing: Sources of Funding and Factors Influencing Choice of Journal' *Jasist*, vol 63, no. 1, pp. 98-107

Solomon, D & Björk, B-C (2012b), 'A study of Open Access Journals using article processing charges' *Jasist*, vol 63, no. 8, pp. 1485-1495.

Solomon, D, Laakso, M & Björk, B-C 2013, 'A longitudinal comparison of citation rates and growth among open access journals' *Journal of Informetrics*, vol 7, no. 3, pp. 642-650

Taylor and Francis (2014) *Open Access Survey 2014*
(<http://www.tandfonline.com/page/openaccess/opensurvey/2014>)

Ware, M and Mabe, M, (2015) *The STM Report: an overview of scientific and scholarly journal publishing*, fourth edition, (http://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf)

Willets, D. (2014). Progress review: Implementing Finch report recommendations. *Letter to Prof Dame Janet Finch*. (<http://www.researchinfonet.org/wp-content/uploads/2013/02/BIS-Transparency-Letter-to-Janet-Finch-One-Year-On-Response-January-2014.pdf>)

Woodward, H. M., & Henderson, H. L. (2014). *Report for Jisc Collections on total cost of ownership project: Data capture and process*. Kirkcudbright: Information Power Ltd. (https://www.jisc-collections.ac.uk/Global/News_files_and_docs/IPL-Jisc-Total-Cost-of-Ownership-Data-Capture-Report.pdf)

Annex K

Abbreviations and Glossary

AAM	Author accepted manuscript: the version of an article that has been accepted for publication, but before it has been formatted by the publisher
APC	Article processing (or publishing) charge: a fee levied by journals for publication of articles on OA terms.
ArXiv	a repository of pre-prints, particularly in the physical sciences
CCBY	The Creative Commons Attribution licence, which allows usage for all purposes, subject to content being attributed to the original authors. Other Creative Commons licences are available which restrict use to non-commercial purposes (CCNC), which forbid the creation of derivatives (CCND), or require users to share content that they create on a share alike basis (CCSA). These various terms may also be combined.
COAF	A partnership of major UK medical research charities to provide block grants to universities to meet the costs of APCs
COUNTER	The Counting Online Usage of Networked Electronic Resources initiative which sets standards to facilitate the recording and reporting of online usage statistics in a consistent credible and compatible fashion
Delayed OA	Journals that make their contents freely available on the publisher's site after an embargo period (though the terms on which content is made accessible often differ from those of fully-OA or hybrid journals).
DOAJ	The Directory of Open Access Journals: a listing of fully-OA journals
FWCI	Field Weighted Citation Impact: a measure of the relative quality of publications (see Annex C for a full definition).
Gold OA	Journals and articles that are freely accessible on the publisher's site
Green OA	A term used to describe the posting of versions of articles in repositories and other web-sites where they can be accessed freely, often after an embargo period
HEI	Higher Education Institution
HSS	Humanities and social sciences
Hybrid	Journals that offer to authors the option of making their articles freely-accessible, usually after payment of an APC
IR	Institutional repository
IRUS UK	Institutional Repository Usage Statistics UK: an aggregator of download data from UK IRs;
ISSN	International Standard Serial Number: a unique eight-digit number used to identify a periodical publication
JUSP	Journal Usage Statistics Portal: an aggregator of publishers' data on downloads from UK universities and colleges;

NIH	National Institutes of Health: the primary US agency responsible for biomedical and health-related research.
OA	Open access
OASPA	Open Access Scholarly Publishers Association
OpenDOAR	A directory of OA repositories
PIRUS	The Publisher and Institutional Repository Usage Statistics project which developed models for gathering and aggregating article-level download data from a range of sites.
PJA	Published journal article: a term used by some publishers instead of version of record (VoR)
Pre-print	A version of an article before it is submitted to a journal
PMC	PubMedCentral: a repository for scholarly articles in the health and life sciences
RCUK	Research Councils UK: the partnership between the seven Research Councils in the UK
REF	Research Excellence Framework: the procedures for assessing the quality and impact of the research undertaken in UK universities
SCOPUS	a bibliographic database containing abstracts and citations for journal articles, owned by Elsevier
Self-archiving	A term sometimes used to describe the process of posting or depositing versions of articles in a repository or other website, with a view to making them freely accessible
SHERPA RoMEO	a searchable database of publishers' copyright and self-archiving policies
STEM	Science, technology, engineering and medicine
TCP	'Total Cost of Publication': a term coined by the former UK Minister for Universities and Science, David Willetts, to describe the costs to universities of paying APCs alongside subscriptions to journals. See Section 5.
VAT	Valued added tax, charged on subscriptions to journal, and on APCs, at the rate of 20% in the UK
VoR	Version of record: the final published version of an article

Annex L

Units of Assessment in the Research Excellence Framework 2014

Main panel	Unit of assessment
A	1 Clinical Medicine
	2 Public Health, Health Services and Primary Care
	3 Allied Health Professions, Dentistry, Nursing and Pharmacy
	4 Psychology, Psychiatry and Neuroscience
	5 Biological Sciences
	6 Agriculture, Veterinary and Food Science
	7 Earth Systems and Environmental Sciences
	8 Chemistry
	9 Physics
	10 Mathematical Sciences
B	11 Computer Science and Informatics
	12 Aeronautical, Mechanical, Chemical and Manufacturing Engineering
	13 Electrical and Electronic Engineering, Metallurgy and Materials
	14 Civil and Construction Engineering
	15 General Engineering
	16 Architecture, Built Environment and Planning
	17 Geography, Environmental Studies and Archaeology
	18 Economics and Econometrics
	19 Business and Management Studies
	20 Law
C	21 Politics and International Studies
	22 Social Work and Social Policy
	23 Sociology
	24 Anthropology and Development Studies
	25 Education
	26 Sport and Exercise Sciences, Leisure and Tourism
	27 Area Studies
	28 Modern Languages and Linguistics
	29 English Language and Literature
	30 History
D	31 Classics
	32 Philosophy
	33 Theology and Religious Studies
	34 Art and Design: History, Practice and Theory
	35 Music, Drama, Dance and Performing Arts
	36 Communication, Cultural and Media Studies, Library and Information Management

Annex M

Desirable improvements to the quality and availability of data

In the course of our work, we have noted a number of issues relating to the quality and availability of data, and we note some of the more significant ones below. L

Accessibility and availability

Future exercises would be facilitated and rendered more accurate by measures including

- ❑ Full publisher-wide adoption of the NISO standard (<http://www.niso.org/publications/rp/RP-8-2008.pdf>) on the naming of different versions of articles, or the creation of a new simplified naming convention.
- ❑ Creation of a master lists of categories of online locations where articles may be posted, with examples, so that publishers can make their policies explicit in relation to those categories.
- ❑ Publishers to provide full public listing of their journals and their OA policies, including embargo periods, in accordance with the terminology and categories noted above.
- ❑ Full publisher-wide implementation of the NISO standard on Access and Licence Indicators (<http://www.niso.org/workrooms/ali/>)

Usage

- ❑ Article-level download data should be made openly-available. If that is not feasible on grounds of commercial sensitivities, then the PIRUS code of practice should be adopted and implemented by publishers, libraries, repositories and other sites from which different versions of articles can be downloaded, and the data made available to those who have a legitimate interest in it.

Financial sustainability for universities

- ❑ Data on expenditure in institutions for subscriptions, APCs and administration costs should continue to be collected and made public on an ongoing basis
- ❑ Reporting of APC expenditure data should be further standardised, preferably using the Jisc reporting template, including standardisation in the reporting of:
 - 'Publication date'
 - APCs distinguished from many additional charges (e.g. colour and page charges)
 - Any splitting of single payments between different funders
 - Consistent inclusion of VAT (if paid)
- ❑ Subscription expenditure should also be reported and made publically-available for as wide a range of publishers as possible
- ❑ Further work should be done to clarify administrative costs, particularly those associated with activities such as APC payments
- ❑ Approaches need to be agreed for estimating and, where possible, recording payment of non-centrally-managed payment of APCs
- ❑ Further work on publisher costs could also be carried out, although it is appreciated that this can often prove to be difficult because of commercial confidentiality.

Financial sustainability for learned societies

An improved evidence base in these areas would be particularly beneficial for the large number of small- and medium-sized societies that generate significant surpluses from a single journal (particularly, but by

no means exclusively, in the social sciences), and are thus constrained in their ability to experiment with new OA business models. Further work should be undertaken to examine whether and how societies could make the transition to OA, and the financial implications of doing so. Future studies should include

- ❑ more in-depth engagement with societies themselves, and access to their data on publication revenue and expenditure (with appropriate safeguards to preserve confidentiality).
- ❑ examination of the relationship between publishing and membership subscriptions.