

The Reckoning: An Analysis of Wellcome Trust Open Access Spend 2013-14

3 Mar, 2015

by [Wellcome Trust](#)

*To help make the costs around open access more transparent, the Wellcome Trust has [published details](#) on how much it spent on article processing charges in the year 2013-14. The data also shows to what extent the services the Trust has paid for – i.e. that the final, publisher-formatted versions of articles are freely available in Europe PubMed Central with a CC-BY licence – have been delivered. **Robert Kiley**, Head of Digital Services at the Wellcome Library, provides an analysis of what the latest data reveals...*

Once a year we ask all those institutions in receipt of an [open access \(OA\) grant](#) from the Trust to provide details on how the grant has been spent. In addition to simply calculating the total spend, and the average article processing charge (APC), we have, for the first time, done an analysis of the licence types that have been applied to these articles. The licence analysis was undertaken to determine to what extent articles had been published under the Creative Commons Attribution licence (CC-BY), in line with our OA policy.

Cost data

In headline terms, some 2556 articles were published under an APC model, and the total spend for the year 2013-14 was just under £4.7 million. The average APC was around £1837, whilst the median cost was £1800. Table one, shows how these figures compare with the previous year (2012-13).

The takeaway from this is that more Trust-funded articles are being published via the OA APC route – an increase of 20% – whilst APC fees are relatively static (an increase in the average APC of just 0.8%).

Focusing just on 2013-14 data, Table two provides a breakdown of the type of journal where the articles were published, and the average APC for each type. A fully OA journal is one in which every article is made OA (e.g. *PLOS One*, *Cell Reports* etc.), whereas a hybrid journal is one which is still published under a subscription model, but where individual articles can be made OA. The “unknowns” in table two are those that we were unable (programmatically) to determine if the article was published in a fully OA or a hybrid journal.

Although Wellcome Trust authors are publishing large numbers of papers in fully OA journals (around 24% of all papers in this cohort), there is still a very strong desire to publish in traditional, subscription-based journals, albeit via the OA, hybrid route (in line with the Wellcome Trust [policy](#)).

The second key finding is that the average APC levied by hybrid journals is 64% higher than the average APC charged by a fully OA title. This higher average fee is despite the fact that hybrid journals also enjoy a revenue stream from subscriptions.

Faced with similar data showing the high price of APCs charged by hybrid journals, a number of organisations – including the German Research Foundation, [DFG](#), and the [Norwegian Research Council](#) – do not allow OA grants to be used to fund APCs in hybrid journals.

Given the scale of hybrid OA publishing amongst Trust-funded researchers, such an approach would seriously limit the ability of many of our researchers to publish in their journal of choice, and therefore we have (to date) decided not to go down this path.

However, we do feel that the hybrid APC market is not working effectively and we are actively exploring other ways to tackle this marketplace's current dysfunctions. By way of example, the [FWF](#) (Austrian Science Fund) have recently imposed a cap of €1500 for hybrid articles. We will watch with interest whether this cap leads to any change in author behaviour (i.e. authors start to consider price when they are submitting an article for publication), or changes to publishers' APCs.

Our final table on costs provides an analysis of the top five publishers (by volume of Trust-attributed articles published).

What is significant here is the degree to which the two traditional, subscription-based publishers (Elsevier and Wiley) still dominate the field where Trust authors seek to publish. Together, these two publishers represent some 40% of our total APC spend, and are responsible for 35% of all Trust-funded papers published under the APC model.

It is also interesting to note that the average APCs for the fully OA journals published by Elsevier and Wiley are significantly higher than those charged by competitors at PLOS, BMC and OUP.

Compliance data

The other reason for collecting this data is to check whether the publisher has provided the service we demand in return for paying an APC.

When Trust funding is used to cover an APC, the publisher has to provide a number of specific services. These include depositing the final version of the article in PubMed Central (PMC) at the time of publication (this content is then mirrored to Europe PMC) and attaching a CC-BY licence to the article, thus making it clear that anyone can re-use that work for any purpose, subject to the norm of attribution.

To help us check compliance in a more automated way, we commissioned [Cottage Labs](#) to develop a “compliance monitoring” tool to programmatically determine whether the paper is in the Europe PMC repository and what licence (if any) is attached.

A summary of the key compliance findings is presented in table four:

The headline data is disappointing – particularly in terms of the relatively high numbers of papers which are not in the Europe PMC repository.

To investigate this further, we undertook a manual exercise to determine if these articles were freely available on the publisher site and/or were published “ahead of print”. The latter distinction is important as the PMC repository only accepts the final version of the article, and thus it not possible for a publisher to deposit a paper that is released online “ahead of print”.

The results of this exercise are shown in table five:

Table five shows that typically publishers *are* making APC-funded papers OA on their sites. This suggests therefore that, for some publishers the problem is confined to being able to consistently deposit papers at PMC. By way of example, of the 325 papers that were not in Europe PMC, 110 of these (34%) were published by Wiley.

It is also interesting to note that 23% of the papers are classified as “early view” papers, and that in the fullness of time we can expect most (all) of these papers will be deposited in PMC (and mirrored to Europe PMC). Looking across this cohort of papers we can see a wide range of early view publication dates, with the oldest article (not yet formally published) originally published online in October 2013.

However, even if we assume that *all* the early view papers will be deposited, this still means that 237 papers, available in final published form, for which APCs totalling over £480,000 have been paid, are not available in Europe PMC.

This is unacceptable.

Licence compliance was always going to be more problematic – authors do not always fully understand the Trust’s requirements, a situation compounded by some publishers who persist in offering a choice of licences to Trust funded authors – but we do now have a baseline compliance rate (61%) to build on.

Action Areas

As a matter of urgency the Wellcome Trust has started to follow up with all the publishers concerned, to better understand the problems they are experiencing, and how they plan to rectify them. It is encouraging how responsive publishers have been so far.

Their responses have also highlighted the number and variety of obstacles facing articles destined for open access; obstacles which publishers, funders and institutions must identify and navigate when tracking compliance. Problems identified include:

- Content not available in PMC, even though the publisher believes they have successfully deposited it;
- Institutions reporting an APC payment – but publisher has no record of that article ever being flagged as an OA paper (and no invoice was ever raised);

- Articles submitted for publication *before* the CC-BY requirement was introduced (1st April 2013);
- Ambiguous (or no) licence information in the article;
- Contradictory licence information (the human readable text says one licence, the licence encoded in the XML cites a different one);
- Author has selected an alternative to the CC-BY licence (either erroneously, or otherwise), presumably unaware of the [sanctions](#) the Trust has defined for non-compliance).

Pinning down which of the above issues relates to which specific articles is a moving target, but the numbers involved indicate clear trends and challenges.

It is clear that early view articles pose another challenge; it is only possible to identify whether the article is early view manually, and if it is, the months which elapse before its final publication (and deposit in PMC/Europe PMC) are still troubling. The Trust is discussing this issue with staff at PMC to see if a policy change can be made, such that publishers can (for papers funded by named funders) deposit an early view paper and follow-up with the final, published version, in due course.

We will also be following up with institutions with regards to our CC-BY requirement and we will make it clear that, going forward, any APC costs associated with papers that are **not** licenced CC-BY cannot be met from Trust funds.

We will also be ensuring publishers are aware of their obligations to ensure papers associated with our funding are deposited and licensed in line with our requirements where an APC is being charged. We will also make clear our expectation that the licences are encoded in a consistent, machine-readable way.

Moving forward, in cases where particular journals or publishers consistently fail to deliver the services that have been paid for, we will declare them to be non-compliant with our OA policy, and alert researchers and institutions to this, including through the [Sherpa FACT](#) service. More generally, we will encourage all researchers and institutions to actively consider whether the APC they are being charged by a particular journal is justified in light of the quality of service they receive.

Conclusion

Last year, when we released our first set of APC data, I published a [blog post](#) that concluded by saying that we expect *every* publisher who levies on open access fee to provide a first class service to our researchers and their institutions.

One year on, there are still significant problems – especially in terms of depositing content in PMC and licensing this in accordance with our requirements – and the first class service called for still seems to be some way off. These problems are particularly prevalent amongst publishers offering a hybrid OA option, which, as discussed above, is also the more expensive way to comply with our OA policy.

The Trust remains committed to its OA policy, but ultimately if we are to be successful and get to the point where all Trust funded research is OA, then the friction in the system – which is all too evident in this analysis – has to be resolved.

You can read the [Wellcome Trust's OA policy](#) on our website and if you're interested in

exploring our 2013-14 APC data set you can find it on [Figshare](#). The underlying code for the Cottage Lab compliance tool we used has been [This piece of software has been made available on github](#) under an [Apache open source licence](#), to enable others to take this code and adapt it to suit their specific needs.