# Charity Open Access Fund APC spend (2016-17): an explanation of the data

## Overview

Since 2014 the Wellcome Trust has published details of its open access spend as reported by UK institutions in receipt of an OA block grant in the previous year. Since March 2016 we have reported on the APC spend of the Charity Open Access Fund ([COAF](http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Open-access/Charity-open-access-fund/)), a collaboration between research charities to provide a single funding mechanism to cover APC’s.

This year we are publishing the data for 2016-2017 Charity Open Access Fund (COAF) open access spend. During this period the members of COAF were Wellcome Trust, Arthritis Research UK, British Heart Foundation, Bloodwise, Cancer Research UK and Parkinson’s UK. As in previous years in addition to simply reporting spend, we have run the data through our “compliance monitoring” tool (developed for us by [Cottage Labs](http://cottagelabs.com/)) to help us determine programmatically whether the paper is in the Europe PMC repository and what licence (if any) is attached to the article.

The dataset includes information when an article processing charge (APC) was levied or when an APC was waived as part of an offsetting agreement between publisher and institution. If an author has self-archived a paper, this information is not included in this dataset.

Equally, data are not included in cases where a researcher (based at an institution not in receipt of a COAF block grant) received a supplement to their grant to cover OA publishing costs.

## The dataset

The data used for this analysis was self-reported by institutions between November 2017 and February 2018. The analysis was carried out using the Wellcome’s CottageLabs [Compliance Tool](https://compliance.cottagelabs.com/) on 8 and 13 March 2017. The analysis was conducted using the raw data provided by 35 of the 36 institutions in receipt of a COAF grant. While every effort has been made to provide accurate information, there may be errors within the analysed data.

The dataset contains the following elements:

### PMCID/PMID/DOI

To help COAF identify whether a work is in PMC/Europe PMC, we ask institutions to cite the PMCID (or the PubMed ID if a PMC ID is not known). We also ask for the article DOI. Where possible, any missing IDs have been inserted by the automated compliance checking tool.

### Publisher

This lists the name of the publisher, as reported by the compliance monitoring tool.

### Journal name

This lists the name of the journal (which published the COAF-funded work), as reported by the compliance monitoring tool.

**ISSN**

The automated compliance monitoring tool supplemented the data provided by the institutions with the journal ISSN where it was able to identify it in Europe PMC.

**Fully Open Access Journal?**

If the automated compliance monitoring tool can find the journal in [DOAJ](http://doaj.org/) it will be marked as TRUE, if it does not, it will be marked as FALSE.

### Article title

This is the title of the article, as reported by the institution.

### Cost

The cost listed is the figure which the institution has reported as the full cost of the APC. These costs should exclude any page or colour charges which the publisher may levy.  
  
The cost quoted, includes VAT where applicable. All costs have been converted into £ sterling.  
  
As you look through the data you will see (for example) different prices cited for articles published in the same journal. These differences can be due to fluctuations in currency prices, the use of institution/author membership discounts, the inclusion of additional costs (e.g. page and colour charges) within the APC, or that an error has occurred.

**Discounts, memberships & pre-payment agreements**

If the article APC reported received a publisher discount or benefitted from membership schemes or pre-payment agreements, it is stated here. This data is presented as reported by the institution. Where an institution did not return any data in this column articles have been marked as “Not Reported”. Where an institution reported data in this column for some articles but left others blank, we have left this blank with the assumption that no form of discount was applied to that particular article.

**Full text in Europe PMC?**

The automated compliance checking tool checks whether the full text of the article appears in Europe PMC.

If it does, it will be marked as TRUE, if not, it will be marked as FALSE.

**Licence in Europe PMC**

The automated compliance monitoring tool attempts to detect the licence for the article in Europe PMC. Initially this is attempted by looking at the article XML, if no licence information is detected there then the tool looks at the HTML version of the article.

If the licence can be identified as a Creative Commons licence it is marked as such. If another licence can be found it will be marked as “non-standard-licence”. If no licence information can be found it will be marked as “unknown”.

**Licence on Publisher website**

The automated compliance monitoring tool attempts to detect the licence for the article on the publishers website. This will be "not applicable" if the licence was found elsewhere, or "unknown" if the licence could not be detected on the publisher's site.

**Accepted author manuscript?**

The automated compliance monitoring tool checks whether the version of the paper in Europe PMC is the Accepted Author Manuscript or the final published version. If an APC has been paid the final version of the paper should be in Europe PMC, otherwise it is not compliant.

If the article is the author manuscript, it will be marked as TRUE. If it is the final version it will be marked as FALSE. If it was not possible to identify its status it will be marked as “unknown”.

**Ahead of print?**

Some journals release articles “ahead of print” before assigning them to an issue and depositing them in Europe PMC.

If the full text is in Europe PMC this value will be "not applicable".

If the full text of an article is not in Europe PMC this field will tell you whether the article has been published ahead of print. If an article has only been published ahead of print this will be marked TRUE. If an article has been assigned to an issue this be marked FALSE. If the tool was unable to determine the status of the article this will be marked "unknown".

**COAF compliant?**

The automated compliance checking tool calculates whether an article is compliant based on whether the full text is in Europe PMC **and** the detected licence is CC-BY or CC0. As APCs were paid for all articles in this dataset compliance has been manually set to false for all accepted author manuscripts.

**Funder of Research/Grant ID**

This information is presented as reported by the institutions.

## Comment

We hope that this data will be of use to help better understand the cost of OA publishing.

If through use of this data you identify errors or believe the status of an article to be incorrect please notify me. Working with colleagues at institutions and publishers we will endeavour to investigate these issues and correct errors where identified. Corrections to the dataset will be published as a new version of the dataset along with a note explaining the changes.

An article summarising the findings from the data can be found at: [www.wellcome.ac.uk/](http://www.wellcome.ac.uk/)

Hannah Hope

h.hope@wellcome.ac.uk

Open Research Coordinator, Wellcome

30 April 2018