

# Monitoring Organisational Article Processing Charges (APCs) using Bibliographic Information Sources: Turku University Library Case

**Anna-Kaarina Linna**

University of Turku, Finland and Federation of Finnish Learned Societies, Turku, Helsinki, Finland, [anna-kaarina.linna@tsv.fi](mailto:anna-kaarina.linna@tsv.fi), [orcid.org/0000-0002-7901-2789](https://orcid.org/0000-0002-7901-2789)

**Irene Ylönen**

Open Science Centre, University of Jyväskylä, Finland, [irene.ylonen@jyu.fi](mailto:irene.ylonen@jyu.fi), [orcid.org/0000-0002-1374-2375](https://orcid.org/0000-0002-1374-2375)

**Anna Salmi**

Open Science Centre, University of Jyväskylä, Finland, [anna.m.salmi@jyu.fi](mailto:anna.m.salmi@jyu.fi), [orcid.org/0000-0002-7599-9870](https://orcid.org/0000-0002-7599-9870)

## Abstract

As open access publishing has become more widespread and required by research funders and the research community, the management and monitoring of article processing charges (APCs) have emerged as an important task in research organisations around the world. Within this tendency, a question of the comprehensiveness of organisational APC monitoring has become relevant. This case study demonstrates how the comprehensiveness of in-house APC monitoring can be evaluated using international bibliographic information sources like Web of Science and Scopus, where it is possible to identify the

corresponding author, as well as Unpaywall and DOAJ, which contain information about the open access statuses and APCs of articles. Based on study results, it can be assumed that the organisation's in-house bookkeeping has succeeded in registering 52 percent of APC invoices while 48 percent have not been identified. The results show that the number of unreported publications that have been openly published and whose corresponding author is affiliated with the university is almost equal to those registered in the university's institutional APC report. The study describes the stages of data collection and processing in order of implementation, which allows a similar review to be feasible in another organisation. At the end of the article, development proposals are presented for both the organisations' in-house data collection and the content of publishers' invoices.

**Keywords:** open access publishing; open access fees; article processing charges; open access policy

## **1. Introduction**

Over the last few years, development in open access publishing has increasingly led to Gold Open Access based on article-specific publication fees. In many countries, the development has been accelerated by the transition to nationwide transformative open access agreements. Many universities have also created centralised funds to pay article processing charges (APCs). Because of the above, the monitoring of article payments has become pivotal. To be able to outline the total cost of scientific publication, it should be clear what is being paid and to whom. A better and wider knowledge base on the cost of publishing has been identified as an important goal, especially in scientific libraries (see e.g. Lawson, 2015; Pinfield et al., 2016). However, estimating the total cost has proven difficult.

In Finland, 12 universities (86 percent of all universities) followed APCs in 2019 (Vilén et al., 2021). In recent years, the motivation of Finnish universities to monitor the costs of open access publications has been partly influenced by a national policy. One main goal of the policy is to make the total cost of scholarly publishing transparent and publicly available (Open Science Coordination in Finland, Federation of Finnish Learned Societies, 2020).

As the first university library in Finland, Turku University Library started the systematic collection of APCs in co-operation with the university's financial services in the spring of 2017. Monitoring was organised on two levels: the processing of purchase invoices centrally in the financial services and a more detailed analysis of payments in the library. It is noteworthy that the model is not based on a centralised fund, but invoices are paid from unit-specific budgets. Payments made with credit cards were retroactively included in the data collection.

At the University of Turku (UTU), invoice processing is centralised in the financial services. Financial secretaries identify the APC payments from amongst all invoices received by the university and allocate them to the correct general ledger account. Before the organisation-level monitoring started, the library arranged a training session for financial secretaries on open access publishing and APCs to help them identify the correct invoices.

The library receives a compilation of the APC invoices based on the data of the APC general ledger account as an Excel file three times per year. The original invoices are available as PDF files in the local accounting system where an information specialist has a user account. However, library access is limited to the specific general ledger account only. With authorised access to the APC account, it is possible to view the original invoices more closely and identify the publication, Digital Object Identifier (DOI), journal title, publisher and whether the publication channel is full open access or hybrid. After identification, the information specialist creates an Excel table where the identifying data of the publication and publication channel are recorded. If the fee is not APC but instead, for example, a submission fee, it will be deleted from the report.

In addition, the library checks whether and under what conditions the publisher allows the publication to be self-archived, as well as whether the information about the publication has been deposited in the university's current research information system UTUCRIS. If not, the publication and APC information are recorded to UTUCRIS, and the full article is self-archived and made openly available via the institutional repository. Finally, the compilation of the APC data is submitted to the international Open APC service (University of Turku APC data available at <https://treemaps.openapc.net/apcdata/turku-u/>).

Open APC is a service managed by the Bielefeld University Library. Open APC releases data sets on fees paid for open access journal articles by universities and research institutions (Pieper & Broschinski, 2018). In the future, the separate data transfer from Turku University Library to Open APC will become unnecessary because APC data can be transferred from UTUCRIS to the national VIRTA publication service<sup>1</sup> automatically. After this, VIRTA delivers the data received from the Finnish organisations centrally to Open APC.

### **1.1. Research Questions**

Recently, the library began planning a study to examine the comprehensiveness and quality of the current APC data collection. One premise was to harvest APC information from international bibliographic information sources and compare it with in-house APC bookkeeping to find out how many paid articles were actually discerned by the organisation's two-stage monitoring model. The years 2017 and 2018 were selected for the analysis for two reasons: the collection of APCs had just started and the transformative agreements, where APCs are often included in the subscription price, were not yet concluded.

In this study, we aim to answer the following questions:

- Is it possible to identify articles that may have been published with an article processing charge (APC) by using international bibliographic sources?
- How does the information retrieved from external sources compare with the organisation's own bookkeeping?
- Do articles not identified by in-house reports share common characteristics (e.g. a publisher)?

## **2. Literature Review**

There have been a few attempts to study the monitoring of article fees by research organisations. The latest review can be found from Andre Bruns and Niels Taubert (2021). They have reviewed Gold Open Access payment

data compiled from German universities into the Open APC data collection. In their study, Bruns and Taubert (2021) found two significant problems in tracking publication fees. They call them blind spots. Comprehensive data are not available for organisations, as they do not have a centralised fund to facilitate the monitoring of payments. In these organisations, fees are paid in different units, and collecting centralised data has not been simple. There may also be a complex set of funding sources. In addition, even if an organisation has a centralised fund, monitoring payments can still be challenging, as payments are made by different entities of a research organisation and processed in a variety of ways.

To look at the blind spots they perceived, Bruns and Taubert (2021) used, in their own words, an indirect approach to identify APC-specific publications. They used bibliometric data sources (Web of Science and SSN-Gold-OA-list) for the identification instead of relying on data collected by organisations. As a result, for a total of more than 40 German universities, the authors concluded that the data reported in the Open APC database was notably incomplete. In universities that transferred data to Open APC, more than half of the estimated costs were not covered. However, the reasons for the incompleteness of the data were beyond the scope of their study.

A local knowledge gap in monitoring publication fees was also noticed by Stephen Cramond et al. (2019) a couple of years earlier. In 2018, the Council of Australian University Librarians (CAUL) carried out a project to study the financial impact of APC payments on universities in Australia and New Zealand. One of the main results of the pilot project was how little is actually known about the nature and amounts of publication fees in universities. An important finding was the variance of different sources of funding such as research funders, endowment funds, departmental funds, transfers from research partners or researchers' personal funds. They also found that only a few universities had a centralised APC fund, and only some of the universities collected data on publication fees at all.

Cramond et al. (2019) reported that, due to the limitations of centrally reported figures, the Australian project decided to look at established APC payments in an indirect way. They evaluated the volumes of open access articles and potential APC payments on the basis of publication data in the Web of Science and Scopus and compared the data obtained with data from Unpaywall. The methodology used was the same as in this article.

APC payments that are paid from different funds and from different sources outside the centralised organisational fund are often referred to as 'APCs in the wild' (see e.g. Andrew, 2016; Monaghan et al., 2020). Theo Andrew (2016) noted in a review of the University of Edinburgh's 'APCs in the wild' that there was a significant amount of APC fees, of which the university was unaware. He estimated that 'APCs in the wild' could account for up to 20 percent more in the total cost of publication than previously thought.

In Sweden, the monitoring of APC payments in research organisations has also been examined (Gustafsson et al., 2017; Lovén, 2019). Stockholm University Library and Gothenburg University Library have been tracking open access publishing costs since 2016. Sweden has a national centralised service run by the National Library of Sweden to collect APC fees from universities, called Open APC Sweden<sup>2</sup> (Gustafsson et al., 2017).

In her article, Lisa Lovén (2019) has described the process of collecting fees at Stockholm University. The process within the university works similar as at the University of Turku. However, Stockholm University seems to have a more transparent accounting system, in which all payments are accessible for all authorised users. This makes it much easier to track invoices from the system if they were not initially recorded for the correct APC account. In her review, Lovén (2019) has looked more closely at the publishers to whom the largest amount of APC fees is paid, as well as the sources of funding.

As of 2018, Stockholm University Library managed a significant centralised fund for APC payments (Lovén, 2019). At that time, Lovén (2019) noted that Sweden decided to withdraw from the agreement with Elsevier, and the saved funds were used to finance article fees for full open access journals. Many benefits were identified by the centralised fund, including easier tracking of article fees, greater understanding of the open access publishing landscape and clear benefits for researchers who no longer had to worry about paying article fees themselves.

In a report from the University of Gothenburg, Gustafsson et al. (2017) noted how difficult it was to find out the exact amount of APC payments. In Gothenburg, publication fees for 2016 were analysed. Due to the monitoring problems, two different methods were used in this analysis to identify the payments. First, the authors reviewed the organisation's invoices in the

financial management system, and second, they requested information directly from the publisher. Excluding Elsevier, data were obtained from several significant publishers. Combining these methods, the aim was to estimate the minimum APC costs for the University of Gothenburg in 2016. The authors concluded that at least SEK 5.2 million (approximately EUR 490,000) was paid in 2016 for the openness of articles.

According to the Gustafsson et al. (2017) report, APC payments had been credited with a total of 18 different accounts in the university's financial management system. APCs paid with a credit card were also a problem in terms of tracking. Additionally, there were major shortcomings in the publishers' invoices from the perspective of identifying the article.

When it comes to tracking APC fees at the organisational level utilising financial management systems, all the articles and reports above (Bruns & Taubert, 2021; Cramond et al., 2019; Gustafsson et al., 2017; Lovén, 2019) recognise the great difficulty of collecting data. To improve the quality of the data collected by organisations, further consideration needs to be given to the reasons why the coverage of the data is largely poor. To get to the root causes, it is necessary to continue to observe the data collected by organisations and compare it with data from external sources.

### **3. Methodology**

The objective of this study was to identify APC-liable publications from bibliographic information sources and to compare the results with centrally collected APC data from the University of Turku. The initial assumption was that the university's own bookkeeping has managed to collect APCs comprehensively. The analysis is based on comparison of in-house APC reports and harvested University of Turku-affiliated open access articles from the bibliographic information sources Web of Science and Scopus. Unpaywall and DOAJ are used for identifying the open access statuses of articles more closely and to exclude articles from so-called Diamond publication channels, which do not charge APCs. The analysis focuses on open access articles whose corresponding authors are affiliated with the University of Turku or Turku University Hospital and which are not reported in internal APC bookkeeping at the university or hospital.

Publications were included in the analysis based on the institutional affiliation of the corresponding author (Scopus) or the reprint author (Web of Science) to the University of Turku or the Turku University Hospital. Although there is disciplinary variation in determining the roles and responsibilities among authors, including the role of the corresponding or reprint author, they are generally likely to be the principal authors and therefore most likely to be responsible for paying APCs (Cramond et al., 2019; González-Alcaide et al., 2017). The analysis in this study is based on this assumption. As the reprint or corresponding author needs to remain locatable for requests as far to the future as possible, the role is typically assigned to an author with a stable position in the organisation (Weiss, 2012).

However, even though the corresponding author is usually the person responsible for assuring the publication costs, the APC invoice may be distributed to a co-author in another institution or split by several authors (Lovén, 2019). For this study, it was reasonable to assume that the corresponding or reprint author was most likely to be a person responsible for submitting the APC fee, even though there could be no absolute certainty about it. This is one of the limitations of the study.

### **3.1. Data Collection from in-House Bookkeeping**

For this study, a separate Excel spreadsheet of all identified APCs during 2017 and 2018 was received from bookkeeping. In total, the data contained information on 233 individual articles whose payment year was either 2017 or 2018. When the APC data were compared to publications reported by the university to the Ministry of Education and Culture,<sup>3</sup> it was found that the year of payment is not necessarily the same as the publication year. Two of the articles had been published in 2016, 17 in 2019 and two in 2020, which were excluded from the analysis to ensure a consistent comparison based on publication years. Since the Unpaywall identification is based on DOIs, 12 articles without DOIs need to be removed from the analysis.

To supplement the university's in-house data, similar information on APCs paid by the Turku University Hospital was also received. The hospital's publications are recorded in the university's current research information system (UTUCRIS) and reported to the Ministry of Education and Culture as outputs of the University of Turku. The information on APCs paid by the hospital



*Table 1: Total number of publications in University of Turku and Turku University Hospital APC bookkeeping by payment year.*

	University of Turku	Turku University Hospital	Total
Payment year 2017	78	26	104
Payment year 2018	122	39	161
Total number of publications	200	65	265

was therefore natural to include in the analysis. The hospital data were interpreted by a librarian trainee preparing his bachelor's thesis on APC collection.<sup>4</sup> From these data, a total of 65 articles with paid APCs were identified. A total of 265 publications from the organisation's own bookkeeping ended up in the analysis (Table 1).

### 3.2. Data Collection from External Sources

The following external sources were used to collect and analyse the article metadata:

- Web of Science
- Scopus
- Unpaywall
- DOAJ

The search from Web of Science and Scopus was conducted in April 2020. The filtered results of the search were submitted to Unpaywall in the same month. The bibliographic knowledge base of the analysis was created. At the final stage, in the autumn of 2021, the DOAJ metadata were integrated into the data set to exclude Diamond open access journals from the analysis.

#### 3.2.1. Web of Science and Scopus

The first phase of the project consisted of gleaning the metadata of the journal articles that were published in 2017 and 2018 and were affiliated with the University of Turku or Turku University Hospital from 2017 to 2018. The decision to use and combine data from both Web of Science Core Collection and Scopus was based on the observation, made over time by the publication

information reporting team at the Turku University Library, that a consistent number of university-affiliated publications were indexed in Scopus as online first (early access) versions before becoming available in Web of Science. Moreover, there is variation between the two databases in terms of disciplinary coverage, e.g. in social sciences and humanities that are prevalent at the University of Turku (for a comparison of document overlap between Web of Science and Scopus, see Visser et al., 2021).

The Web of Science results were used as the baseline, and the result lists of the two databases were matched based on DOIs. Duplicates that were found in both Web of Science and Scopus were excluded from the Scopus data. This choice was made due to the rich open-access-related metadata available in the former. In addition, author affiliation information was found to be particularly clear in the Web of Science metadata since it included the name and address of the reprint author plus a separate metadata field for the corresponding author.

Article types selected for the study were original refereed articles and review articles. These two articles types were selected because they were assumed to be particularly APC liable. Both Web of Science and Scopus index the name and affiliation of the corresponding author or authors as well as information on the open availability of the articles. The information on corresponding authors was found to be crucial, as it is not available in any other internal or external sources.

Next, the article entries were filtered based on whether or not they had a clearly indicated University of Turku or Turku University Hospital affiliation in the corresponding author field. Publications where the corresponding author had two or more affiliations were excluded from the analysis because it was not possible to unambiguously identify which organisation had been paid the possible APC payment. In addition, if the corresponding author had not provided the University of Turku or Turku University Hospital email address, it was assumed that the author's primary affiliation was neither the university nor the hospital.

### ***3.2.2. Unpaywall***

The second phase of the project consisted of Unpaywall analysis for those articles where corresponding authors were affiliated with the university or the hospital.

Unpaywall is a non-profit service run by OurResearch that curates a list of almost 20 million freely available scholarly articles, and it is commonly used by browser plug-ins (Else, 2018). In this study, however, the Unpaywall Simple Query tool was used for receiving a structured open access report. The report was created in two batches because only 1,000 DOIs were possible to submit at once. To ensure accurate matching with open access status data, articles without a DOI were excluded in this phase. Finally, the received Excel sheets were combined and copied to the analysis.

### 3.2.3. DOAJ

Although DOAJ information is aggregated in Unpaywall, information was also needed on whether or not the publication channel charges APCs from authors. Unpaywall does not include information on APCs, but DOAJ does. In order to exclude the Diamond open access journals, i.e. journals that do not collect APCs from authors or readers, the original Web of Science and Scopus data enriched with the Unpaywall results were compared with DOAJ information. DOAJ journal-level metadata were downloaded in .csv file format and imported into the analysis Excel. The comparison was made by ISSN IDs.

## 4. Analysis and Results

### 4.1. Data Analysis

The analysis focuses on open access publications whose corresponding author is affiliated with the University of Turku or Turku University Hospital and which are not reported in the university's or hospital's in-house APC bookkeeping.

To perform the analysis, all collected data were compiled into one Excel file and compared using Excel's VLOOKUP syntax based on the articles' permanent identifiers (DOIs, ISSNs). After comparison, the analysis data sheet contained 1,897 rows of articles and a 39-column information matrix compiling the in-house bookkeeping, Web of Science, Scopus, Unpaywall, DOAJ and university publication data reports from 2017 and 2018. For the final analysis, the following indicators were selected for further examination.

- |                              |  |
|------------------------------|--|
| 1) In-house bookkeeping      | DOI, APC payment year  |
| 2) University data warehouse | Publication year   |
| 3) Web of Science            | Reprint author, email, open access information   |
| 4) Scopus                    | Correspondence address, affiliation, access type   |
| 5) Unpaywall                 | best_oa_evidence, best_oa_host, oa_status, best_oa_license, journal_name, journal_issn_l, journal_issns, publisher |
| 6) DOAJ                      | APC payment  |

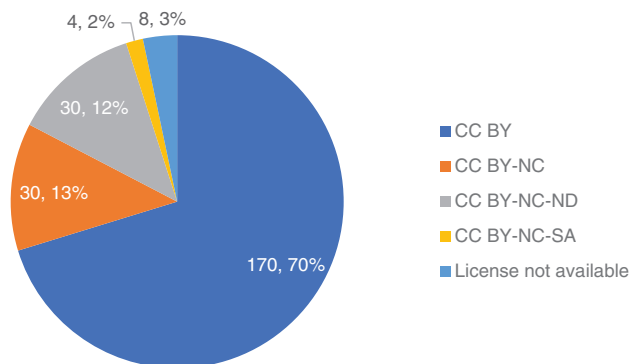
In manual revision, a total of 199 publications were excluded because the corresponding author had not provided the university's or hospital's email addresses, or they had a double affiliation. After this procedure, 1,698 rows remained in the data. Next, the publications recognised in either the university's or hospital's APC bookkeeping were removed from the analysis. After that, the data set contained 1,433 articles. Finally, the openly available articles were extracted from the data using Unpaywall open access statuses.

Unpaywall uses four different open access statuses: Green, Gold, Hybrid and Bronze, which are commonly used in open access contexts (Priem, 2021). It is also widely agreed that publications identified as Bronze open access cannot be defined as genuinely open access because there is no information about their long-term availability and license (Piwowar et al., 2018). In total, the analysis data contained 298 publications that were indexed as Gold or Hybrid status in Unpaywall. A closer examination of these publications showed that there can be identified both Creative Commons licensed publications and publications that were published in DOAJ-indexed journals without a license. Finally, based on the information in DOAJ, journals that do not charge an APC were removed from the analysis.

## **4.2. Results**

The purpose of this study was to survey how comprehensively the organisation's own accounting reaches the publications where the APC has been paid. Based on the data analysis from different bibliographic and open access sources, we identified 242 articles (Figure 1) that have been published openly

Fig. 1: The share of identified licensed open access articles that were not reported in organisation APC reports and that were published in journals that charge an APC.



available in Gold Open Access or Hybrid journals and are mostly licensed under Creative Commons licenses.

One of the goals of this analysis was to determine whether unidentified article processing fees would have a connecting factor, such as a specific publisher.

After analysing both internal and external data, no significant differences between publishers were found. As expected, a significant number of articles had been published in Springer, Wiley and Elsevier journals (see Figure 2). Interestingly, articles from these publishers came out on top in both analyses. Some of their articles were well founded in the university's bookkeeping (Figure 3), but also plenty of these publishers' articles have been unnoticed in the university's monitoring (Figure 2). The reason for this is difficult to assess. In-house follow-up at the University of Turku is based on observations and interpretations of the content of invoices by individual financial secretaries.

Most clearly, the in-house follow-up was unable to report the articles of MDPI and American Chemical Society (see Figure 4). This would indicate that the publisher in question may have room for improvement in the information recorded in the invoices.

Notably, this final comparison (Figure 5) shows that there can be identified almost as many reported as unreported open access articles.

Fig. 2: Distribution of unreported open access articles (Gold and Hybrid) by publishers based on the number of their articles. Publishers with four or less articles are combined with 'Other publishers'.

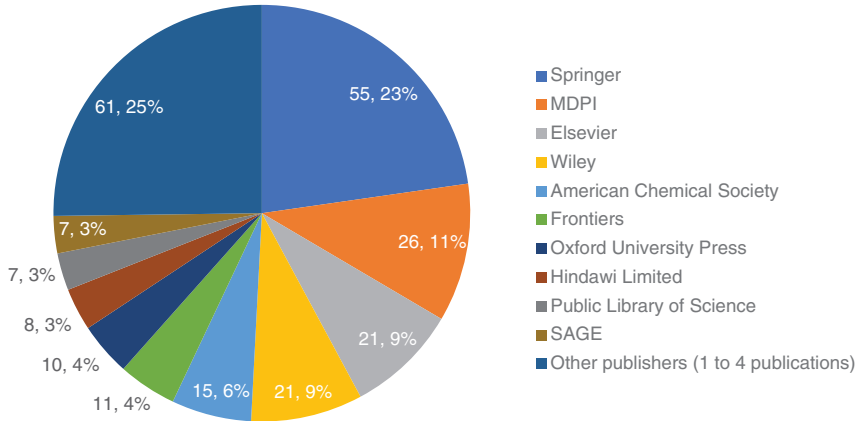
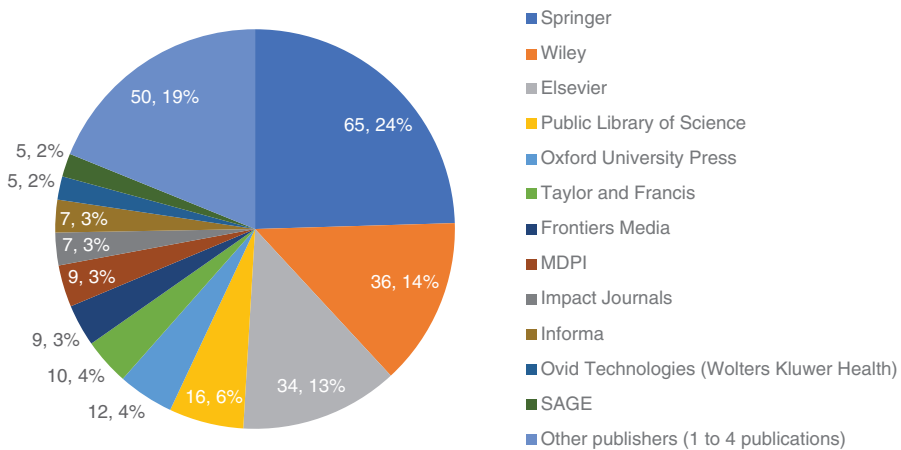


Fig. 3: Distribution of in-house APC reports by publishers based on the number of their articles. Publishers with four or less articles are combined with 'Other publishers'.



Based on the analysis, it can be assumed that the organisation's bookkeeping has succeeded in registering 52 percent of APC invoices, while 48 percent (242 articles) have not been identified (Figure 5). When focusing on the

Fig. 4: Comparison of publishers of reported and unreported publications based on article volumes. Publishers with four or less articles are combined with 'Other publishers'.

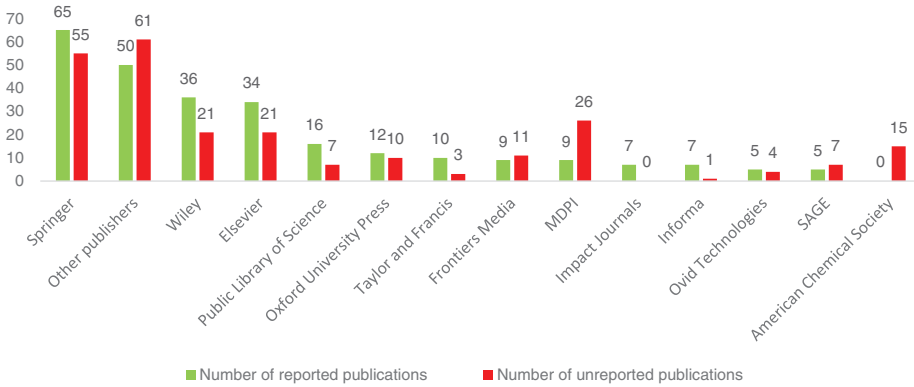
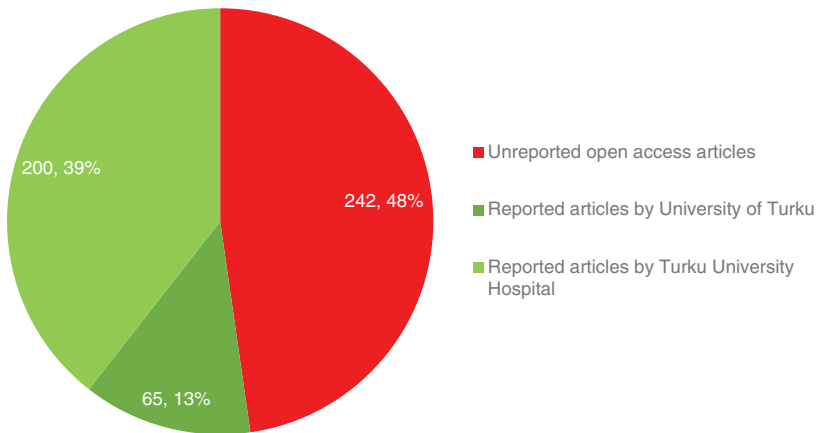
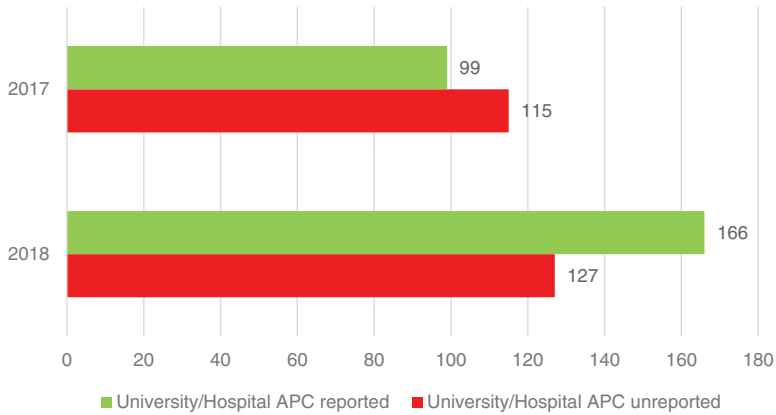


Fig. 5: The share of articles reported in university or hospital APC bookkeeping and identified open access articles from bibliographic information sources.



annual level, a total of 99 articles have been found in 2017 and 166 in 2018 (Figure 6). Considering that only two years of statistics are under review, the proportion of identified payments has clearly increased.

Fig. 6: The comparison of reported and unreported APCs by publication year.



## 5. Discussion and Conclusions

The first research question of this study was to find out whether it is possible to obtain information on paid APCs by using external information sources. Based on the analysis, and paying attention to certain limitations, it seems possible to harvest and identify APCs from external sources. External data sources have also been utilised previously, for example at Stockholm University where the Web of Science is used to complete the annual institutional APC report by identifying open access articles in which University of Stockholm researchers are identified as corresponding authors (Lovén, 2019).

However, a few limitations must be remembered when using bibliographic information sources. Payments cannot be clearly defined from them. A particular problem is caused by the roles of the corresponding author and publication fee payer. In this study, it has been assumed that the corresponding author is also the payer of the article fee, but in fact, this is not always the case. The paying party cannot be found clearly from any available source other than the organisation's own monitoring. In addition, when using external sources, it is impossible to find out the payment year of the publication fee. The publication fee may have been paid in a different year than the publication has been published.



The email address used by the author can also cause problems. If the researcher is not using the organisation's email domain when submitting the article to the journal, it can be challenging to identify the affiliation to a certain organisation. Consequently, articles of this type may be completely excluded from the review, as decided in this study.

Using external sources to monitor publication fees is possible but rather laborious and contains the above-mentioned limitations. The most ideal situation would be if the payments could be identified within the paying organisation. Although the internal monitoring of organisations does not seem very accurate in light of this and previous research, it is possible to develop it.

Although in this study we did not focus on the reasons why the organisation's own monitoring is so inaccurate, some possible causes can be presented. Currently in Finland, the organisations' own monitoring is largely based on the manual work of the financial secretaries. They are usually not familiar with open access terminology. If the APC is not clearly mentioned in the invoice, it is possible that it is not registered. Publication fees paid by credit card are particularly difficult because there is very little information on the credit company's invoice.

The second research question concerned how information retrieved from other information sources compares with the organisation's internal accounting. Based on information received from Web of Science, Scopus, Unpaywall and DOAJ, it can be assumed that the University of Turku has paid APCs for more than 500 articles published in 2017–2018 (see Figure 5). Most of these articles were indexed in Web of Science (444 articles, 86 percent), and only 34 (7 percent) articles were published in journals that were indexed only in Scopus. In this context, it is important to note that Web of Science was assessed as the primary data source, so if the same article was found in both databases, only Web of Science data were taken for analysis. In the university's own bookkeeping were also 29 articles that were not indexed by either Web of Science or Scopus.

The third and final research question concerned whether open access APC-liable publications identified from external bibliographic sources share some common characteristics (e.g. a publisher). A significant number of articles (51 percent) reported in the organisation's own bookkeeping had been published

in Springer, Wiley and Elsevier journals (see Figure 3). When comparing the same distribution with publishers of non-reported open access articles, one can observe that, although Springer is the largest of these, MDPI rises second (Figure 2). Both statistics also highlight the large deviation of articles among different publishers, which suggests that in 2017–2018, open access publishing was not centralised for journals of large commercial scientific publishers. In fact, the data from DOAJ demonstrate that there were 32 articles that were published in open access journals that did not charge APCs, so they were removed from the final analysis for that reason.

The final result shows (Figure 5) that the University of Turku in-house book-keeping has succeeded in registering 52 percent of APC invoices, while 48 percent have not been identified. The results show that the number of unreported publications that have been openly published and whose corresponding author is affiliated with the university is almost equal to those registered in the university's institutional APC report. Based on the result, it can be assumed that the APC monitoring for the years 2017–2018 has not produced a particularly reliable result, and based on that, the APC statistics collected by the organisation should not be used exclusively as a basis for evaluating the total costs.

### **5.1. Recommendations for Organisations**

To develop local monitoring, it is important to be able to identify payments as precisely as possible. Therefore, the practice mentioned by Lisa Lovén (2019) in her article at Stockholm University would be applicable in other organisations as well. If invoice management were as open as possible and access to the system was also possible for library and experts in open science within the institution, it would improve the identification of publication fees. Lovén mentions that, in Stockholm's invoice management system, it is possible to search by the name of the publisher, which enables finding invoices that were inadvertently credited to something other than the publication fee account. The more centralised the monitoring of payment information is and the more the expertise of information specialists can be utilized, the more accurate the result could be.

Organisations themselves can facilitate the tracking and identification of invoices by instructing researchers to provide invoicing information in the

correct format in publishers' systems. Additionally, by using different types of references that can be added to invoices, they can be more easily identified in the organisation's financial management.

Naturally, centralised funds facilitate the monitoring of publication fees. However, it must be remembered that they also need to be communicated clearly and sufficiently widely within the university, so that all individual researchers and financial secretaries would remember their existence when the payment of the publication fee becomes relevant to them. Especially in situations where the organisation is unable to cover all publication fees, but there are internal criteria for centralised payment of fees, this is particularly important.

## 5.2. Recommendations for Publishers

Publishers would have an opportunity to clarify and facilitate the monitoring of publication fees. From the paying organisation's point of view, the most important issue would be the adequate information on the invoices received from the publishers.

The ESAC community (2017) has listed recommendations of the information an APC invoice should contain in machine-readable format:

- *Name and email address of the author who is affiliated to the paying institution (must be the corresponding author)*
- *Complete statement of the author's affiliation to the paying institution (e.g. university, institute, department)*
- *Funding organisation (research funder and grant ID)*
- *Date of acceptance*
- *Date of publication*
- *Journal title*
- *ISSN*
- *Article title*
- *Article type*
- *DOI and link to the published article*
- *Amount due*
- *Discounts and discount group (if applicable)*

- CC license (ESAC community, 2017, <https://esac-initiative.org/about/oa-workflows/>)

Currently, not all publishers' invoices have sufficient information. In practice, the invoice may have contained so little information that it has not been possible to identify the paid article at all. The invoice with unclear or incomplete information can remain unnoticed by the financial services, and even if it is recognised, further processing in the library is laborious.

### 5.3. Study Limitations

The study has concentrated on using and analysing information on international bibliographic sources. One can justifiably ask why the study did not make use of the lists of paid open access articles available from publishers. During the research, pre-compiled publisher-specific lists were centrally available only from large hybrid publishers, so getting them from smaller publishers would have been comparatively difficult. Furthermore, analysis based solely on data from Web of Science and Scopus by default excludes publications issued in national languages in Finnish publication channels. On the other hand, based on an extensive survey conducted in 2020, APCs were very rare in Finnish peer-reviewed publication channels; only 2 percent of 332 of them charged APCs at that time (Linna et al., 2020). The choice to use Web of Science and Scopus as sources was based on their comprehensiveness and ready availability at the time of the study.

### 5.4. Availability of Data

The data can be accessed through Jyväskylä University Digital Repository at <https://doi.org/10.17011/jyx/dataset/83986>.

## References

Andrew, T. (2016). *Improving estimates of the total cost of publication by recognising 'APCs paid in the wild'*. The Winnower. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=9f4798017e3acfd62057a1bc0d3d03acaba47606>

- Bruns, A., & Taubert, N. (2021). Investigating the blind spot of a monitoring system for article processing charges. *Publications*, 9(3), 41. <https://doi.org/10.3390/publications9030041>
- Cramond, S., Barnes, C., Lafferty, S., Barbour, V., Booth, D., Brown, K., Costello, D., Croker, K., O'Connor, R., Rolf, H., Ruthven, T., & Scholfield, S. (2019). Fair, affordable and open access to knowledge: The Caul collection and reporting of APC information project. *Proceedings of the IATUL Conferences*. Paper 2. <https://docs.lib.purdue.edu/iatul/2019/fair/2>
- Else, H. (2018). How Unpaywall is transforming open science. *Nature*, 560, 290. <https://doi.org/10.1038/d41586-018-05968-3>
- ESAC community. (2017). ESAC workflow recommendations for transformative and open access agreements. *ESAC Initiative*. <https://esac-initiative.org/about/oa-workflows/>
- González-Alcaide, G., Park, J., Huamaní, C., & Ramos, J. M. (2017). Dominance and leadership in research activities: Collaboration between countries of differing human development is reflected through authorship order and designation as corresponding authors in scientific publications. *PLoS One* 12(8), Article e0182513. <https://doi.org/10.1371/journal.pone.0182513>
- Gustafsson, K., Henning, K., Jönsson, P., Nilsson, R., Petersson, I., & Rydholm, A. (2017). *Kartläggning av OA-publicering och APC vid Göteborgs universitet*. Göteborgs universitetsbibliotek. <https://gupea.ub.gu.se/handle/2077/54026>
- Lawson, S. (2015). 'Total cost of ownership' of scholarly communication: Managing subscription and APC payments together. *Learned Publishing*, 28(1), 9–13. <https://doi.org/10.1087/20150103>
- Linna, A.-K., Holopainen, M., Ikonen, A., & Ylönen, I. (2020). Kotimaiset tieteelliset julkaisusarjat ja avoimuus. *Informaatiotutkimus*, 39(4), 4–32. <https://doi.org/10.23978/inf.98656>
- Lovén, L. (2019). Monitoring open access publishing costs at Stockholm University. *Insights* 32(1), 3. <http://doi.org/10.1629/uksg.451>
- Monaghan, J., Lucraft, M., Allin, K., van der Graaf, M., & Clarke, T. (2020). 'APCs in the wild': Exploring funding streams for an accelerated transition to open access SpringerNature. <https://doi.org/10.6084/m9.figshare.11988123.v4>
- Open Science Coordination in Finland, Federation of Finnish Learned Societies. (2020). *Open access to scholarly publications. National Policy and executive plan by the research community in Finland for 2020–2025 (1)* (Responsible Research Series 3:2019, 2<sup>nd</sup> volume). <https://doi.org/10.23847/isbn.9789525995343>
- Pieper, D., & Broschinski, C. (2018). Openapc: A contribution to a transparent and reproducible monitoring of fee-based open access publishing across institutions and nations. *Insights*, 31, 1–39. <http://doi.org/10.1629/uksg.439>

Pinfield, S., Salter, J., & Bath, P. A. (2016). The 'total cost of publication' in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions. *Journal of the Association for Information Science and Technology*, 67(7), 1751–1766. <https://doi.org/10.1002/asi.23446>

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of open access articles. *PeerJ* 6, Article e4375. <https://doi.org/10.7717/peerj.4375>

Priem, J. (2021, Decembre 1). What do the types of oa\_status (green, gold, hybrid, and bronze) mean? *Unpaywall support portal*. <https://support.unpaywall.org/support/solutions/articles/4400177288-what-do-the-types-of-oa-status-green-gold-hybrid-and-bronze-mean->

Vilén, T., Danielsson, M., Ikonen, A., Jussila, J., Laakkonen, A., Linna, A.-K., Merimaa, M., Suonpää, S., Ylönen, I., & Xu, Q. (2021). *Kirjoittajamaksut ja niiden seuranta: Havaintoja ja kehitysehdotuksia. FinELibin Open APC -projektin loppuraportti* (Kansalliskirjaston raportteja ja selvityksiä 2/2021). Kansalliskirjasto Nationalbiblioteket. <http://urn.fi/URN:ISBN:978-951-51-7163-4>

Visser, M., van Eck, N. J., & Waltman, L. (2021). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quantitative Science Studies*, 2(1), 20–41. [https://doi.org/10.1162/qss\\_a\\_00112](https://doi.org/10.1162/qss_a_00112)

Weiss, P. S. (2012). Who are corresponding authors? *ACS Nano*, 6(4), 2861. <https://doi.org/10.1021/nn301566x>

## Notes

---

<sup>1</sup> VIRTA Publication Information Service, implemented in 2016, is a data hub that compiles bibliographic information of scientific publications from Finnish higher education institutions and research organisations using different local solutions for the publication data collection, such as Current Research Information Systems (CRIS). The Ministry of Education and Culture's annual publication data collection is carried out through VIRTA. Further information is available at <https://wiki.eduuni.fi/display/cscvirtajtp/VIRTA+in+English>.

<sup>2</sup> <https://www.kb.se/samverkan-och-utveckling/oppen-tillgang-och-bibsamkonsortiet/open-apc-sweden.html>

<sup>3</sup> The Ministry of Education and Culture collects publication data from Finnish higher education institutions, state research institutions and university hospitals on a yearly basis to produce a knowledge base on the research activities and the social impact of the Finnish research system. For data collection, the publications are classified based on their format (publication type classification), scientific quality (Publication Forum) and contents (field of scientific classification).

<sup>4</sup> Vuori, Petri (2021): Kirjoittajamaksutietojen kerääminen ulkoisista lähteistä. Bachelor's thesis of the Turku University of Applied Sciences, available at <https://urn.fi/URN:NBN:fi:amk-2021061816385>. Although Anna-Kaarina Linna acted as one of the supervisors of the thesis and the library produced Web of Science, Scopus and Unpaywall data for the work, the analysis in this article has been carried out separately.