The 5th SPARC Japan Seminar 2011

"Burgeoning Open Access MegaJournals"

PLoS ONE and the Rise of the Open Access MegaJournal

Peter Binfield
(PUBLIC LIBRARY of SCIENCE)

Abstract

PLoS ONE is the archetypal "OpenAccess MegaJournal" – in 2011 the journal published almost 14,000 articles, constituting approximately 1 in 60 of all STM articles published in that year. Despite this success, MegaJournals are a relatively recent phenomena. PLoS ONE itself is five years old, but it is only in the last 18 months that similar titles have been launched by the likes of Nature, Springer, and SAGE. Clearly this is a publishing model which can be extremely successful, and this is the reason that other publishers are experimenting with their own launches. Dr. Peter Binfield (the Publisher of PLoS ONE since 2008) will present an overview of what constitutes a "MegaJournal", he will review the performance to date of PLoS ONE; and he will survey the more recent entrants. He will conclude with some predictions about where he believes this model will take the publishing world in the next 5 years.

Peter Binfield

After graduating with a PhD in Optical Physics from Aberdeen University, Peter Binfield began his publishing career at Institute of Physics Publishing in Bristol, UK as a Commissioning Editor in their books program. From IoPP, he moved to Kluwer Academic Publishers (KAP) in the Netherlands, where he ran their Major Reference Work program (encyclopedias and handbooks) and then held a variety of positions at KAP, including the management of the Physical Sciences group (Physics, Materials Science, Chemistry) and the directorship of the Plant Sciences and Earth & Environmental Sciences division. KAP merged into Springer Publishers, and during that period he held a position in Business Development, working on projects such as e-books, e-reference works and the Springer Open Choice program. In 2005 he moved to the US to live in California and work for SAGE Publications running their successful US Journals Division (some 220 journals across medical and social sciences). Since April 2008, Peter has been running PLoS ONE, which is now the largest peer reviewed journal in the world, publishing almost 14,000 articles in 2011 alone.

The Public Library of Science

The Public Library of Science is a publisher. We have been an organization since October 2000 and a publisher since October 2003. We published seven open access. We also have a blog site, where we have a number of bloggers and a couple of other products called PLoS Currents and PLoS Hubs. We are the largest not-for-profit open access publisher. We are one of the big three (PLoS, Hindawi and BioMed Central). We are the only major open access publisher based in the US. Geographically, we are based in San Francisco that is where I am from and also Cambridge in the UK. We have about 120 people in total, about 100 in the US. We have been self-sustaining since 2010. We are a not-for-profit, but we produce a surplus now, which is largely due to PLoS ONE.

Of seven journals we published, PLoS Biology and PLoS Medicine are highly selective journals run by professional staff editors. The four journals

in the middle are what we call our community journals and they are very similar to an average society journal that you might encounter in the world (Figure 1). We are here to talk about PLoS ONE which launched in December 2006. It has just turned 5 years old. We would say that PLoS ONE has really launched the era of the open access mega journals.

PLoS ONE's editional process

The most interesting thing about PLoS ONE is out editorial process. All papers are fully and formally peer reviewed. When we peer-review papers, we are only looking to determine whether the content is scientifically sound. That means we ask very objective questions; is the work rigorous, is it ethical, was it well reported, do the conclusions follow the data and so on.

There, we are only looking to decide whether this is a scientific article and does it deserve to join the literature. Specifically what we are not asking is how impactful is that article, what kind of degree of advance is it, how important is the work because we believe that those kinds of questions are very subjective and that they are most appropriately answered after publication not before publication. What that means is everything we publish has been vetted to decide whether it should be published, but as a reader you cannot then determine how good or bad that article might be on some scale above the minimum. We aim to provide online tools that actually allow the reader to make that determination after publication and not try to make it before.

A 'First Choice' journal

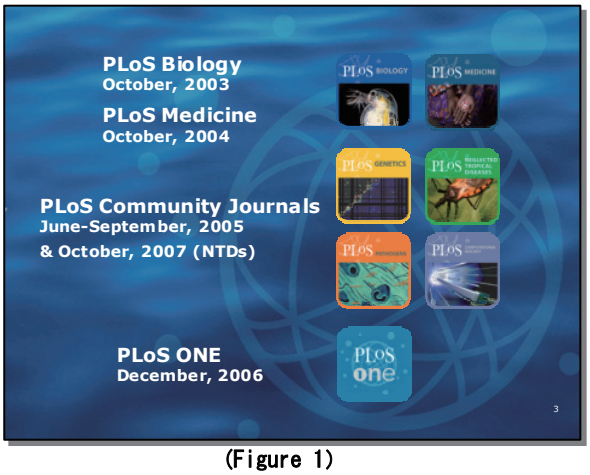
We heard earlier today about the cascade model of peer-review. I wanted to emphasize that PLoS ONE does not necessarily operate a cascade model. We do take transfers from our other journals, but that is less than 5% of our submissions. In fact, what we find in our author survey is that we are the first choice journal for over 40% of all of our submissions. We are the first or second choice for three quarters of all of our authors. What that means is that people either submit directly to us or they submit to their top journal first, get rejected and then come straight to PLoS ONE. Typically, we find that their top choice journal is Nature, Science or Cell, which are very high-level journals.

In addition, we find that our authors have had a very good experience publishing with us. 40% claim it is the best publishing experience they have ever had, and as many as 90% would rank it the best or almost the best (Figure 2).

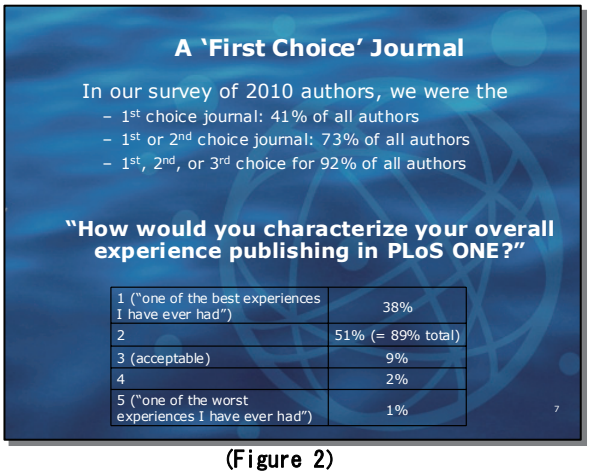
PLoS ONE published articles

The interesting reason that we are here today is because of the success of PLoS ONE. This graph shows the number of published articles every 6 months since launch (Figure 3). In 2007, which was our first year of existence, we published 1200 articles, and that is a larger number than probably 99.7 % of all journals in the world. In 2010, we became the largest journal in the world. Just 3 years after launch, we published almost 7,000 articles that year. In 2011, we published almost 14,000 articles. PubMed index is about 900,000 articles a year. If you imagine that the entire universe of scientific content is roughly 1 million articles and the rest might be social sciences and humanities, we basically published about 1.5% of all articles in the world last year.

It is just worth noting the break in the graph, where the curve accelerates, is the day we got an impact factor. PLoS disagrees with the impact factor. We actually campaign against it. But for the outside world, it is a very important number, which meant that a lot of people could submit to us



(Figure 1)



(Figure 2)

that were not allowed to before.

Features of open access mega journals

How do we define an open access mega journal? It has to be open access. In order to be very large, it has to cover a very broad subject area or preferably all subject areas. PLoS ONE covers all subject areas.

The content needs to be peer-reviewed for rigor, not for impact. Because if you peer-review for impact, you are going to be rejecting some large proportion of content and therefore not grow in a way you should. Because of that, you need to provide some sort tools or metrics that allow the author to decide how impactful or how good or bad the articles were.

Ideally, of course, the model needs to be scalable, which means each article has to have a revenue source that covers the cost of that article and typically that is an APC fee. The organization of the journal also needs to be set up to be scalable. These journals grow very fast so you need to scale very fast.

The inherent advantages of a mega journal

There are many advantages to being a mega journal (Figure 4). For example, compared to a program with perhaps 200 journals, when you are one large journal, you only need to be indexed once in MedLine, Web of Science and so on. As an author, you only need to be evaluated once. There is no being rejected by a journal or moving down and moving down being rejected. Because you are very large, you get a lot of visibility, a lot of usage. Because your scope covers everything in a subject area, there is no need for an author to publish anywhere else. As the publisher, you can consolidate many things that would normally have to be done many times, one marketing plan, one blog, one Twitter stream.

You realize economies of scale that make you very efficient. In this model, there is no economic reason to limit the size of the journal. This is not a scarcity model. You can grow as big as you need. We would say that filtering before publication is an outdated way of doing things that belongs in a scarcity model, not in an open access distribution model.

The end result is a healthy environment for the authors. They no longer have to beg to be published in certain journals. They no longer have to waste time attempting to get into the top journal. Finally, because of your size and visibility, you can set standards. You could set peer-review standards or presentation standards in a field.

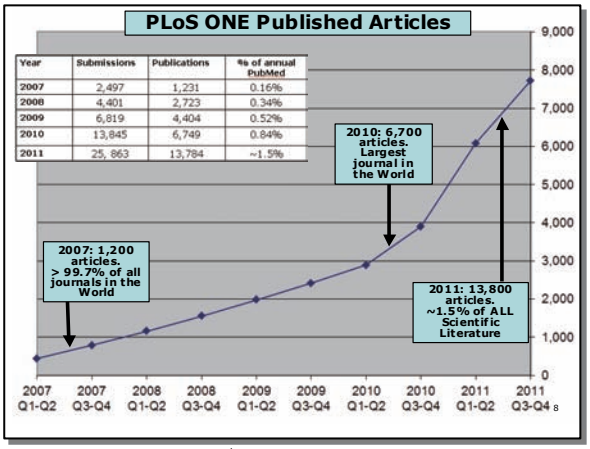
Recent launches of PLoS ONE 'clones'

PLoS ONE has been very successful and has now been emulated or copied by many other publishers (Figure 5). This is a list of the recent launches of PLoS ONE clones but it is not completed (Figure 6). They are ordered by the APC charge. Remember that the PLoS ONE charge is $1,350, so there is a clustering. I wanted to highlight four of these journals that I think are particularly interesting.

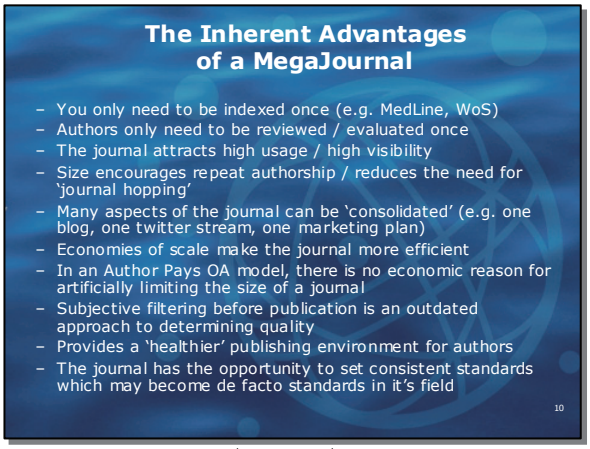
The first is Scientific Reports by Nature. What I think is interesting there is the fact that it was launched by Nature. Even though they did not put their name on the journal, it is clear that it is a Nature journal and the Nature brand carries a lot of weight.

Second is Springer Plus. What is interesting about Springer is the size of their organization. They have perhaps almost 2,000 journals which they could funnel rejected papers from that corpus into Springer Plus. It is also interesting that their price is $300 cheaper than PLoS ONE.

You may not have heard of the Scientific World



(Figure 3)



(Figure 4)

Journal, which is owned by Hindawi, an Egyptian publisher. This journal was a subscription journal, but Hindawi bought it in about September last year and converted it to an open access mega journal. What is interesting here is that it was the journal that they bought had an impact factor. As we know from PLoS ONE, the impact factor is extremely important for the success of a journal like this. None of the other journals on this list have an impact factor because they have all been launched from scratch and it will take perhaps 3 years to be assigned one.

Then, SAGE Open is interesting because this is a journal in the social sciences. You do not see many open access journals in the social sciences and their APC is very low, $700.

When PLoS ONE launched, Harold Varmus, the Nobel Prize Winner that founded PLoS ONE described it with this quote. "PLoS ONE will be a very large compendium of papers that have been vetted for scientific quality, but which will not be confined in terms of their likely importance." The same statement, obviously, applies to the collection of PLoS ONE clones.

How could we measure 'importance'

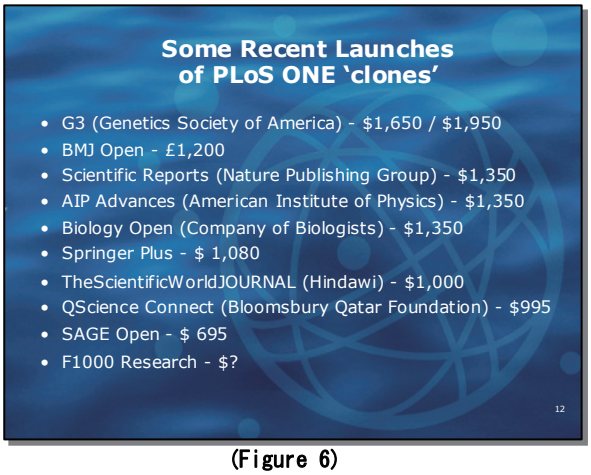
Then we ask ourselves if that is the case, how could we measure importance of these articles? We could measure a number of things such as citations to the articles, the usage of the articles, social metric such as bookmarking, community ratings, star ratings and so on, media coverage, blog coverage.

This is what we have done at PloS (Figure 7). We have attempted to collect metrics under these headings. We have created something called our article-level metrics program to present these data on each article. On every article, we have a metrics tab (Figure 8). If you click the metrics tab, you get the article-level metrics. First of all, we display a table of the usage of this article. Each bar is a month. If you hover over the month, it gives you a breakout of the detail for that month. We provide usage from the PLoS platform and also from PubMed Central.

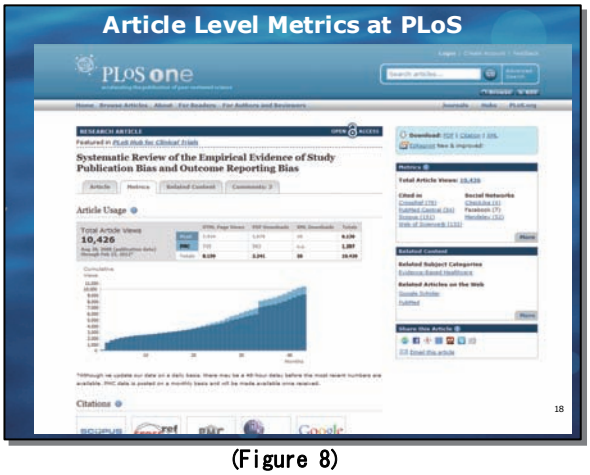
In addition, we show the number of citations to

(Figure 5)

(Figure 7)



(Figure 6)



(Figure 8)

this article as measured by four different citation counting services. For example, Scopus has found 19 citations of this article, and if you click, you go to the Scopus homepage where it shows that data (Figure 9). In addition, we attempt to measure the activity on social networks. We measure CiteULike, Delicious, Facebook likes, and Mendeley users. We will be adding Tweets in the next few weeks. For example, here are the 96 readers of this article on Mendeley. Then, we have a count of the number of blog articles that have written about this article and trackbacks to other locations on the Internet. We allow users to give article star ratings and have discussions around each article. However, star ratings and comments are not very well used, to be honest.

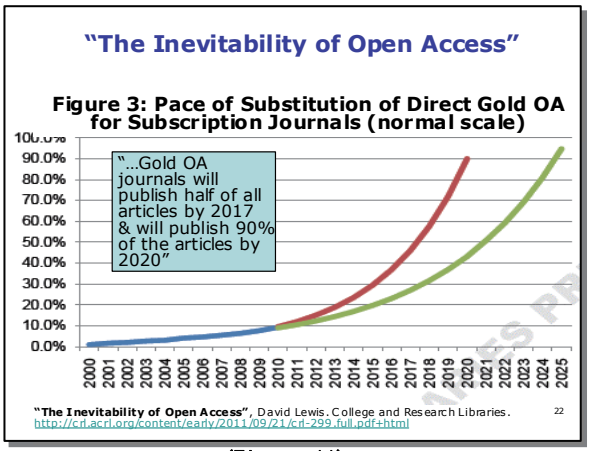
Having compiled that kind of article of a metric data, we can then use it as a discovery tool in our search forms. For example, you can run a search and sort the results by most views, least views, most cited, and so on (Figure 10). If you hover over, you get number of citations per service and so forth.

Open questions for being a mega journal

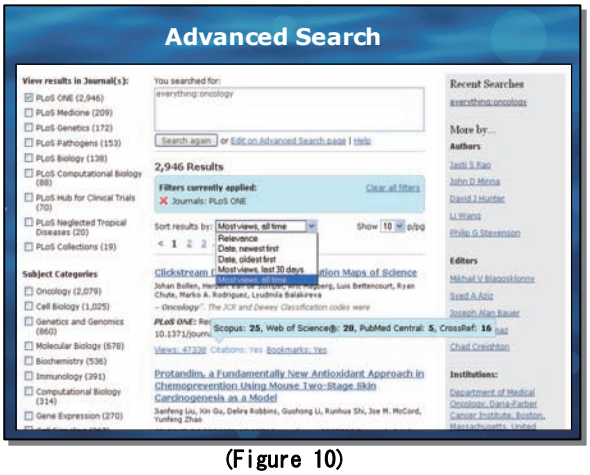
What are the open questions for being a mega journal? Can we actually develop tools that measure impact? This year PLoS ONE may publish 3% of the literature. If you are that big, are you really a journal anymore? When you are publishing much more than your entire organization, how do you interact with that organization? For example, PLoS believes in PLoS ONE and they believe PLoS ONE is the future. If you are a competitor publisher that simply launched it because you see an opportunity in the market, will you believe in your product in the same way PLoS does? If PLoS ONE continues to succeed and the clones of PLoS ONE succeed in the same way, very soon we could have just a very few, very large journals. We could have less than 100 journals, each publishing about 1% of the literature. If that happens, what does that mean for the ecosystem of 25,000 journals that exist right now?

This is a model that attempts to predict how fast the open access model will displace the subscription

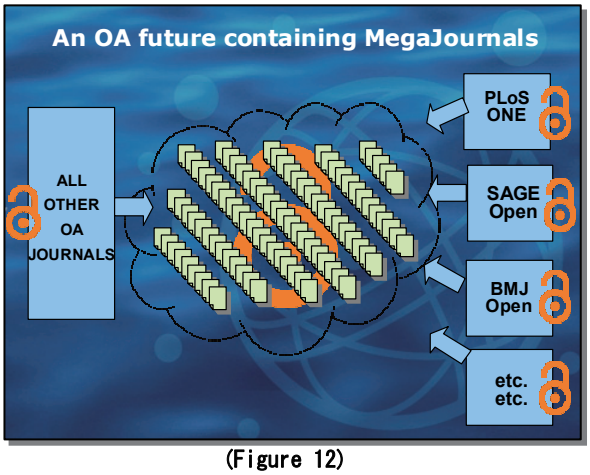
(Figure 9)



(Figure 11)



(Figure 10)



(Figure 12)

model (Figure 11). It was put together before open access mega journals really hit the scene. Personally, I think the most likely curve is the red one. I think within 5 years, 50% of content will be open access in large part because of these mega journals.

What does their future hold? There is a world in the future where there is small number of very large open access mega journals, putting open access content into the world that has not been differentiated in quality (Figure 12). On the left are all other open access journals and there are thousands of them, of course. Those journals typically have tried to stratify their content based on impact. We end up with a cloud of open access content, some of which has been stamped with an impact measure and some of which has not. I think the opportunity for the future is for people to develop tools and services that look at that cloud and attempt to differentiate it to provide citation metrics around it, discovery tools, filters, and so on. We already see evidence of this. We have services like Mendeley or Faculty of 1000.

Summary

I think we have shown at PLoS ONE that impact and technical assessment can be separated in a successful journal. It is possible to create post-publication mechanisms, which allow you to enhance the content and show how important each content is. I think what has not been proven yet is how effective those metrics are in the real world. With the launch of the clones and the success of PLoS ONE, I think we can say that the open access mega journal is not going to be going away in a hurry.

If PLoS ONE continues to be successful, and if the clones are successful, and if more of them are launched, I think we could see a dramatic change in the publication landscape. There are currently 25,000 journals in the world and perhaps you only need 100 which is clearly a scary prospect for many people, but hopefully research communication and research itself will be accelerated and improved.

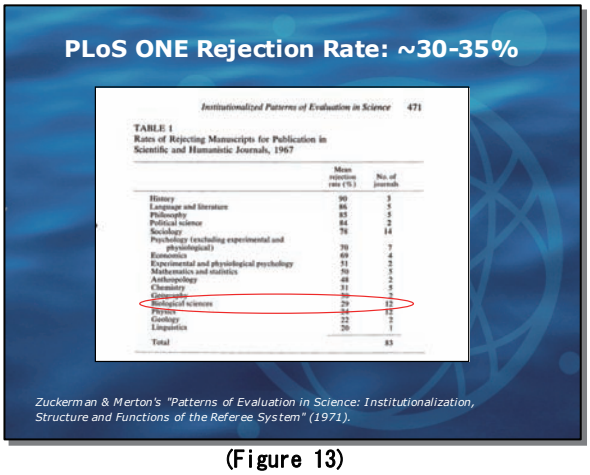
(Q1) It was interesting to look at the figures of published papers to submissions. It is interesting to see that when your impact factor did come out, the acceptance rate actually did fall of PLoS ONE. I am interested in Peter's comment. Is that a reflection of science in general that about half of the articles are scientifically sound? I am interested to know how you see that trend emerging. It is for the previous 2 years your acceptance rate had been about 65% or about two-thirds of what you had received and after the impact factor came down that dropped to about half.

The other comment is related to that is I am interested in your views on the future of peer-review. If we draw the process to perhaps the logical conclusion of the metrics that are offered post-publication and we are not making comments on impact, is there a place where post-review will replace peer-review and the community will judge, as it does already in some circumstances, the importance of a paper after it has published?

(Binfield) To take the first question, the numbers actually are not as they appear. In fact, before we had an impact factor, our acceptance rate was roughly 71% or 72% and now it is roughly 4% lower than that about 67% or 68%. The reason that it looks different is partly because the journal is growing so fast month on month, there is more submissions every month and they take 3 or 4 months to work through the process and the math is just you cannot divide one by the other one this chart.

But even with that said, our rejection rate did change by about 5% and that was because different contents started being submitted to us. For instance, we are opened up to the whole of China the moment we got an impact factor and we got a large proportion of Chinese submissions and their acceptance rate was different to our average.

If we only reject papers that are unpublishable and we reject about 30%, is that appropriate? There was a study done in 1971 (Figure 13). I think of this is an ice core. It looks back in a period before the world was tainted by impact factors and "publish or perish" and a lot of commercial publication activity. This was a small study. There were less journals back then but this study



(Figure 13)

found that in a survey of 12 biosciences journals, the average rejection rate was about 30%, which is exactly what we see due to our natural behavior. We do not set that rejection rate. It just happens.

You also asked whether post-publication peer-review might work in the future. PLoS ONE made a bit of a marketing misstep in its early days. I think it promoted some of these metrics as some sort of post-publication peer-review, which clearly they are not. They are certainly post-publication evaluation mechanisms. But I think taking it to its logical conclusion, you could imagine something even lighter as it were than the decision that we make, and simply should it be published or not, and then a post-publication process that really says and how could it be improved. We have some experiments going there, but I think that is in the future.

(Q2) As a not-for-profit organization and an organization that truly believes in open access to science, I am just wondering if you consider these other open access clones to be sort of a threat or are you just happy to see other publishing companies and other journals moving in this direction?

(Binfield) Absolutely, we see them as a good thing. It was a very happy day when Nature launched Scientific Reports. We felt that that really validated the model and that many others will be copying and as a result this model would propagate, open access would move faster and so on.

Also, it does make the healthy competition. It would be unhealthy if PLoS ONE were the only journal in the world. We need competition to keep us sharp. In addition, we actually support these other journals. I go out and give talks similar to this to the publishers of those other journals and try and tell them how we have done it and how they could improve.

(Q3) Can you tell us how PLoS is going to work on the further development of community journals such as PLoS Biology and PLoS Medicine in terms of their relationships with PLoS ONE?

(Binfield) The community journals were launched to demonstrate a self-sustaining open access business model. All four are now self-sustaining and they publish about 50 articles a month. They are good models for a small to medium-sized society journal.

But the only way that model could have changed the world would be to launch 25,000 of them and that was never going to happen. Other publishers like BioMed Central have gone that route launching 300 or 400 journals, but we then went the PLoS ONE route. We have no plans to launch any more community journals.

(Q4) You said that a journal could sustain itself if it publishes 50 articles a month and you also said that the APC, which supports the journal's sustainability, is not dependent on the scale of the journal and does not influence the number of articles published. But portions other than article processing costs are increasing each year, due to an impact of inflation and rising personnel costs. Are they going to be reflected in the APC? At the same time, unless research funds as a whole maintain certain levels, and hopefully increase, people who are supposed to pay the APC may not be able to do so. I am interested to understand the correlation of all of these.

(Binfield) Our APC for journals is really set by the rejection rate perhaps that those journals aim for. They are selective and they actually reject about 70% of their content. It is not so much a function of the cost and the cost of labor increasing and so on. It is more how many articles are published pay that amount. But we are a not-for-profit and I think we do have an expectation that at some point our prices will go down rather than up.

(Q5) I believe PLoS ONE was launched with a policy to challenge the peer-review process itself, but you have now adopted it. Do you plan to scale up PLoS ONE with an assumption that the peer-review process will function appropriately going forward?

(Binfield) We actually operate a very traditional peer-review process now with the simple difference that we only ask is whether it is scientific or not. We do not ask its impact. Other than that it is a very rigorous, normal peer-review process that any other journal operates. I would agree. I think that that is less radical than perhaps it was originally intended to be.

But our journal is run by the Editorial Board. We have 3,000 academic editors who handle the peer-review process for each paper and they have as it were brought us to a place that they are comfortable with and it is not productive to try and force them to be radical when the community is not ready to accept that. I think the more radical elements of that original idea now are being demonstrated in new products of ours. We have PLoS Currents, for instance, which is a much more radical experiment in peer-review.

(Q6) You said you expect that by 2020, 90% of all articles will be open access. But I suspect a bias might occur as whether someone is well funded or not will determine whether they can publish an article or not. Could you tell us how frequently you are reviewing OA author charges and also whether you have a discount program in place for the current OA author charges?

(Binfield) We are a non-for-profit mission-driven organization. We allow anybody to take a fee waiver for any reason. We have a belief that the ability to pay should not influence the ability to publish. In reality, over 90% of people end up paying the correct amount. But I think your broader question was if everything goes to this model, what is the ability to pay. We heard from the earlier presenter that there are ranges of journals with a range of APCs going down to $8 in some places. I think there is a lot of choice and there is a lot of competition.

For the question whether we have the discount program in place already for the author pay, anyone can take away for any reason so we have no need for a discount program.

(Q7) We can read articles of open access journals on the Internet, and I have the image that despite copyrights, these articles are being copied largely without prior consent of authors. What is your view about revenue from copyrights in the case of open access journals?

(Binfield) We use a Creative Commons Attribution license, the CC BY license, which allows full and unrestricted reuse even for commercial purposes. The author retains their copyright, but they giveaway all rights and restrictions for reuse. Anybody can use that content for any purpose. The person reusing simply has to credit the original author, the original source. I mean clearly we make no money out of retaining copyright in that way. Typically, nobody does. It is very hard to make a profit-making model on content which is completely free.

(Q7) There is an organization for copyright management, called the Copyright Clearance Center, in the United States. You have no relation whatsoever with that organization?

(Binfield) No.

(Q8) People from publishing companies are posing questions quite vigorously in today's Q&A session, and I am listening to them with an impatient feeling about whether libraries can play any role related to OA mega journals.
The question of discount was raised a few moments ago. Are you thinking of a business model where a discount is offered under a contract with an organization? In addition to the system where an author pays directly and individually, do you have another model where an organization pays in bulk for all authors who belong to it and gets a discount as a result? Another question – do you feel that you need to do promotion efforts? If you do, whom do you want to target for those efforts?

(Binfield) I think you are right. Libraries and librarians have an interesting challenge. At the moment, they are gatekeepers and they curate a collection of journals. In the future, that is perhaps not necessary because the authors themselves choose which journal to publish in and everyone can access it for free. But I think there are opportunities for librarians to create services that look at that cloud of content and curate that cloud. I think you are right as well that librarians could be the middleman who pays the publisher. They collect money from their university, apply to their authors, and pay the publisher on bulk.

Then, I think your last point is very interesting. In the current world, publishers market directly to librarians because they are the customers. However, in the open access world, the customer is the author. We need to reach authors who are in the university environment. Librarians could be that group of people that advocate, market, and promote our services to their researchers and explain why that is important.

(Q9) I have a question regarding the cascading. It sounds to me that the journal is trying to get more papers and trying to expand their market. Is it causing some problem or impact onto the scientific world in terms of lowering some reputation of the journal so that they are not using the same brand as their main journal?

(Binfield) I think it is dangerous to set yourself up as a journal of rejected papers. I think nobody wants to publish in a journal of rejected papers or the bottom of a cascade. As you said that has the potential to damage the brand of the publisher and perhaps that is why some of these publishers, Nature for instance, did not use that branding. At least on PLoS ONE, we absolutely try and avoid any association with being a journal of rejected

papers. As I said, most people see us as a first choice journal. My recommendation for anyone trying this would be to set it up similarly. I would not recommend the cascade model personally.