

## **Episode: Ethics Talk Videocast Transcript – Publishing Science and Accelerating Discovery During and After COVID-19**

Guests: Stefano Bertozzi, MD, PhD and Jeffrey Mackie-Mason, PhD, MPP

Host: Tim Hoff and Audiey Kao

Transcript by: Tim Hoff

Access the video and podcast [here](#)

This interview has been lightly edited for clarity.

Tim Hoff (Host): Welcome to another special edition of Ethics Talk, the *American Medical Association Journal of Ethics* podcast on ethics and health and health care. I'm your host, Tim Hoff.

This episode is an audio version of a video interview conducted by the journal's editor in chief, Dr Audiey Kao, with Drs Stefano Bertozzi and Jeffrey Mackie-Mason.

Professor Bertozzi is dean emeritus and professor of health policy and management at the UC Berkeley School of Public Health and Professor Mackie-Mason is Berkeley's university librarian and a professor at the UC Berkeley School of Information and the Department of Economics. They joined us to discuss sustainable models of scientific publishing that advanced bio medical knowledge and discovery during and after the COVID-19 pandemic.

To watch the full video interview, head to our site, [JournalofEthics.org](https://JournalofEthics.org) or visit our [YouTube channel](#).

Audiey Kao: Professors Bertozzi and Mackie-Mason, thank you for being guest on Ethics Talk today.

Stefano Bertozzi/Jeffrey Mackie-Mason: Thank you so much for having.

KAO: So would our COVID-19 response be better if all scientific papers were available for free upon its publication? The World Health Organization thinks so and over 150 scientific publishers, biopharmaceutical companies, and research institutions have temporarily made all published COVID-19 content accessible at no cost. Some believe that this pandemic marks the beginning of the end of the journal subscription model. What do you think a sustainable business model for scientific publishing looks like going forward?

MACKIE-MASON: Well, Audiey, there's two questions there: whether or not it's a good thing, and whether or not there's sustainable business model. As far as this being good for our response to COVID-19, the answer is surely a

resounding yes. There's so many critical things we don't know about the SARS-CoV-2 virus and the disease it causes, COVID-19. There're questions about how it's transmitted, how long it thrives outside a host, why it's so contagious, why it affects so many physiological systems and causes so many different symptoms, how well different public health strategies like masks and social distancing work. Not to mention the economic, educational, and social and political consequences of a modern pandemic and the different responses to it. This pandemic is so serious and it's causing so many health and social disruptions that more researchers want to try to discover the answers to these questions than for any other pressing question I think we've ever faced. And how does science make progress? It makes progress by learning from previous discoveries and building on those.

Now, many sites, scientists and doctors around the world can't afford and do not have subscriptions to all of the relevant scientific journals. If the relevant articles were published, open access with no journal subscription or article reading fee charge, we would get more, better, and faster answers to the critical questions we need to conquer COVID-19. It's impossible to prove that we would make certain discoveries that we otherwise wouldn't have made or would have made more slowly. We're scientists. We want to be careful about what we claim. But let's face it, if science is of any use, and of course it is in the face of a pandemic, making more scientific results available to more researchers and doctors, can't but help improve things.

BERTOZZI: So, Audiey, I completely agree with Jeff. I'm a veteran of the HIV pandemic and in many ways there's lots of similarities. But there's one very important difference, and the difference is that HIV moved like molasses and compared to this pandemic. So while you could make the same arguments Jeff did for publications about HIV, especially early on in the pandemic the temporal urgency was nowhere near what it is for this pandemic because the transmission rate is so much faster, and we see things changing on a daily basis.

So you know the other day, I was reviewing a manuscript which was looking at the comparative effectiveness of different types of masks. Now, if those findings are borne out, what that paper suggests is that if you use a fleece neck gaiter as your mask, you actually make things worse because the fleece neck gaiter turns your larger droplets into smaller droplets, which stay in the air longer. Right? That's a really important finding, and if that that kind of finding waits for three to six months to come out in a usual publishing process, and then a large proportion of those go into the proprietary access publishing world, we miss an opportunity to act on findings quickly, and so I completely

agree with Jeff, this is more important and more urgent than it has been for previous situations, including the one that I know best which is HIV.

KAO: Yeah, well given what you just both said, what then is the quote unquote "sustainable business model" for scientific publishing then, if not for journal subscription?

MACKIE-MASON: Well, the greatest value in scientific research publications comes from the research itself, which is performed by the scientists who are not being paid for the articles. But there's also some value of course provided by the publishers. What they do takes work, and they have wages to pay. If we want high quality scientific publishing, we need to find some way to pay those costs. However, we don't need to pay the publishers with reading fees or subscription fees which block people from getting access to the results. Instead, we can pay the publishers for their publishing services that they provide. That is, pay them upfront to do the work needed for quality publishing. And then once they've done their work, make the articles available to everyone to read for free. This business model, and it's the only one that economically makes sense in a digital world, is "paid to publish, then read for free." It's clearly sustainable. The world is already providing enough funds - about 10 billion dollars a year that pay for scientific publishing. All we need to do is pay that 10 billion dollars a year in publishing fees rather than in reading fees. The industry will get the money it needs to do its publishing services, and everybody who wants to read the results will get access to the results without putting a paywall in front of them.

KAO: Sure, so if I can just follow up on what you just said. earlier this year, the White House Office of Science and Technology Policy issued a notice seeking public comment on waves of speeding up access to publicly funded research results - taxpayer funded search results. Some argued, however, that results were immediately available at no cost. America would be giving up intellectual property for the rest of the world for free. So how should one think about the issue of intellectual property balanced against the need to speed up access to scientific information?

BERTOZZI: So I wonder whether you let me make a comment about the previous question before we get to this question of intellectual property. Because I agree with Jeff that it doesn't matter whether you pay after you pay in advance in terms of the total amount of funds. In fact, it could easily be that it's more efficient to pay it upfront. The question is for whom does that potentially create barriers? That is an important question, and because I work disproportionately in the developing world, it creates a different kind of a problem where publishers who, in the past may have given preferential

access to published materials to libraries and researchers in developing countries, need to be able to mirror that kind of differential pricing policy for creators of scientific knowledge in the developing world, and we have to work out a different way to make that work. I think the other thing to consider on that is that to the extent that creating that publishing platform which enables researchers to quickly disseminate validated research findings, it's a public good and the role of the public sector in supporting a public good of that sort is something that we also need to consider. To a large extent, it is, in fact, funded by the public sector through academia, but we need to think about whether that's the most efficient way to do it, or whether, in fact, more direct support such as you might have with federal support to National Public Radio kinds of support might be something to consider. I mean, I would certainly want the NIH to consider the possibility that instead of funding publication through the university and through their research grants, it does some direct funding of public good research platforms directly, as an example. So on the IP issue, though I want Jeff to take the first whack at it because he's much more in tune with that than I am.

MACKIE-MASON: Personally, I think the argument that the publishers presented in response to the Trump Administration call for information which you mentioned, Audiey, was... I think the argument is nonsensical. I'm just going to call it what it is. I think it was offered by the publishers as political pandering to the Trump Administration, which frankly is obsessed with trade wars. When American researchers publish in subscription journals, no money comes back to the researchers to compensate them for their intellectual property. The subscription revenues all go to the publishers. And as if that weren't enough to dispense with this argument to a large extent, the publishers collecting these payments aren't even US companies. The top four publishers in the world top four scientific publishers publish over 50% of scientific articles. Of those 4, three of them are foreign. The number one, publisher, Elsevier, is Dutch. The number 2 publisher, Springer Nature, is German, and the number 4 publisher, Taylor and Francis, is British. So these are not American revenues that we would be protecting, and, in any case, they're not intellectual property revenues, they're publisher revenues.

Scientists give away their discoveries for free for the most part, and that's what we as scientists want to do. We want the whole world to benefit from our discoveries. That's why most of us are scholars. There is an exception, some discoveries with commercial value are patented so that the scientist or her employer can earn a return on the investment in research, but patents require the publication of the results and, in fact, patents are freely readable because their government documents. Open access publication doesn't undercut

patents. In the least it's completely... it has nothing to do with it. It's just a It's a nonsensical argument.

BERTOZZI: I completely agree with Jeff, and that's why I was happy that he laid it out so clearly. It's just a true, true, unrelated kind of a situation, right? If there is intellectual property that needs to be protected, that has nothing to do with publishing. I mean, it's just frankly ridiculous. Of course scientists will sometimes delay dissemination of a new finding while they file for patent protection or things of that sort, but that doesn't change in any way with what the publishing model is, and I agree with Jeff, and I would have called it out that I think it's literally pretending that this is a trade issue when it really isn't at all.

KAO: Yeah. One of the scientific trends that we've been observing is one of the preprints. Preprint research findings are now freely accessible via various preprint platforms before a manuscript is accepted for peer review publication. So through mid-September more than 9000 preprints on the novel coronavirus have been posted to bioRxiv and medRxiv. What are the opportunities and challenges that preprints pose to advancing biomedical knowledge and discovery?

BERTOZZI: Well, I have to say that for me, this is definitely a glass that's more half full than it is half empty. But as you point out, there are opportunities and challenges with this new model. So, when I trained as a scientist in the biomedical space, you basically had to keep things secret until they were submitted to a prestigious journal because the prestigious journals weren't interested in publishing anything that was already in the public domain.

KAO: Right.

BERTOZZI: The other half of me, the economist, has a completely different experience because in the economics world, you have to present your paper over and over again and multiple for-public fora and get feedback on it before it's considered vetted enough by the community to warrant publishing in a top journal. I mean, it's an opposite publishing model between the two areas, so I've always been struck by this seeming lack of logic in in the publishing world. What has happened dramatically in the biomedical world is that that requirement disappeared with the advent of the preprint servers and one after the other. The major journals agreed that publishing something on a preprint server was not disqualifying in terms of subsequent publishing. So in economics, for example, the *National Bureau of Economic Research* has been publishing working papers for years, which in fact is not very different than a preprint server, right? It's making the scientific work accessible to the

broader community before it's published in a peer-reviewed Journal. I think that *NBER* is a little bit of a hybrid because it's, you know you can't just randomly anybody post something on *NBER*. There's a filter mechanism there which is in some ways akin to a preliminary peer review, but that's not the case with medRxiv or bioRxiv. And therein lies the rub. So while it's wonderful that somebody can finish up manuscript and post it on medRxiv and the next day policymakers can act on it, at the same time there's zero filter so that the policymaker or the journalist or the tweeter understands whether this is valid scientific work or something that needs to be debunked.

And the reason that I got involved in in this new experiment with the *MIT press* with this new journal is precisely to address that deficit right where we want rapid peer review of things that are already in the public domain so that the things that are validated can be accelerated in their spread, and the things that are that need to be debunked are not further disseminated as if they were valid. So I think that's... therein lies the rub in the sense that you have the potential for things getting posted on what looks like a legitimate academic site as if it had the same validity as the journal that comes out in *The Lancet* or the *New England Journal of Medicine*.

MACKIE-MASON: Yeah, I agree with Stefan, this completely... and also as an economist, I had the same experiences he did in my early years. Of course, in my early years, and Stef, when we did pre-print distribution with standard back then 35 years ago, we distributed our prepaid by photocopying our manuscripts and mailing them at the post office. It was a very primitive system. And it actually is an important point because now with digital distribution, preprint servers can get things out instantly, and that is a great win. We can get distribution of scientific results much faster, much wider, much more easily. But as Stef says, there's a, there's a big trade-off, and the trade-off is that if you get them out there before they're fully vetted before they've been peer reviewed, there's a more greater chance that they're wrong, there's mistakes. And the peer review process is so critical for academic science, it really is the bedrock of us making sure that we make good decisions. But there's no perfect solution. Ideally, all results will be peer reviewed before we rely on them so that we have some reassurance that correct. Peer review can make mistakes, too, but it's a good process to help. But also, ideally all results would be available sooner, so we can solve problems like COVID-19 faster. So it's a tradeoff. Good scientists and doctors and policymakers will read preprints, but they'll be more skeptical, and they'll be more cautious before they rely on them. They'll recognize that there's a greater chance there are errors, and so they'll be a bit more careful. But they'll

at least see what's out there and use that and take that into account in their research and their policy and their practice.

KAO: So Dr Bertozzi, you just mentioned a moment ago of this new MIT press journal, for which you are the editor in chief, is called "*Rapid Reviews: COVID-19*," which has been described as an open access overlay journal that seeks to accelerate peer review of COVID-19-related research preprints and prevent the dissemination of false or misleading scientific news. Can you explain a little more about how this journal aims to achieve this editorial mission and what do you see as its main challenges to achieving this mission and how they can be overcome?

BERTOZZI: Well to start with, it's a communications challenge because it's really two efforts in one, and I don't think that we've done a good enough job of explaining that. So the first effort is a rapid peer review effort, which is open, transparent, and in the public domain. So what we do is we look every day, every week, at what's coming out on the preprint servers, the two that you mentioned, as well as many others that are publishing in the social sciences and the humanities. We consider things like the *National Bureau of Economic Research* to be the equivalent of a preprint server. And by scraping those we add, we identify manuscripts that we think are either important because our pre-reviewers, if you will, think that they have the potential to have significant impact or because we notice that they are already getting attention, right?

So it could be the manuscript went through, we didn't think it was especially important, but then two weeks later, it's lighting up the Twitter-sphere. You know, and in that case, that means that people are paying attention to it, and whether, and if it is, lighting up the Twitter-sphere and people are paying attention to it, then it's important that it be validated, or debunked. So we can either filter things when they first come through, or we can pick them up once they start to be noticed by the media, by scholars who are citing the work, or by social media.

And what we do is we send out those manuscripts for what we hope is a very rapid peer review. We're giving our reviewers a one week turn around time, and we're asking them to do a rigorous peer review, which is scientifically rigorous, but not necessarily as complete as a traditional review that would be suggesting all kinds of editorial changes, future experiments that might be done, et cetera. What we want is a review that says, "These are the main claims of this manuscript, and we think that they are very valid. Not so medium valid, not so valid, etc." So that that information can be publicly available very quickly.

Now our second activity is that if we receive strongly positive reviews of a manuscript, we are going to offer those authors the opportunity to publish that manuscript with us. Now that's a very unusual model because normally an author submits their manuscript to a publisher. In this case, it's like we're offering you the job, but you never applied for the job, right? And this is a little bit tricky because some authors are very nervous that we might somehow interfere with their ability to submit their manuscripts somewhere else. And so we're trying to be very clear about the fact that the reviews that we're doing are open, transparent, public, and they can take those reviews to any journal they want. In fact, they don't need to take them, the journal can find them. And if that makes their review process more efficient, all the better, right? So that we think of this as one of the public goods that comes out of an effort like this is that somebody publishes something, 2 weeks later, there are reviews published, they take it to the *Journal of the American Medical Association* and the journal decides that because of the reviews that already exist, they need to do fewer reviews and that might even accelerate publication. It also might be that looking at the reviews they make it more rapid decision about whether they want to accept a paper for review or not.

So in that sense, it's purely, sort of, not directed toward our own journal, but then what we have to remind people is that if we do offer them the opportunity to publish with us once they say yes, that's the equivalent of submitting their manuscript, right? But until they say yes, they haven't broken any of the of the journal publishing rules, although, I certainly do understand why it's a confusing topic in process and we've been getting, you know, various feedback about the clarity of our communication.

KAO: I appreciate that clarification. I think it's something that's pretty innovative and new to our learning audience. So as we near the end of our conversation back in June, the *New England Journal Medicine* retracted an article entitled "Cardiovascular disease, drug therapy and mortality in COVID-19" because of concerns about the integrity of study data. Now open journals have require sharing of raw data for a while, but compliance has been spotty. What do you see as the future of open data?

MACKIE-MASON: It's a maturing area. We want to share data for two reasons. One is so that people can verify our results before relying on them. It's like the peer review process. If you're doing database research and people want to know you got it right, they would like the data so they can test, make sure you got it right. The next is so that they can use that data to run new tests and experiments to build on the science. This is a longstanding concern in the profession. The data should be available, but as the amount of data has

exploded and more and more research is based on data, policymakers and academics have gotten more serious about this, and we're trying to standardize data policies for making the datasets reliable and shareable and findable, but it's very immature. There actually turned out to be a lot of challenges. It's not as easy sharing data as it is to share text. There are challenges with having privacy protections, for instance in health data. There are challenges in documentation. We don't know how to document data reliably or comprehensively so that we can replicate science and a lot of other challenges. I won't go into them all. It's an area that's maturing, but it's so important it's so critical to, again, the advancement of knowledge and the advancement of science as we rely more and more on the greater availability of data that we're going to solve it, or we're going to certainly make it better and we are making it better all the time. So I have great hopes and expectations that open data and reliable data storage for scientific work will be very far along in the next decade. We were already making progress and it's so important. But these days, mistakes can happen. It's quite immature.

BERTOZZI: I agree with Jeff, I think that we're in the beginning of a transformation. So, if I generate data in a research study, and then I lock those data up so that they are not available to other people, that's very costly because we are not reaping the benefits of other people using data that has already been created for additional science, right? The flip part of that is that when I generate data for my own use, it's not ready to be shared. It's not ready to be shared because other people can't understand my data without significant documentation and explanation for how those data are organized. And if I don't do a very good job at that, then they'll be asking me for lots of technical assistance to figure out how to work with my data. Now, historically, scientists haven't been compensated for either of those activities. Either the activity of transforming the data that's suitable for your own use into data that's suitable for other people to use, nor for providing technical assistance on use of the data that you've generated.

But while that is an additional cost, it's not as big as the forgone benefit of locking those data up. So I think what that really means is that the funders have to, on the one hand, obligate people to share their data so that they can be used by others, and at the same time provide the funding so that that's possible without detracting from the research, and I agree with Jeff, that's an evolving process. And, you know, it's harder in some areas. I mean, I work in health which has lots of privacy issues. And I think that there are interesting solutions to that, such as, for example, some data being stored in a central warehouse where other researchers can query the data without necessarily owning the confidential part of the data. And these sort of service firewalls,

I've called them is, I think, an interesting innovation in the health or privacy protected data space.

But the other problem is I work a lot internationally where you're dealing with national government data or performance data from health systems or educational systems and countries who don't necessarily want to put that in the global public domain. They're generating those data to improve their own performance. They don't necessarily relish the idea that their flaws are open to the world. So there are privacy issues, confidential issues, and willingness to provide access issues which are really, really complicated.

KAO: So on that note, I want to thank Professors Bertozzi and Mackie-Mason for sharing their deep insights and expertise with our audience. Thank you both again for being guest on ethics talk today.

MACKIE-MASON: Thank you.

BERTOZZI: Thanks for inviting us. It was a real pleasure to be with you.

KAO: For more COVID ethics resources, please visit the *AMA Journal of Ethics* at [JournalofEthics.org](https://www.journalofethics.org).

And to our viewing audience out there, be safe and be well. We'll see you next time on Ethics Talk.