

On June 11, 2012, Holden Karnofsky and Cari Tuna spoke with John Ioannidis

Holden: There are a couple of topics that I want to talk with you about. The first topic is the Cochrane Collaboration because we're interested in that group and want your perspective. The second topic is what we're calling meta-research. I've heard your name come up a lot in this context, I really like the work you've done, and I'm interested in your thoughts about what a funder could do to be helpful and speed along improvement in this area.

John: I think these are very useful topics to discuss.

Holden: Let's start with Cochrane. The story that we've been hearing is that they need a lot more funding than they have, especially for the basic infrastructure and for things like the US Cochrane center. I don't know how well positioned you are to have an opinion on that, but what do you think about the group in general and of the idea of funding them?

John: I have a very positive feeling about Cochrane Collaboration. I have been involved with them. Although I have not been a key player for many years, I was probably among the first people to be involved with the Collaboration. It's a non-profit group with very strict rules about what kind of funding they can get for the review projects that they perform. That creates a set of limitations for it, which I think are appropriate, because the authors cannot be objective when performing a systematic review if they are funded by sponsors of whatever's being appraised in that review. I think that the funding situation is probably more difficult in the US compared to some countries like the UK that have more of a long standing appreciation of the need to synthesize evidence in order to arrive at a rational decision about what should be used and what should not be used in medical care. I think that probably some of the people in the US Cochrane centers have tried to support their centers through some parallel initiatives like the evidence-based practice centers (EPCs) that are performing systematic reviews for projects that are sponsored by AHRQ. But I'm not sure whether they've had any major funding directly covering their infrastructure as Cochrane centers. So I think that supporting that initiative would be a good idea. I think that the people are excellent. It's an organization that has recognition and participation around the globe. I think that they're doing something very useful and I can't see a reason why they should not be supported for what they do.

Holden: You raise the issue of these evidence-based practice centers. That makes me wonder – one of the things we've been trying to figure out is, "Does Cochrane have competitors in a sense, other groups that are doing similar work?" One of the things that we've wondered is why the group AHRQ which funds the evidence-based practice centers, doesn't support Cochrane with significant infrastructure support. We haven't gotten a response from that group yet. Do you have thoughts on the evidence-based practice centers? Is their work as good as Cochrane? Are they a reasonable substitute for it?

John: I think that there's a lot of overlap between these entities. Many, if not most, of the people who are at the EPCs have had a relationship with the Cochrane Collaboration at

some point. Some of them are leaders in the Cochrane Collaboration, others came and went and preferred to have more independent routes in their work. And I think that AHRQ has tried to create its own network through the evidence-based practice centers. That's also a nice idea. I'm not saying that only one should survive and not the other. I have not been involved in the mainstream of either Cochrane or the EPCs because my type of work is a little more unusual, rather than focusing on doing specific systematic reviews on one topic. I think this is very important work and I have enjoyed working on single, topic-specific reviews, but I cannot give this activity a lot of priority for my personal current research interests. I like to study some more high-risk bigger picture questions.

Holden: One of the things I wonder is: if Cochrane is getting funding in other countries and AHRQ is doing its evidence-based practice centers in the US, where is the call for more, why is more needed? Is it the case that more funds are needed and specifically in the US? How would you go about answering the question? Do we basically have enough resources to get the most important systematic reviews done? Or is it important that there be more of a presence, especially in the US?

John: I think there's definitely plenty of room to have more. It's an issue of how to prioritize research funding allocations. But probably the ratio of funding given to this type of activity vs. just producing another piece of information with *de novo* data collection and *de novo* samples is very suboptimal.

I think that in many fields it's not that we don't have enough data, it's that we have too much data and don't know what to do with it. The highest priority would be to integrate what we know already and try to see if it's enough. Maybe it's more than enough. Even if it's not enough, we could learn exactly where the investments should be for the next step. Maybe we shouldn't be spending money on ten different directions but just focus on one or two that seem to be where the most formative research would be. So I think that this sort of work is extremely useful. I think that it's underfunded, and it's even more underfunded in the US because there's less of a tradition of that sort of research being important. I think that people in Europe, particularly in the UK but also Canada, Australia, New Zealand, have a very high appreciation of these tools of evidence-based medicine. In the US, the research has been more focused on early translational stages. So T0 research is absorbing the vast majority of funds. People just want to make new discoveries and they don't care if these discoveries will be used, how they will be used, the benefit of patients, and cost-effectiveness.

This is probably why Cochrane is having a hard time to sell their work to traditional funding agencies.

Holden: You said that you think that there are fields where we have a lot of data and we just need to pull it together. Do you know offhand what some of those fields are? I take it that you think that there's still a lot more room for this kind of work.

John: Yeah, so Cochrane for example is focusing mostly on clinical trials and in

particular randomized evidence, although they have tried to expand a little bit to observational data – on medical interventions. I think that they're very comprehensive about that. They try to cover all different types of interventions and specialties. But this is only one part of the biomedical information agenda. There are many types of studies and probably the minority of public money spent for research is spent on randomized trials.

My perspective is that it's even more useful or as useful to try to understand patterns of research across very different fields. That could include, for example, diagnostic tests, prognosis, predictive tests, animal studies, laboratory research and imaging studies. These are huge research areas in which sometimes more than a million papers have been published. There's far less of a tradition of integrating evidence. It's not the same in all fields; there are some fields that probably do it. Yet many of the problems that we have recognized through meta-research on randomized trials – which is the work that Cochrane is doing – are the same or worse in other fields in biomedical investigation.

I would argue that even beyond biomedical investigation, every empirical field of research is likely to be affected more or less by similar biases that affect empirical collection of information, design, dissemination of that information and integration of that information. This is what I think is worthwhile expanding into and this is what I see as a goal for meta-research: trying to understand research processes empirically and at a large scale for very diverse types of investigation.

Holden: Do you see other groups out there that you think would do as well as Cochrane or do you think Cochrane is the group that's most likely to bring systematic reviews into other fields?

John: I think that Cochrane is extremely unlikely to do this, because they have a very specific mission which they're doing their best to accomplish and it's a huge mission. They have already collected a huge data set of clinical trials and they have come up with about 7,000 reviews. I think that with about 10,000 or 15,000, most, if not all, interventions in medicine would be covered, but then obviously one needs to keep these updated. So Cochrane has a very specific mission. Some of the people involved in Cochrane may have some interest in venturing into other areas as well. It varies from field to field. Sometimes you feel like you are the only person who is interested in taking a big picture look at that field and other times you realize that there are many others who share the same concerns. I cannot say that there's one pattern across all sciences.

Holden: You mentioned that you think that this T0 research is a big focus and then other kinds of research don't get as much attention. Why do you think that is?

John: Maybe it's a misunderstanding of what research is about. Probably there are just too many people who have been crowded into T0 research. Maybe they get powerful positions as department chairs, leaders of institutions, grant reviewers and journal editors and they feel that we just need to do more of the same. I think that there are far fewer people who venture into trying to understand larger pieces of evidence and more overarching pictures of where the evidence in a given field lies. There's also some risk

about that, because when you take a view of the bigger picture sometimes you realize that some research fields are operating at levels that have extremely low efficiency and extremely low performance. Let's say that you realize that the field has had a replication rate of 1%. This is not very good news. So you need to have people in the field realize that and be willing to do something about it, improve their reproducibility practices, their validation efforts, create larger collaborations and consortia with transparency and data sharing. The willingness to do that is not always there. Some people are happy to just continue doing what they used to do with very low efficiency and very low replication rates, but they get promoted, they get funded, they review their peers' grants and papers and life goes on.

Holden: If you were us and you were trying to figure out which subfields of medicine are in good shape and which ones are not doing so well, is that something that easy to determine? Are replication rates easy to get?

John: I think that in order to have a comprehensive picture one would need to create a much larger number of meta-researchers. I think that for most fields, the number of meta-researchers is extremely small at the moment. One reason might be that there's not really any good funding for that. Unless you have a critical mass of people who are willing to take big picture shots of the evidence and map many different areas I don't think that we'll be in a position to tell across very different fields which ones are doing better and which ones are doing worse.

Holden: I see. So there are scattered people out there right now...

John: Right. The Cochrane Collaboration currently has thousands of people who are involved in one way or another and this is why they have accomplished their mission to a good extent, capturing 7,000 systematic reviews. But in other fields there's not a very large number of investigators doing this type of work.

Holden: What do you think about T1 research? I've heard it argued that in addition to overlooking systematic reviews and meta-research, the NIH also underinvests in bench to bedside research. What do you think about that kind of hypothesis?

John: I think that the numbers decrease once you go from T0 to T1 to T2 to T3 to T4. I'm not trying to say that one translational stage is better than the other. I think that we need all of those. It's just that the amount of money, manpower, time and investment is sharply decreasing as you move away from T0.

Holden: And you feel that it's needed in perhaps equal amounts; that the decreasing is because of underinvestment?

John: Right. Perhaps not equal amounts, because obviously you don't expect all discoveries to move through the whole process. You'll have attrition no matter what. And the attrition is most likely going to be severe. In some empirical studies that we've done, we've seen that maybe 5 out of 100 major promises published in the most high profile

basic science journals eventually manage to materialize. And even that is probably a select sample, so maybe the success rate is even lower than that on average. This means that probably there will be some decline as you move across translational stages. But regardless of what stage we're talking about, there's meta-research. You need meta-research for all of these stages; for example you need meta-research to appraise T0.

Holden: When you say “meta-research” are you talking about systematic reviews?

John: Right, you can think of systematic reviews not just for clinical trials but also for laboratory studies, animal studies, biomarker studies, and genomic association studies. Having worked in several of these fields, I've found it very revealing to see how you can map the patterns of biases and inefficiencies that exist in different processes, how you can make some suggestions, and how you can improve things by paying attention to the steps that seem to have the highest inefficiency.

Holden: We're basically looking at the world of all research, but especially medical research, and thinking about, “How can we be helpful, how can we help the whole system run better?” We have this impression that making the whole system run better is not as popular as trying to go after a particular disease and that therefore we might be able to accomplish more good with the same amount of dollars by focusing on it. So what would you be doing if you were us? Where would you be looking and what ideas are important for improving medical research?

John: I think that one has to find ways to convey the message to funders and people who are interested in supporting research that while it's very nice to have researchers who try to come up with something new, over the years research has moved away from the situation where there are very few researchers to a situation where we have a very large number of researchers who are publishing scientific papers, and it's not enough to just support their individual work. It's important to tell what the totality of the evidence is and take very abstract looks at a large scale at what's happening here. Otherwise we run the risk of funding literally a million studies in one direction and getting nothing out of doing so.

The number that I'm using is not an extreme number. There are some fields like some areas of observational epidemiology that probably have more than a million studies and what we have done there has mostly been leading us down the wrong path. If we had been able to look at that big field in its totality, maybe we could have decided early on that we need to fix a few things. Otherwise we're wasting billions of dollars and hundreds of thousands of scientists' time on that and we could have done better. This is a message that some people may have difficulty hearing because they want to hear “we will make big discoveries about diet” or “we will make big discoveries about exercise” or about specific drugs or the treatment of specific cancers. They may have a very specific aim in their mind, while meta-research usually involves looking at very large-scale pictures of the evidence. You're trying to improve a whole field of research which could include hundreds of drugs or thousands of associations. Obviously the impact of meta-research could be larger in that regard.

Holden: Right. We largely have the same intuition, but the issue is that we're not very familiar with the field; we got interested in this idea by our observations of the problems with development economics. As somebody who knows the field a little better, where would you be looking for specific opportunities to make a difference and push along meta-research?

John: Meaning what kind of funders?

Holden: We're sort of a funder; we're recommending funding opportunities. So if you were a funder then where would you write a check to accomplish good?

Cari: And Good Ventures, the foundation that I run, is a funder and is interested in this field and therefore specific opportunities we could act on.

John: Right, so I think some themes would be trying to appraise biases in different research fields, in particular biases in the domain of significance chasing, including publication biases, selective analysis and selective outcome reporting, and even fraud, although fraud I don't think is a major component. Looking at what I call "vibration of effects": the capacity to change results using different analytical protocols so that they fit a particular expectation.

Another field would be that of improving reproducibility. One aspect of this is promoting open access to raw data sets, protocols and analysis codes. Another aspect is reducing the time to replication of research discoveries and making replication processes more robust. Maybe it's also worthwhile to look at dynamics of information. Some fields have a strong preference for "positive" results but in other fields it may be considered very important to be able to kill a proposed discovery. There's something termed the "Proteus phenomenon" which is when you have a very extreme claim and then a very extreme refutation coming very shortly thereafter. This doesn't occur in all fields. Some fields like to get significant results and they're happy with that. Other fields have more of a built-in emphasis both on discovery and on refutation.

Holden: Interesting. So you're saying that there are sometimes incentives to data mine for refutation, not just for positive results?

John: Right. I think we have some indicative examples. Most fields like significant results. I think there's no doubt about that, many people (including myself) have shown that. But there are some fields that also like refuting results. For example, in pharmacoepidemiology there's sometimes an attraction to not finding an adverse effect. There could be different motives for that. If it's a study done by the pharmaceutical industry then maybe they want to show that their drug does not have a toxicity problem. It could be completely unrelated to that, it could be that an investigator wants to "kill" a discovery that was made by another team, showing "no, you were not correct."

In some other fields like genetic associations we've seen this phenomenon where first

there is a study finding a huge association and then within a year there is another study showing that there's no effect. This is likely to happen when (a) the proposed discovery is not true or exaggerated and (b) the replication study can happen very quickly. If you need ten years of follow-up you cannot refute the discovery in a timely fashion. But if the replication effort can be done very quickly, let's say it's just measuring a few samples that you already have stored in a refrigerator and you can do it over night and write the paper very quickly, then that Proteus phenomenon will happen. We do see that happening in fields that have very rapid turn-around in terms of their ability to measure things. In genetics we have tremendous measurement capacity, we can measure millions of gene variants overnight literally. There we see that back and forth between proposed discoveries and refutation happening.

Holden: So you think that there could be some legitimate positive results there that are getting covered up by refutations that are driven by publication bias? Is that what you're saying?

John: It is possible.

Another area is registration, an area that crosses many disciplines. I think that for optimal functionality, different fields would have very different requirements. For randomized trials registration of the mere existence of a trial is useful, although it doesn't solve the problem of selective outcome and analysis reporting. A sponsor may register a trial but then use the analysis to come up with whatever results they want. So unless you also register the protocol and analysis plans, registering trials not going to solve the problem.

In other fields there's often no protocol. There's no protocol in much of nonrandomized research both in medicine and in other fields. Astrophysicists looking at the collection of data from telescopes in the galaxy may go wild exploring observations and come up with something interesting and then build a protocol around it. There you would probably need registration of data sets and opportunities to perform studies rather than registration of particular protocols.

Holden: What do you mean by that, registration of data sets and opportunities to perform studies?

John: Let's suppose you know that there is a telescope out there that has collected that information and you know the people who can analyze it. This is a registration of the data set of that telescope. The same thing could apply to a bio-bank or information that will be collected from a cohort study.

Holden: Are you saying that the observational scientists register their predictions and protocols in advance of the data being released?

John: You can ask for registration of the data set for observational studies. As I said, much of the time the protocol will not be there upfront.

Holden: What would registration of a data set accomplish?

John: It would tell us what is available for people to explore. It makes a difference whether there are a hundred data sets that people could be performing analysis on or just one. If there's just one data set that leads to a publication, then that's it. If there are a hundred then you don't know how much selective data peeking or opportunistic observations may have occurred.

Holden: What fields do you think are most desperately in need of meta-research right now? What fields are in need of things like efforts to promote registration and data sharing?

John: I think it's difficult to prioritize. All fields in biomedicine, but even beyond biomedicine, all fields with empirical research should have meta-research efforts attached to them. If you want to prioritize, then maybe two parameters are "What is the investment in that field? How many papers are published in the field and how extensive are the funds that are spent in that particular field?" and also "What is the perception of the community of researchers and possibly regulators about what kind of an impact that field has?" If everybody says "oh that's very nice but I'm not going to take it seriously really" about a field then probably it doesn't matter. But I think that fields that people are using to inform decisions, clinical care, public health or the next research steps should have high priority.

Holden: How would you investigate that second part, how would you figure out which fields are influential?

John: That's an interesting meta-research question by itself. You can ask experts, but I'm not sure that this is the best way to answer it. Probably the best way to answer it would be to look at large segments of the literature and look at what its impact has been on subsequent research. I have some gut feelings about this but they may be completely wrong once you look at the data more closely.

Holden: I'd still be interested in them, just as a starting point.

John: That would require creating a map of science and the directions of influences. You can take a field like biomarkers where we know that there are about half a million papers and try to see what kind of fields have been influenced by that research. Is it a field that has pretty much created a hub on its own and nothing else has happened outside of it, where people just publish papers and they're happy with that? Or is it a field that has made claims for applications, for changing medicine and so forth?

Holden: I also wonder if some of the fields that are not influential are not influential because they're most in need of more credible research methods. There are some fields in which it's known that the methods are very poor and so people ignore whatever comes out of them, but it doesn't mean that the fields are not important. They may present the best opportunity to improve methods and make a difference.

John: I think you're definitely right about that. It could be that the inability of some fields to make a difference has been a function of their inefficiency and poor replication patterns.

But at least the first criterion that I mentioned about how many scientists are working in a field and how many papers are produced in the field and how much funding is given to the field is one criterion that would be relatively objective.

Holden: Then there's this other question distinguishing fields where there's already been a lot of progress in improving practices and fields where there's still a long way to go. Do you have any idea of how to gauge the state of the fields on that metric? What the prevalence of registration, protocol registration, data and code sharing is in a given field?

John: I think it would be worthwhile to map that. We've done some empirical work, for example we wrote a paper where we looked at the 50 most widely cited journals across different scientific fields and tried to map what kinds of policies they had for requiring reproducible research including access to raw data, protocols, and data sharing. There were a lot of policies in place. Most of these journals had some, if not full coverage on reproducibility issues. But when we looked at specific papers that were published in journals with these policies, the policies were not implemented most of the time. So there's also an implementation gap.

Even when they're implemented there are still problems. A couple of years ago I published a paper in Nature Genetics where we got four teams of microarray analysts and looked at papers in Nature Genetics which is the most highly cited genetics journal. It has had a reproducibility policy since 2005 where it requires authors of papers in the journal to make raw data, protocols and analysis codes available in public. We took 18 papers and we tried to reproduce what they had published in figures and tables in these papers. We could do that for only two of the 18 papers. This is the best journal, with the best teams of scientists publishing there, with reproducibility policies in place, and nevertheless the reproducibility rate is about 10%, which is pretty scary.

Holden: Do you know particular groups or people working on these issues that you would recommend funding?

John: I think it depends on the field. I think in each field there are some people who are interested in doing this type of work, but as I told you there are not that many. It's not work that's traditionally funded. There's no NIHS that I know of that have made the work a part of their mission.

Holden: This is where we're trying to figure out where we come in. Good Ventures is a funder and GiveWell makes funding recommendations. We hear a lot of ideas and there are a lot of things we'd like to do. We'd like to promote more preregistration, more adherence to replicability guidelines and more sharing of data and code in a way that makes the studies reproducible. But we don't know where to write a check to make those things happen. Who are the people working on these things, who are the organizations

working on these things, if there aren't any then what can we do to get some people and organizations working on these things?

John: Are you thinking of a particular research area?

Holden: Not really. Our way of approaching giving is trying to do the most good possible and so whatever fields have some combination of the greatest humanitarian value and the most room for improvement on meta-research issues, and perhaps a lot of money in them – those are fields that we'd want to start with for improving. But we would also be willing to start where it's easier to begin, learn as we go, and start extending to other fields later on.

John: I think that there are a lot of people who would be interested if they knew there were funding for this type of work. I have a very large team that has been working on meta-research. There are people from all over the world, not necessarily from Stanford, who have been interested in working on some project. They would approach me and they would spend some time with me at Stanford or in Europe and some of them continue to do that kind of research even after moving to another institution. You have scattered people who have some expertise, who have some interest, who have worked on some projects in that direction.

Holden: How do we find and bring together those people?

John: Probably by looking at the literature – looking at who is doing this type of work and what they publish. I'm sure that there will be many people, although probably most of them will not be fully committed to doing this type of work at this point for the reasons that I explained.

Holden: Do any people jump to mind as people who would be particularly good for us to talk to about these issues?

John:

- Some of them are at NIH or CDC, because that's a type of research career you can have when you're interested in the big picture.
- Muin Khoury has a joint appointment with the CDC and NIH. He's interested in knowledge integration issues. I've known him for many years; we've worked together for a long time.
- Neal Young at NIH in the hematology branch. He and I have been interested in issues of publication processes. We have been thinking about how you can apply game theory modeling to the publication processes and selection of particular types of results.
- Thomas Pfeiffer who I've worked with on selective reporting modeling. He was at

Harvard and recently moved to New Zealand.

- David Chavalarias – he’s at the Ecole Polytechnique in Paris. He’s interested in large scale mapping of research fields, for example we had done some work together looking at mapping 235 types of biases in medical research. Currently we’re working on trying to model the evolution of p-values across the biomedical literature in 20 million papers.
- There are several teams interested in improving the reporting of research. David Moher at University of Ottawa has been a leader in that field. I’ve worked with him on several projects. He’s been a leader in reporting guidelines for randomized trials and other designs.
- Doug Altman, the director of the Center for Statistics and Medicine at Oxford. He’s interested in the use and misuse of statistics in different fields and he’s done great work in several fields. We have overlapped a little bit in predictive models and randomized trials and meta-analysis related themes.

There are many people. I can probably come up with many more names. I think if you start probing the literature you’ll see who has done what and what the direction of their research has been.

Holden: Do you mind if I name a few groups that we’ve run across and ask if you’ve heard of them and have a view on them?

John: Yes, tell me.

Holden: One is the Equator Network, which I believe works on quality of reporting.

John: Doug Altman leads The Equator Network. They’re having their annual meeting in October and I’m giving the plenary lecture in that meeting. So it’s a small world.

Holden: Do you have a high opinion of that group?

John: I think Equator is an excellent initiative. It’s an umbrella effort to cross-fertilize initiatives for reporting of research across different domains. There are many reporting standards, I think reporting standards are flourishing at the moment; many people are interested in creating them. Equator is just putting them under the same umbrella. I think this is very useful.

Holden: Do you have any idea of whether they need more funding?

John: My default answer to that would be “yes.” I think that this is a field that has been so underfunded and where there’s so much important work that is been done. Doug Altman has done such superb work that I cannot believe he would not be able to use some additional funding very creatively.

Holden: How about the Open Knowledge Foundation?

John: The name vaguely rings a bell but I'm not sure that I'm very familiar with them.

Holden: They're not medicine specific, it's a group focused on general knowledge sharing and data sharing in general.

John: I'm not sure that I know them. But again, efforts in the direction of making information open and accessible are definitely worthwhile no matter what. I just don't know exactly what that team might be doing.

Holden: How about Open Science Framework?

John: Same thing, I'm not sure who's involved with this group.

Holden: How about Sage Bionetworks? I think this is a group that's trying to get more data on cancer available publicly.

John: I'm not sure. I haven't heard of that either, but the effort sounds laudable.

Holden: Do you have other views on what the blind spots of the NIH are? You've been calling the area of meta-research, systematic reviews and quality of research promotion underfunded. We also have the sense that it's underfunded. Are there other areas that you think are underfunded, areas that you think the NIH systematically overlooks?

John: That's a tough question to answer on the spot. I would have to think about it very carefully.

Holden: We're always happy to hear from you if you have ideas in the future. I guess those are all my questions for now. Is there anything else that I should add, anything I should have asked, anything you want to know about us?

John: I'm very excited that you're taking such a close look at meta-research issues. I think that this been a very fruitful discussion and I hope that some of it will be useful for you.

Holden: Terrific. We hope to be in touch. If you run across more ideas or more initiatives that you think could use more funding we'd be happy to hear from you. If it's okay with you we may be in touch in the future.

John: Perfect.

Holden: Thanks so much for taking the time to talk with us. I really appreciate it.

John: Have a good day.