Unlocking the True Potential of Clinical Research Data - Aaron Mann

In Teckro's podcast episode "Unlocking the True Potential of Clinical Research Data," Aaron Mann discusses the importance of unleashing the power of clinical research data. Aaron explains why he's optimistic about the potential of data sharing among sponsors to revolutionize trial design and improve patient recruitment – and why he believes patients really want their data to be shared.

HANNAH LIPPITT: Hello and welcome to the Totally Clinical podcast brought to you by Teckro. Totally Clinical is a deep dive into the freshest trends, big-time challenges and most excellent triumphs of clinical trials. I'm Hannah, your host. Join me as I chat with industry experts, trailblazers, thought leaders and, most importantly, the people benefiting from clinical research. So, tune in, settle back and don't touch that dial. It's time to get Totally Clinical.

HANNAH LIPPITT: This week, data scientist Aaron Mann, CEO and co-founder of the Clinical Research Data Sharing Alliance, joins the podcast. Now, Aaron has decades of experience working in data science. Before his current role at CRDSA, Aaron was Roche and Genentech's global program lead for industry collaborations data sharing initiatives.
As part of his role at Roche, Aaron served as the industry lead for DataCelerate, TransCelerate BioPharma's data-sharing platform, and platform lead for the industry's COVID-19 patient-level data-sharing initiative. In this podcast, we discuss how patients can be reassured their data is safe, how the health care industry can improve communication to build trust, and why we're not heading for a dystopian future.

HANNAH LIPPITT: Hello, Aaron and welcome to the podcast. Could you start by explaining a little about your organization and how it came to be?

AARON MANN: Yeah – CRDSA was born out of frustration when talking to colleagues that work in clinical research and use secondary-use data. We were all very good about talking about the problem statement. We're not sharing data as readily as we could be. The data utility wasn't high enough for researchers, so we're really good on the problem statement and the beginning of last year we started talking about, "Well, how do we come together to solve that as an ecosystem?" And one of the things that we realized was that it really had to be a multi-stakeholder voice. So for us, Clinical Research Data Sharing Alliance, when we came together, was bringing together industry, it was bringing together academic institutions, nonprofits, data sharing platforms, all being represented as a voice of the data sharing ecosystem.
In sum, the vision of CRDSA is to dramatically improve the sharing and reuse of clinical research data to accelerate drug discovery. Fundamentally, we want to make it easy to share and easy to use for researchers.

HANNAH LIPPITT: Now let's dive into the key question on most peoples' lips when they consider data and health care: how can we be sure our data is safe?

AARON MANN: I think a lot of it comes down to where is data being shared and how is it being used? In clinical research you tend to see the data held by the sponsor of, for example, clinical trial. That changes when you go to the secondary use of data. Then the question becomes, "Has it been properly protected from a patient privacy standpoint?" I think it has been a learning experience for all organizations how to navigate protecting patient privacy responsibly and ensuring maximum data utility for downstream uses, and kind of emphasizing downstream uses, because we're talking about not the primary use of data – that is very secure behind sets of

---

![CRDSA logo]

firewalls. The question that CRDSA is really involved with is, "What happens after the trial?" How do we reuse those patients' data? Because we know patients want their data to be used to further the science. It's why they join the clinical trial in the first place. So how does that responsibly get shared downstream? The response is both security and an evolving set of privacy laws that are intended to protect individuals.

HANNAH LIPPITT: I think what worries people are the many stories about hackers and data leaks that we hear about in the media. What can companies do to convince patients that giving access to their data is worthwhile?

AARON MANN: A lot of those hacks, data leaks, things that you may have heard of aren't in the clinical research side, because the level of security that tends to live in this particular sector. I think actually we have to convince many companies within the industry that these stories aren't, that there isn't a big danger that we see from hackers as a reason not to share data, because we have heard that from some sponsors. But these systems are highly secure. And if you've protected the patient privacy, in which you've taken out all the things that might identify an individual, then you're not sharing personal data and therefore any kind of a hack can't be used to target individuals. But at the same time, I think they're also not excuses for, frankly, for industry not to share data and share data as openly as possible to further the research.
One of the things that we consistently hear from patients and you see in surveys and I'm sure sites hear: "Will it be used to further the science around whatever therapeutic area?" And so better outcomes become core to why patients participate in the first place.

HANNAH LIPPITT: Yes – people definitely want to feel they are contributing to a good cause. Now, moving on to the subject of data sharing and interoperability – why is progress so slow in this area?

AARON MANN: I think a lot of progress has happened. But if we look at a historical context, data sharing didn't really start till about 2013 at any scale when a number of sponsors came together and created CSDR which is the clinical study data sharing side, one of the first ones, and that started in 2013. And of course, there were all the types of questions: "When something is new, how much should we share? How fast should we share it?" And then 2018, 2019, along comes the GDPR – the General Data Protection Act in the EU – that then imposed a set of responsibilities in terms of protecting patient privacy, but also had sponsors taking a big pause and saying, "OK, how do we interpret that? What does that really mean? What is anonymized? What is that threshold? And how do we react to this piece of major legislation?" And I think in some ways, that's slowed down a process that had been gathering speed.
I think where we are today is that there is a bit of a tension between patient privacy and sharing with high data utility or that open science side. And a lot of sponsors are trying to figure out where is the right place on a continuum there to fall, where you are responsibly protecting patient privacy? So, the individual is protected – but what you're sharing has high data utility. So, what we see, our sponsors are all over the spectrum. So, some sponsors, when they contribute a trial for secondary use, have stripped out, say, most of the demographics and all the adverse events. Of course, a researcher gets that and says, "That's not helpful. That's not going to help me answer my research question." There are a lot of responsible ways to approach patient privacy that still maintain the data utility. And what CSDRA is trying to do is really have the industry come together on what "good" looks like. What's a good standard that protects patients and also maximizes the utility and the research value of the data that they contribute?

HANNAH LIPPITT: I've had quite a lot of guests on the podcast from sites, and they've talked about the problems they've had with communication, especially with sponsors. Now, this is partly due to a lack of understanding of daily pressures on sites. What are your thoughts about how improved data sharing can help to improve this situation?

AARON MANN: Two thoughts on that, because at first flush, sharing secondary use of data sharing is pretty downstream from the day-to-day pressures that sites deal with. But taking a step back, I think there are two answers to that, on the patient side and on the patient recruitment side, that ability to really connect with patients and say, "These data, whether you're in the control or the investigative arm, your data is valuable, and it is going to be used in this clinical trial to further the science, but it is also going to be responsibly reused." Downstream, it's not a given. It is a contribution that they will make of time and data that will keep giving.

So, I think the first thing that this helps with is conversations that sites would have with patients about why participating in a clinical trial is important. The second reason it's important is because we have a real opportunity to change trial design. We are collecting so much data as part of the trial process, and if we don't reuse it – if it sits behind gates and locks and put away in a vault as if it is going to grow in value over time – then we miss the opportunity to design better trials. So, for example, if we have patients today that are approved in a clinical trial – the indication or the compound is approved, it becomes the standard of care – why can't we reuse that trial data in the next trial and reduce the number of patients we have to recruit? Or reduce the number of patients that we have to randomized to the control arm and have more on the investigative product? Because we have all these data that's just traditionally just been sitting there and not reused.

So, I think the second place the data sharing can really have an impact is in making trial design more efficient for sites so they can concentrate on what they do best. And frankly, as an industry and as an ecosystem, we're not recruiting patients when we've got data that can help us minimize the burden on patients and increase the ability to have trials or ramp up more quickly and run more efficiently.

HANNAH LIPPITT: This truly would be groundbreaking. If trial design could change in this way which trials would be most impacted?

AARON MANN: You do hear about it more when you start talking about rare diseases or single arm trials where you're trying to then construct a control. But it is a topic that is of increasing interest to sponsors, but as importantly, it's increasing interest to regulators. So, part of what we're doing is working with regulators like FDA on, "Well, what are the guardrails for creating a supplemental control and reducing the concurrent control population. What does good data even look like? What are good selection practices?" So, these are things that we're actively working with regulators on.

There is a real promise from the trial design side for what shared data can do, and it's up to us as an industry to realize that promise by working together with regulators, multiple, all the stakeholders that are concerned to figure out what's the best way to do it.

HANNAH LIPPITT: So what's holding the regulators back? Is it a question of patient privacy?

AARON MANN: OK so in this case, I would say it's not patient privacy. I think it is that regulators are – they understand the clinical trial randomization process, and so they know what they're looking at. And the idea of using a concurrent control with a hybrid supplement to that control, that's historical data from trials. So post-trial clinical data could be EHR or real-world data or a combination of both. It's a really new concept because that's not exactly a clinical trial. So they'll start asking very good questions like, "Well, how do we know it's randomized?" And there are good answers to that. But that is something that is kind of cutting edge. What we're working on with regulators is, "How does that work? Does it really work then the same way as a fully randomized clinical trial?"

I think the benefits – there is improving patient recruitment – I think there's another big benefit in diversity. So even if you're looking at a big therapeutic area like non-small cell lung cancer in oncology or some neuro indications where they're certainly not rare diseases, but we already know that we as an industry don't recruit a

diverse enough patient population. And what a tragedy it is to have done all the work recruiting on a clinical trial – a particular subgroup or subpopulation you're really interested in – and then not reuse those data again. So here you are trying to recruit the same subpopulation in a new trial and in randomizing part of it to the control. When wouldn't it be better to have them all on the investigative arm? So, there is a lot of reasons to do this, but I think you're seeing a lot of caution on the regulatory side that it gets done in a way that is consistent with how we're used to doing. The gold standard RCT.

HANNAH LIPPITT: Now there's a lot of dystopian thinking – especially around AI, the future and concerns about privacy. Who will be able to control our data... what are your thoughts on this?

AARON MANN: They're all legitimate concerns that we, again, as patients, as consumers, individually, we do need to be concerned about. I think the future is going to be less dystopian than some people say it is because there's a lot of regulatory pressure, particularly out of the EU coming around making sure that there isn't a dystopian future. So, for example, when we look at AI, there's a lot of work being done on how AI functions, how it shouldn't be a black box that just comes out with an answer and you don't know how it got there. So, I think there's a lot being done around that.
The most interesting part of your question is controlling the data, particularly in health care. So, I think there's a real debate right now over who owns and controls individual patient data, your EHR records, and it's not consistent across geographies. But the way that we look at it at CRDSA is that companies – commercial enterprises – are stewards of the data and that those data, for example, in a clinical trial are donated by patients – so companies don't own that data, they are stewards of it. And I think the same applies to EHR data. So, when you go in to see your doctor, it is the same principle that that EHR record is contributed data – if it's going to be used for secondary use, it should protect patient privacy – but also that's a societal good. It's not necessarily owned by the company that happened to provide the software that your physician used when you came in for a visit. Now, that's a hot ethical debate right now. But in terms of controlling data, I think we all need to be very vocal and who actually owns our own data? And my personal opinion on that answer is, "We as patients, as consumers, we own that data."

HANNAH LIPPITT: And staying on the topic of the future – even if it is not going to be as dystopian as some people think – what are you most optimistic about?

AARON MANN: I'm excited about the possibilities that are now unfolding in terms of powering research with amounts of data that we just couldn't navigate before. We either didn't have it or we didn't have the tools to understand it. So, I think we're at this point where everything can sometimes seem a little scary and we've talked about it, you know, previously, it's the previous questions. On the other side, we've also got the tools to understand big data sets. We've got the methods to protect patient or individual privacy responsibly and maintain high data utility for researchers and so, as we navigate and understand how these come together, we're able to unlock insights and speed drug discovery in a way that just wasn't accessible to us even 10 years ago. I think that, to me, is what makes it exciting to get up in the morning and it makes it exciting to come to work every day and it is an exciting time to be working on data sharing and looking at how we unfold the real power of what it is that we're collecting and the contribution that patients are making.

HANNAH LIPPITT: And that's your dose of Totally Clinical. For all the listeners out there, you can follow Teckro on Twitter – the handle is @TeckroOfficial – LinkedIn and Facebook and subscribe to our YouTube channel. And, of course, download the Totally Clinical podcast on Apple, Spotify and Google. See you on your next visit and remember to bring your friends. Thanks for listening! Goodbye!

CRDSA