Tokyo Workshop on Genomic Medi...eutics, and Health 20240411-09

Thu, Apr 11, 2024 4:48PM • 39:31

**SUMMARY KEYWORDS**

biobank, uk biobank, data, people, smaller, privacy, project, facilitate, genomic, standards, policy, analysis, cohort, terms, important, uk, participants, world, process, question

00:00

And there are other final rounds obviously. And then in the context of data analysis, conversation, experience and trying to combine data from events. So, with that again, no matter its format was new to this discussion about this. One of the approach

00:38

here for the age and some of the general principles that are really support, harmonization of data sets and using standards that does not facilitate some aspects of it. Sharing information I noticed the pipeline in the UK Biobank results, similar pipeline biobank all of us so, you know, that's, that's sort of my introduction, but there's not a whole lot more than that. And then after this, I was just thinking that if people want to comment on their experience, in terms of trying to replicate the analysis in some of your data sets, or some of the files come from some of their bio banks, again for facilitating this you know, that was a plan for today. But again, tomorrow, there'll be one discussion as well and presentations from other environments. So, I mean, maybe I'll turn it over first again to talk a little bit about again, as he did before, differentiation, some of

02:14

y'all can help me because he's also Arkansas. So well, g4 GH is an extremely large organization. It's one of the largest organization in genomics at the international level. It's an organization that is completely dedicated to basically facilitating data sharing and also the metadata and when we talk about that, but it's all sorts of omics genomics, yes, but other types of omics data. The organization is an international organization and it is it has offices in the United States. In Canada and in the UK, but really just members from all over the world. And the way that the Global Alliance works is true workstream and driver projects. And so the workstreams are just specialized way of global iamcp, if you will, meaning that there's different area where policies and tools are needed to facilitate data sharing. Some of those area are very technical and meets all kinds of bioinformatics standards and goals are areas that do most well. That other areas are more policy oriented. For example, there is a work stream that is interested in security questions adding to do with the data and protocols in the work stream which I direct is really interested in developing policies to make this possible. So for example, we have developed a privacy policy we have developed recently a very interesting policy on diversity. Because as you know, everybody is saying these days well, you have to have diverse data set in your analysis. But But what does that exactly mean to our diverse data set out you make sure you have diverse data set the new analysis, so we have a policy to help you determine that. That was recently, for example, and the driver project? Well, their role is that they're actually genomic biobanks like, for example, that all of us project is one of the drivers project. You'll be at the MP share project, which was an epigenetic project. That was also a driver project and what they do is they can help develop the various tools and product of the Global Alliance, some of which I've mentioned, but they're also responsible for trying some of those tools and protocols as well. So that really they can provide us feedback. We have feedback from real genomic project about well, is this working is this what we need? Because the idea is we really want to be helping the genomic community so we really need it from people on the ground is that the tool that I need to facilitate my analysis? Is this the type of protocol that that is needed as this type of app that is needed to facilitate girl life? And that's pretty much

05:42

it's in one of the first project GH was the Beacon project was meant to announce. The concept there makes a whole lot of sense in the context of rare disease. You know, if you want to be able to link up biobanks identify patients that have specific rare variants. You know, that's an example of something you know, is there a way to facilitate is to identify this this manner and then observe in any of the existing way

06:14

and now they're expanding it so at first it was simply a query to know basically, do you have this kind of variant in your own database? Is it present yes or not? Now you can do it you're trying. Again, that's where all the SQL for example can be helpful. They're trying to implement a version that will allow to share more information. So not just a simple yes, no, but I don't know what kind of phenotype what kind, what additional variant is there that could be relevant. So you might basically require more information. And there might be ways also to do this to multiple site as well, you probably would know more about that. But the question is, is it possible to do that with the current privacy standards and if we start implementing this in ostinatos, because the idea is to also have these kinds of tools available in as many Ospital that are doing genomic testing in addition to research academic research lab. So what do they need to know to operate this properly and not provide more information that they're not allowed that they shouldn't be providing because of privacy restrictions?

07:44

So we need to turn it over to you to give us perspective. Okay.

07:51

I have some clients or private systems, I see how people seem to have the issue in privacy is gone, what else is going to change? Right? Initially, we discuss the identification people discover issue differently about you can already have an answer. To me, somehow, I think being you give up on yourself not be a violation of privacy, because subtraction, so then you're going to find another type of privacy people will have to specify and maybe as if the privacy issue can be more flexible than me changing. Changing the topic. So I will so I think that because it is just by Ranking, always the main since the first time in the world, so likely to have such an expected progress. For example, outside Japan is actually what we should do in the right day. We can learn from what you give us to us. Three, once or has provided a policy is really good guidance. Because the privacy is really this kind of LCS changing. We need to be more flexible with the mishap with the privacy's and dignity, for example. For example, the travel since it hasn't discussing more or less the same charging issue several massive, massive, massive I found this the when the technology changes, they didn't see any of them understand. So I see. They faced such new issues every time so right we know how you guys seem or when you first met the new power leader, but tie this into this kind of policy you have on the world and deciding that, of course, outside for all Kevin how you're going to see

10:24

I don't have a complete we have an aspects of you connected in your work that we do that we also have specialist group so we have a specialist group to look at emerging data wanting to Natick data on the health record and it can help provide some insight into that and steroids. They provide expert analysis and perspectives. We obviously have our own partners as well, and a legal team who are very, very hot on privacy and data privacy, and how that can be reasonably interpreted. So I think there's always a balance between protect protecting complete privacy of the participants and ensuring that they want to stay part of the study and that they feel so secure and supported by it. But also making sure that the data is useful. Right and that you can make sure that the data is accessible and available to people and that it can be used well results.

11:34

It gets really important for different why that sound allowed to engage in with governments and health services around these things as well because, of course, it's not just us to decide the policy it's sometimes the government or people who for example, liquidator with our site NHS, giving us the constraints that we have to work with. So that's a particular view at the moment with the chancellor trying to develop secure data environments. So that kind of maintain really good relationships with these bodies, so that scientists can be part of the conversation and not just have those imposed upon us is really, really important. And that means that we have to show that we are interested in what our participants think, what the public thinks. And it also means you have to show that we're willing to engage and to follow and we could care or we can often

12:40

forget also if the if you're thinking about research as as an international activity. So right now the UK Biobank as found the sweet spot where they have something that works very well for their population that meets their legal requirements. But that can be very different in another country. And the solution of UK Biobank can work. In the UK, it might work in other countries, but not necessarily. And as the worker chief was used to try and make things more interoperable, but without you know being very rigid in their approach, the basically will make policies that we felt were that will facilitate the interoperability of various national context will basically be able to work in different country and exchange data, but he will not try to have identical laws and policies and biobanks all over the world because that's that's something that will certainly not happen overnight. And I think the perfect example of that was for me that COVID And he tried to do something internationally globally, to do research to advance on COVID Some countries did very well in terms of and the UK was one of those that managed to get the data out there extremely rapidly. Canada you can attest to it was not that great at it. And other country were somewhere in the middle and why because of all kinds of things you know laws, political context, cultural context, in some country, people are very sensitive to what they want to share about themselves. In other country, people tend to trust researchers very much and then it's much easier to convince them to do that.

14:39

You know, one of the things I just said it was very interesting also is the need there differences in different countries. Taiwan opt out to raise the discussion about

14:50

in the six month delay to access the data was something

14:56

and then within the audience of people who had experience trying to access data from from different cohorts and what's been their experience in trying to replicate and analysis in different data sets. You know, there's experts in the back when we've done some of that so

15:26

maybe we can

15:27

think of obviously UK Biobank but also all of us which is brand new, and for which now, as a Canadian, I can access which we didn't, we couldn't before. So obviously it's tempting to compare the processes as not to say that there's a better one and a worse one. Certainly, there's one that were more useful to the UK Biobank is feels more like natural because that's how we've been accessing data before. All of us were in the room that have been just you're eligible now is very different. The way you access the all of us data. The training is much longer but also the process of accessing goes through. It's a very individual process with every student ever posted up we'll have to go through the training data not depend on their API. So so it's a it's a learning process. And I know in my lab, we felt like there was something wrong because it seems like the students do not depend on me. Although I'm responsible. I have nothing to do in the way the tests answer the question, need the program officers and so certainly, this is evolving. We shall see whether the all of us process or is the one that's going to win in the long run. I certainly don't know but it certainly feels weird. So that's that's that's one thing. And then we're also tempted to compare of run analysis with the all of us data with the DD Nexus where we just heard it and also has a different feeling. So for instance, in my lab, people are very good with demeaning suits now. We're very bad with all of us have been struggling quite a bit. We'll get there but it's it, it's their feeling. And then, you know, I'll just vote by saying that there are also issues are quite different in the US than in the UK. stigmatization of minorities is actually I don't know what it is in the UK, but when you go through this training for all of us, you know, that is a really big thing. So that's also something to pay attention to get different things.

17:38

Like and you'll be speaking tomorrow about the Kinect biobank or with one of the plans are sharing there.

17:47

So I'll talk about this tomorrow, but back, basically Canada has a quarter 40,000 PARTICIPANT it's a small proportion, in terms of comparison to UK Biobank for all of us, but even besides Quebec, it's it's pretty large. It is it's pretty open. Biobank follows the model of the UK biobank that is you can still request the data. It's pretty fast. We're far from six months. You still need it, but you do need RFPs. And then you can download data. Maybe ask for more or there's been so many requests that we'll have to think about it but we're still working under the framework of shutting down allowing people to download data we generate new IDs for every Mr as it used to be with the UK Biobank. So we're still doing this. I'm sure I'll see what's the mall with the mall is going to be but it is very easy, right? Now to download.

18:54

This I read the citizen. I'm not sure how all of us is. I think I massively strongly satisfy the constitution. So I think I couldn't imagine circumcision, because I'm not really satisfied but but you need a different way of assessing. And so I think we are strongly satisfied. We are now starting to recognize that the marriage of the separated data. So we have one is everybody together why not for example, since more than circles all over the world. So people are now satisfied understanding the dots by recognizing system right because you guys are the top adviser, how the next tip I use this and how to combine how we interact the Allied sat resource together the barrier of yes nothing is divided. So anyway, or vocalization when was this company because this is also another big idea if we can Yes, closely analysis

20:17

is really promoting some standards not just with the renovators organized but also in a way that is applied for operators systems. You can run workflows, import workflows from one platform to the next in similar ways. That's where the standards ultimately will hopefully help harmonize and interpret some of these platforms. So I think at this stage, they might feel quite different with the lender. But I think that that's one of the objectives is to try to make them more similar and a bit more accepted available. So live in separate platform, but hopefully some of these repeated analysis will be something

21:03

comes up. Then as we kind of combined more and more wyrex is the potential overlap to individuals. So we see this for a cardiogram, we have many veterans program, and all of us alongside each other, and just kind of work out that's gonna get worse and worse as we get better and better at developing biomax. And I don't know if in particular thought process people thought someone on the second day, that's a problem.

21:38

That's an issue and in some of the issues we talked about the presentation as well as bad chip bags or how to analyze and combine the process in slightly different ways. You know, some of these issues will be amplified once you're away from this as well. So, I mean, different days is a lot of work to do in some ways. It's so instead of focusing on different aspects of it, versus some aspects of the methods. Absolutely thinking about race, sample relatedness, federated search, lacks some discussion I feel

22:25

guilty as the permissions within technology because the privacy missions, the ethics behind which those days were affected in the first place. between a rock and a hard place,

22:42

sometimes, you know, you can amend the protocol and of course, if the data is already there, and you can change what you might the possibility would be to call recontact. Possibly, if that that is feasible. I mean, I agree there is no perfect solution. There are there are things that that you can do but you don't ask the answer for that one. I'll ever one thing I wanted to add, I think we forgot a really important thing about the Global Alliance is that anybody can join. So I think it would be great to have more people from Japan involved in the Global Alliance for example. Because what that means you can actually participate in building those standards. And strategically, it's to your advantage. Because of course, if you see something that from your own perspective, would be interesting to add a feature for example, that would make your life easier when your analysis that you're analyzing across the different database and biobank, then you get to propose that and basically, there are workgroups, as I mentioned, that are interested in all types of tools that are useful and it's quite democratic in a way as a process. So So basically, I really encourage some of you to try and join Global Alliance that is of interest.

24:20

actually wanted to ask a question. So I work with a smaller biobank. And oftentimes what we're being told is, you know, there's biobank and there's biobank and and now there's a bus. What's the relevance of these smaller biobank? And I can think of many, no, I don't think GWAS is the answer, but I can think of many other reasons why we would want to have smaller bio. So I would be curious to hear from the panel and think about Yanni mentioned diversity and then with colleagues, biobank, and then you also have smaller national biomed interval and Cambridge, smaller but, you know, how do these others reality in your countries in terms of just in terms of funding, you know, how do you spin your smaller biobank smaller cohort, despite the fact that there's a huge, very successful project going on?

25:20

Yes, we saw not what the biobank is where this you can actually expose a bunch of scissor paper. There's so many small and I myself, as somebody likes to collaborate with Steve guy, because these guys are always a carrier, the messenger cobalt, they're very happy to help but I know not all properties initially where I was sometimes cannot get enough money or cannot get enough staff. So actually, I love spending big power is important but helping the lady maybe not be was small by Van because they is by what are called the strong independent single principle. That is actually very important, but when such non big corporates stops as its strong blocks for some years for science, that's just not clear. So I like to help them. I like to have them write nice paper, but there's many limitations are changin have you guys or is it only only UK, UK or?

26:37

There's a lot of there's a lot of biobanks. I think there's some What are smaller and specific. So we have quite a few birth cohort, type biobanks and longitudinal studies. So they look at everyone born within a specific year. They enroll as many babies on in one year group as possible. Some like million women study with just women being really really valuable Britain's doctors study with was one of the kind of births of these non maternal prospective cohorts and obviously, because all of us which have slightly different leanings as well, so all of us is often sorry, our future health and so our future health is currently about the same size as UK Biobank but they're in the middle of recruitment. Now I'm going to get to as many as I will say.

27:44

I think most of these blocks are slightly different objectives or they have their representative versus different. So I mean, the blood donors etc tend to be on the younger side. We know all these bytes are gonna take time to come to fruition in terms of developing disease that we're interested in. So as we like to sign up calling something wire bytes, I think there's probably a word that's overused. terms of its of its definition. The Women's Studies prompt is fantastic. And it will answer questions better than up in some respects, but they are specific questions. I think that's that's probably how we have to spend making sure that we keep on abide by back to the vertical is relevant. So what why were they why are they there? But they were generated for a reason. It wasn't just that somebody got more that day. So we just need to reflect back on on that why question, and I really capitalize on that. Some of them are very much more deeply phenotype others. Some of them have tissue for example, or been able to get back to SARS, somebody like besides just looking so I think what most things you kind of have to work out what your USP is. Sell out to the funders like inevitably there's a fixed amount of money, I think

29:15

one of the running red three smokers and the size of the algorithm, the database is always trying to offer you a fixed amount of the budget. And what we are doing and what we want to do is to increase the return rate. So we have some set of the results and materials over time. And I was a bit surprised when I read these addition that they are a very small fraction of people who from whom you get buy materials for for eyes. You it's maybe one or 210 or 15,000 people that have been followed four times. So this is more or less the size of our cohort where we have sort of chronological information and by areas. This is quite important if the study in genomic medicine goes on to be able to be the voice of caramel. Okay, density now and another thing is, you know, the small cohort is very quick in making different projects and for example, we are we are working on the project which otherwise, don't think about for example, we have cooperation with a sample the World Cup and they are very much interested in world of aging of skin by taking images of sites to evaluate the scheme and then try to reach out to genome and omics for the development of snips. And we have a group of economists working to reduce social capital formation because we know very well but there is our mental activity is very best to do this our physical health and there is no ways from happiness technology of meta science to evaluate the fairness for example, regarding on the society and social activity and cultural concepts and so on. So we meet with social science people and if you do that, very scared, it is really difficult. So, we can do sort of a feasibility study in smaller and then once we have set up the board then we can go for that. So the people writing the code can take advantage of this, this is I think this is the quite important point of the smartphone.

33:03

Yeah, I think it's also even with the logic of bots is that you need to try things in small groups. So I made up for participants and like for visits at the moment nice quite small number because it's been piloted. up slowly. So it should be up to 60,000 events. But things like the COVID Please control everything. We knew that was only the capacity to do that well with the painful requirements involved in a small number that's about 3000 participants. It's often advisable to have that really deep thing on small number tend to just not have any data or are very likely to do it at scale. With some projects as well. We start small and pile it up. So normally around 5000 It's a real good way to see how you can do things and how you can do things at scale. Yeah, you always learn a lot when you go up to high numbers. High throughputs learning that in smaller cohort and small subsections.

34:14

Just wanted to add also that your if you're interested in working with indigenous population groups, in many country and that is certainly the case in Canada it is impossible to think about the big court at this point in time. We have project angiomas one of these, I would say it's larger than what we're going to have for for indigenous population to start which app collected its population or biobank. So by nature of some of the people that are participating in this project, far from indigenous population, well, it's already hitting roadblocks, because there's a lot of pushback saying well, these people should have been included using other mechanism and therefore right now it's at the standstill, and people can't access that data. And so we have a separate project now, which is called the silent Genome Project, which is only about collecting data from First Nations. But the problem is you don't have just one First Nation, we've got hundreds. And so and they all think differently about what they would like to see and the complexity of recruiting these participant there's also a question of trust, where at first you'll have to convince and accept that there's not gonna be that many that are probably going to participate. But if they see that the project is serious, then you'll see more the cord will grow and a little bit as you're mentioning, you're going to be testing things to see if that's working. And because also you're integrating member of these groups to the administration of the biobank. It also is a completely different setup, which takes a much longer time to grow, but that's the only way we're going to be able to do to gnomic research on these populations.

36:19

Really, maybe just a final point of all of this, I tried to imagine like an intranet for biobank data really captured all of these biobank into one system. Her respect, privacy and consent. associated with the size of Mobius came in from the UK Biobank with millions and millions of samples. And I think with the kinds of methods that were legal talk about, you know, machine learning, you know, if this was available to these as methods, you know, there's so much discoveries that would be possible. So P for GH in some ways, trying to build an Internet of biobank data in a way that I mean, that's the dream, whether and when you'll achieve that, I think we can learn a lot from these existing large bio banks, with them the idea of trying to find ways to connect across all of these to identify and I was giving the example of their bare patient variants, but to really explore that data in a way that's safe. Radical that that's the dream long term to reduce. So in that sense, encouraging all the bio banks to migrate to these kinds of standards to really allow to these kinds of discovery and interconnection that will really allow the creation of a gigantic world by you know what

37:55

I mean, so, I know that in terms of Canada, we're also setting up a new project and I think the implementation of the standard it takes time. It started 10 or 15 years ago, but you know, there's growing adoption, I think of the standards, and I think there's definitely discussion for projects to have to use those things. I think this is now coming to try to set them up. But I think I would encourage big and small biobank. Look at that. Because just going back to the introduction with the UK Biobank, I mean, ultimately, the goal is to have this data and use it as much as possible in the right way. And I think direct use is one way of being able facilitating these integrations with other datasets. Encourage data access. Anyone, so I think we've talked we've had a long day already. here and again, tomorrow, you'll have additional presentation. Interesting, that maybe a final round of UK Biobank colleagues that really organized quite a day for us.

39:19

Getting started.

39:27

Good evening.