

Adrian Hernandez:

Hey, this is Adrian Hernandez, and welcome to the NIH Collaboratory Grand Rounds Podcast. We're here to give you some extra time with our speaker and ask some of the tough and interesting questions you want to hear most. If you haven't already, we hope you'll watch the full Grand Rounds webinar recording to learn more. All of our Grand Rounds content can be found at rethinkingclinicaltrials.org. Thanks for joining.

Hi, I'm Adrian Hernandez. I'm a moderator for Collaboratory Grand Rounds, and today we're here with Joshua Denny, who will be reflecting on the All of Us Research Program: Improving Health Through Diverse Technology, Huge Cohorts, and Precision Medicine. Josh just recently gave a great Grand Rounds, highlighting what All of Us has been doing and what they're planning next. So Josh, I really thank you for joining us.

Joshua Denny:

Thank you so much, Adrian. It is a pleasure to be here.

Adrian Hernandez:

Well, first let me just start off with, for some people, they may not necessarily know how All of Us got to now. So what was the grand vision for this? Simply just enroll a million people and say, "We're doing something good for science." What's really the vision here?

Joshua Denny:

The vision really began, for me, late in 2014 and through 2015, thinking about the opportunities we have, where the gaps might be to improve health. And one of those things we identified is really the lack of diversity in science in general and especially in genomic studies. And so, this was an opportunity that was really foundational around All of Us. We also thought things like electronic health records had been proving their value as a way to capture a lot of detailed health information across a lot of conditions and drug responses, as well as really wanting to combine opportunities to capture and have participants contribute, donate things like Fitbit data and other kind of activity monitor data to correspond with you and answer questions that you don't get from an electronic health record.

So we thought we could really be holistic across all health conditions with a particular focus on underrepresented populations. And that was really the beginning of the idea in 2015. And then we also decided that we wanted to eventually approach all age ranges. Right now we're 18 and over, but we are working hard on our pediatric cohort.

Adrian Hernandez:

That's great. Now, you touched on this a little bit, but a lot of times people think about electronic health record data as the only digital data source. But All of Us is really tackling a wide range of data. Can you briefly describe what that is and what are the domains?

Joshua Denny:

Definitely. Well, once a participant signs up and consents, the first thing they usually do is fill out a set of surveys. And those surveys cover things from basic demographic information, which gets more details than what you'd get in a typical electronic health record and allows them to specify information around sexual orientation and gender identity, income, education. And then goes to more typical things like personal and family medical history, and a number of other contents for that. Things like healthcare

access, and newer surveys that we're launching this year will address things like mental health and wellbeing.

And so these kinds of questionnaires are some of the first types of information people often provide. They include authorization to access their electronic health record. That's from a number of our primary care providers across the country. We have about 60 or so medical centers that are actively contributing electronic health record data into the dataset, which includes federally qualified health centers and large academic medical centers, and other medical centers that are part of the consortium.

Participants can also link in through patient portals, electronic health information as well. So really that allows us to address almost across the country that kind of information. And then people can link in Apple HealthKit data, they can link in their Fitbit data. Those kinds of things have been shared by a number of individuals. We're starting to give out pilot populations Fitbits in that process.

And then we're generating data from biospecimens. Right now the biggest set of that data is to whole genome sequences. We've now released to researchers, about 245,000 whole genome sequences. We have also released our first data set of long read sequences, about a thousand individuals, which actually all identify as African ancestry.

Adrian Hernandez:

Well, Josh, that's really incredible to hear the amount of data that's been already accumulating here. Can you give us some sense of how is this the same or different compared to previous cohorts that the NIH has funded? So Framingham is something that I know about as a cardiologist, then there's the UK Biobank. Just give us a sense of what the differences are compared to those two.

Joshua Denny:

When I talk about this, I talk about Framingham a lot. It's been such a transformational program of showing the power of observational cohorts over time. As a primary care doc, I used to use Framingham risk scores all the time in clinical care, and I think it's a model for what you can do over time to identify intervenable risk factors and make a difference in disease. And a lot of those cohorts have had similar sorts of findings for specific diseases.

Our goal, as with other such large programs, you mentioned like the UK Biobank, is to really be able to set up the foundation for studying all sorts of diseases and all health statuses. And so that foundation really lets us go across lots of different conditions. And so that's a key difference in Framingham, and it's also one of scale. Notably, Framingham was not very diverse. And so our program is really seeking to include underrepresented populations and close some of that gap we see in genetic studies or other kind of clinical studies available. And then make those data available to as large a researcher audience as we possibly can.

And in that framework UK Biobank's really been an incredible example of how to facilitate open science. The UK Biobank recruited their participants from 2006 to 2010, and then first opened up their data resource in 2012. And it's just become an incredible resource for epidemiological and genetic studies across a whole host of diseases and other conditions as well. People look at all sorts of survey answers, and mental, and societal conditions that all affect health as their key outcome. So UK Biobank has been tremendously powerful.

Some differences between us and UK Biobank is we currently enroll anyone over 18, whereas the UK Biobank focused on older individuals. The UK Biobank also is less diverse. It is a majority white population through the UK. And so our program, again, is really focused on a large definition of diversity.

It's not just race and ethnicity, that's part of it, but it's also sexual gender minorities, it's rural populations, it's low education attainment. It's a number of underserved conditions.

And the most recent one we added to that was disability. That's obviously an important part of the health experience of so many in this country. And so we have added that into our mix. And just like many other things, we have ongoing connection with our participants. We could send them a new survey to ask them questions about that and incorporate that in the metric.

Adrian Hernandez:

That's really incredible to hear really the expansive nature of this and just highlighting when you're talking about underrepresented populations, that includes areas that we may have so-called "forgotten" scientifically. Rural health is an important area as a good example of that, among many of the ones that you said. Every major program like this, and this is huge, has pleasant surprises and, say, challenges that need to be addressed. So what have been some of the pleasant surprises along the way and then just what are some challenges that you all are addressing for the coming years?

Joshua Denny:

Easiest thing to identify would be some of the challenges. And the biggest that hit us was COVID. Like many programs, we were completely reliant on in-person interaction for the vast majority of our recruitment, and certainly collection of blood specimens. So what we've done since is we've developed capabilities of reaching people anywhere in the United States through things like mobile saliva kits that we can send to your house and leveraging other technologies better and relationships like blood banks and our mobile engagement units, which travel the country and allow people to enroll in any place. And then things like, of course, adaptations for social distancing and things like that for a while.

In terms of pleasant surprises, when this all started recruiting at scale a highly diverse population and reaching people was not something that we really knew how to do. And the engagement with just such a diverse set of community partners and engagement partners, both within our healthcare provider organizations and operating independently and collectively, I think has been a real success. Even some simple things like the technical pieces of it, we can call that simple, but the scale of data to work with, the harmonization of electronic health records across so many different providers, so many different EHR vendors into a common data model. Those were all things that really were pretty unsolved problems when we started and we continue to work out the processes for how we manage things at that scale, put them together and make them usable.

I think we've achieved a lot of success with that. It's a journey, and we continue to get better, but I've certainly been pleased with how it's worked to pull those things together to liberate those kinds of data and put them into a common format that people can relatively easily use compared to other processes that have been engaged in the past.

Adrian Hernandez:

This is a very different project. I mean it's not investigator centered, but really community centered and getting research done across a number of different investigators.

Tell us a little bit about the scale of that platform, how many users and investigators you currently have. What's the vision for expanding that? It seems like it's a range of investigators too, established as well as young investigators.

Joshua Denny:

Yes. That's really the vision is to have this platform be as open a science platform as we can make it. And we really want to democratize access to different classes of users. At Vanderbilt, when I was there, I had built a large compute infrastructure. It had to be secured to handle certain kinds of data. And those kinds of requirements to do this kind of analysis limits the researcher audience. Our goal here was to make a cloud-based platform that would allow really any researcher that is approved to be able to access the kind of tools they need to do the analyses that could use incredibly large data sets. So I think we've been successful at doing a lot of those components. Right now we have about 6,000 researchers using the system. We are picking up usually over a 100 new researchers a week. We have north of 500 different institutions that use this.

And one of the things we've really been promoting is engagement with under-resourced universities and research centers as well, HBCUs, as well as the historically Black schools of medicine, other minority serving institutions, your non R1 institutions and even have engagement effort that's working and helping educate high school teachers on use of the platform. We've had a couple of high school students doing research studies using the researcher workbench, which I'm really excited about as we push that real-life STEM access to larger populations. And I think those kinds of things are needed to stimulate interest in science in this country as well.

So most of your researchers are traditional academic researchers, but I think as we engage broader populations, more diverse populations broadly thought about, that also is a win for our country.

Adrian Hernandez:

Yeah, no kidding. Hearing that and all the different programs that happen over the summer where they're really stimulating science and young investigators in high school, be outstanding to see how this evolves.

You've also been committed to participants in return of results and breaking new ground there. It's quite easy to say something in a slideshow. It's quite different to actually return results individually to participants. What's the framework for that and how are you all doing that?

Joshua Denny:

Our framework for return of results really has focused first on genetic return results as a priority. Of course, we let you see things like your survey answers. We've returned, when we did a COVID serology study early on in COVID, we returned those antibody test results. But, you know, what participants told us they wanted back most in our early work before we recruited any participants was genomic health related results in two categories, health disease related genomics and pharmacogenomics. So we have those two components as part of the program that we're returning now. And we started that in December of 2022. And to date, we have notified over 40,000 people that those results are available and that they have gone in and selected that they would like them. And once that happens, then we generate those results.

One of the things we talked to about to participants is they can choose whether or not they want these kinds of results back, and we return results to the participant. That's something else they told us. They wanted to be the arbitrator of that data going back to their provider. And we try to make it easy. There's a provider report that they can print out. There's genetic counselors available to help talk through the results. And those are all really important parts of this to understand what's there.

We also do things like genetic ancestry and non-health related traits, like what kind of earwax you have and cilantro, and we started that a little earlier because it's a lot easier to get that out there. But the health related has really been a focus for us, and I think has the potential for real value to participants.

Adrian Hernandez:

Very cool. Well, last question. Over the coming years, what are you most excited about for All of Us?

Joshua Denny:

I'm right now looking forward, I'm really excited about launching a pediatric cohort and I think that that's a real opportunity for us to think about development, pediatric disease, and the projection of disease over time from pediatric and adulthood, especially when you consider some of the real significant health challenges facing children these days. And I think that that will allow us to look at some of the similarities and differences. So that's one thing.

Growing the genomic data and the power that will have in diverse populations is something that's already starting to show fruit and I think will be a really powerful tool as we go forward. I'm really excited to see what kinds of discoveries come from that, from diverse disease-based genomics in different populations, pharmacogenomics, seeing if some of these things hold up across different populations that we think they do based on allele carriage rates of what we think are functional alleles, but what will that mean?

What will it mean to look at loss of function genetics? People that have loss of function variance in genes that we may not really understand what they do, and what will the power be of that to define what a disease is, maybe find a treatment that will be helpful for a new condition. So I think there's a lot of discovery science will be there, new cohorts and new types of discovery that I'm really looking forward to.

Adrian Hernandez:

Well, it's exciting times in biomedical science where so many things are converging, and so it does seem like it's a perfect storm for All of Us and the benefits that that will generate for the U.S. and beyond.

Well, Josh, hey, thanks for spending time with us on this podcast. Really appreciate your insights and sharing what you all are doing for All of Us.

Please join us for our next podcast as we continue to highlight fascinating and informative changes in the research world.

Joshua Denny:

Thank you very much, Adrian.

Adrian Hernandez:

Thanks for joining today's NIH Collaboratory Grand Rounds Podcast. Let us know what you think by rating this interview on our website, and we hope to see you again on our next Grand Rounds, Fridays at 1:00 PM Eastern Time.