9 Sept 2016 - Personalised Medicine, The Promise the hype and the pitfalls –
Anna Middleton full talk, 18'31"

So my name's Anna Middleton and I'm head of Society and Ethics Research in Cambridge at the Wellcome Genome Campus where the Sanger Institute is.

The Sanger Institute, the home of Sanger sequencing where there are thousands, literally, *thousands* of scientists and bioinformaticians. And I'm one of a very small group of social scientists. And when Alastair said earlier: "patients need to be heard," well, social scientists listen.

A lot of what we do as social scientist, particularly on the campus, is ask patients for their views, their values, their beliefs, their *thoughts* about genetics and genomics and how they want the technology to be used.

I'm also a genetic counsellor so I've come from clinic, I work for 10 years in the NHS working *with* families with genetic conditions, and I feel as if that experience guides a lot of the way that I approach research.

So *I'm* going to give an overview of *why* data sharing, particularly in the genomics world, is necessary.

What are the perceived *risks* from donating your data? Cause *always* there's a person behind the data that is eventually shared.

And then I'm going to end with some preliminary findings from our attitude research that we've done on public and patient attitudes towards genomic data sharing.

So why is genomic data sharing necessary?

Well, *really*, it's because, in order to fulfill the potential of genomics, you have to have a lot of data to look at. So that means a lot of people have to donate their data.

And if you're going to *look* at the DNA of an individual person, in order to be able to interpret it properly, you have to be able to compare it to the DNA from thousands, potentially *hundreds* of thousands, even *millions* of other people.

And those people need to be from different ethnicities. They need to be a *whole mixture* of people, some of whom will have the same clinical condition as, perhaps, your patient does and some of whom may be perfectly healthy and have no disease.

But those comparisons need to be made if you're going to interpret what a "variant" means in an individual.

So there's *a lot* of people that *need to* donate data, to give permission for their DNA and their medical records to be accessed and to be used by others.

And when data is donated, often it's done in a de-identified form so names, addresses, dates of birth are removed and you just have the data there to be accessed from databases.

Those databases might be publically available or those databases might be authorized and you need to get a key to access them.

So what are genomics databases being used *for?* Well *many* different things. They're being accessed every second around the world.

So it could be used within clinic, within clinical healthcare settings, so medical doctors, bioinformaticians might be wanting to know "Is *my* patient's genome result the same as someone else's? Does this condition look like it's *caused* by the variant that we found in this person's sample?"

It could be used in non-profit research, what's the prevalence of a particular genetic change across the UK, for example?

It could be used in *for* profit research, for the development of meds and *is* this particular medicine working in people with a particular genetic change?

Lots of different *types* of databases. Lots of different types of data.

How would *you* describe all of that to the public, in order to be able to *ask questions* of them about it?

And given that we know that 90% or so of the public have never even *heard* of the word "genomics" before, have *no concept* of the spread of genomics through healthcare at the moment. No idea that genomic technology is now having mainstreamed on a *massive* scale and that it's *not only* in the specialists clinical genetics services, it's now *in* dermatology, ENT, pediatrics, obstetrics, oncology. How would you explain all of that to somebody who has no concept, really, of this? So we've had a go.

And a lot of what we do is to *try* and find what hooks, what metaphors, what memes, what phrases do people find resonate for them within this area. And I've been exploring the lot of how to do this using film.

So I'm going to just share with you one of the little films that we've created that sit within a survey that we use to gather attitudes, and this is a public survey, and this describes what genomic data sharing is without using the word "genomics," we use the word "DNA" instead – we *know* that the public have heard of the word "DNA," and without talking about "databases," but using the metaphor of *shelves* in a playroom, of a child's playroom, representing databases. So see what you think.

DATA ACCESS BY OTHERS FILM PLAYS

OK so we have a survey with a series of films that are like that and they give you *enough* information to pique your interest and *hopefully* engage you enough to stay with the survey and answer your questions. Hopefully they don't bias your views, they just give you *enough* to orientate you with the subject and then share your views.

And this is a piece of work for Global Alliance for Genomics and Health, that Mark mentioned earlier. So that's a *great* organization that has incredible people from around the world *trying* to make data sharing possible and *enable* data sharing to happen.

And I realized as a social scientist that there was a *voice* from the patient and the public missing and so hence we set up this study to try and gather some attitudes from the public *towards* this.

And this particular research is being translated into many different languages and if you've got any ideas of languages that you'd *like* to see this survey in, please let me know because there's an offer of first author articles if you want to translate an arm. The Russian translation is going to be one of the first ones to come out at the end of the summer.

So what are the *risks* of donating data. So if we were going to *ask* somebody, "please, will you donate your data for research or for use by others," *really* when you actually have an electronic form of DNA, it looks something like this **[refers to slide]** It's not as if you can *really* look at that and say "Ooh, this person's got an increased risk of breast cancer." You need specialist software to overlay on top of this and interpret this.

But the *risks* that are thought about in terms of data donating and data-consequence sharing is that there's a *theoretical* risk that the person who donated the data could be identified from it.

So although their data has their name and address often severed from it, there may be a *link* to that data that somebody holds *back* to the personal information. There's always a theoretical risk that data could be hacked, maliciously or it could even be accidentally, and that somehow there could be a link between the two made, particularly if people have got a lot of other personal information online.

So this is *thought* to be one of the *problems* with donating data, is that there's this risk, *very small* risk, in theory, of being able to be identified from it. And actually, you know, when we think about our DNA data, it is our most personal profile, it's more personal to us than an iris scan or a fingerprint. So we do need to take the security of it very, very seriously.

So I'm interested in what people *think* about this? What do they perceive the risks to be?

So we've done this survey, we explore *lots* of different things to do with perceived harms, perceived risks, awareness of this whole industry going on. What does *trust* look like? If you're going to donate your data, what makes you *trust* the process, the system? Is it a white lab coat? Is it a fancy website? Is it very *clear* guidance on sanctions that would be in place if there were data breeches? So these are things that we're exploring as well.

And this is the survey **[refers to slide]**. And we have little Charlie in here, who *invites* you to participate. And then you work through the survey and, along the right hand side, are the different sections of the survey.

So we've got the first *thousand* or so responses in and I'm just going to share with you a *little bit* of the data. This is very much preliminary, it only came in yesterday morning so, I was busy yesterday! And I'm hoping for about ten thousand or so responses in total. So this gives you a *little* flavor.

We've got a spread already of ages, demographics, 50/50 men and women. *Most* of the people who filled in the survey know absolutely *nothing* about genomics already, they're completely disconnected from this. And you might say "well *why* is this even relevant to them?" If they're not engaging with health services or research at the moment *why* do they even need to know about this or comment on it?

Well, the *thing* about DNA and genomics is that any result that comes back to an individual is, of course, it *could* be relevant to biological relatives as well. And there is this urgent need *even if you're* not currently using health services, somebody you are biologically related to probably is and we need to be able to, sort of, raise consciousness about what genomics can do for us as individuals and *help* families to talk about what testing they've had, why it might be relevant for other people, *share* information about risks to ourselves. And that's another thing that we're looking at in our department is how to, for example, engage *popular culture* to start conversations in families about genomics.

Anyway, in this sample, 57% or so knew nothing about genomics. And the rest who took part were either health professionals, researchers or patients who were currently engaging in this field.

So one of the things I was quite keen to look at right at the beginning of the survey is: is genomic data special? Or different? to other sorts of data that you may have online. And I often get a sense from working in this industry that we feel that genomics *is* incredibly special, this data. It links people

to each other. It contains data about our past, present and future health. *It's incredibly* special. But do other people see it as special or do they just see it as data?

So one of our first questions was trying to orientate people with the concept of just having online data and genomic data just being added on to that. So the question was: "Let's assume you have the following personal information online: holiday photographs, your bank account details, your medical information and your DNA information," and we explained clearly what that was. So let's assume it's online. Do you think these pieces of information should have the same levels of protection against unauthorized use?

So that gives you an idea of comparing pieces of data between each other. Is there something that is a little bit more risky, would you be more *worried* about it if it was hacked? And *most* people across the board thought that their bank account details required the *most* level of protection. So what their saying is "I'm more worried about my banking details being hacked than I am, really, about my medical and my DNA information. It's *more…* I can see a real risk from that and a real harm from that."

And that was interesting, particularly interesting for me working in genomics cause I *really* was under the impression that genomics was incredibly special to people. And when you think about the systems that *are* being created to protect this data it is more secure than the banking system – well, I exaggerate, but you know there *really,* there are millions and millions of pounds that have been spent on making *the most* un-hackable systems. And of course you're only one step ahead of the hackers anyway.

Would you donate your data? And we presented three different scenarios. Would you donate your data for use by clinical people to help get a diagnosis in a patient, for use by *non*-profit researchers and for use by *for* profit researchers? And this is where we found a difference between the responses from the patients and the health professionals who *knew* something about genomics versus the public who, currently, are very disconnected from it.

So here **[refers to slide]** we've got the results from those who know something about genomics, either through their work or through being a patient. So they said "Yes, I'd donate my DNA and medical information." 70% said for use by medical doctors, 66 for use by non-profit and then we start to drop off when it's for use by *for* profit researchers.

And the comments were from people were saying things like "If I donate my data freely, I don't like the idea of for-profit companies making money and *exploiting* me in some way." So *that's* kind of interesting cause that, sort of, demonstrates a sense of, you know, "do I own my data after it's left me? And I should have some attachment to that and some decision making process about who has access to it or not." And I'm sure Jonathan's going to talk in a minute about ownership. And actually the reality of that is, perhaps, somewhat *different* from what people perceive.

And then **[refers to slide]**, the results from the public were – same pattern, but significantly less of them. So, yes, some of them would still donate but *less than half* of them would donate for these three different situations.

So there's this idea that if you're connected to the field you're *more* likely to *be* altruistic and say "I'd like to donate my data." There was definitely a concern about pharma making profits from free donation and those who know about genomics are more likely to donate.

What about being identified? So this is this *potential* risk if you did choose to donate. Are people worried about that? And if so, what *are* their worries? What do they *think* can happen if they're identified?

So if you think back to that previous image I showed of all the, just, the rows of data. Say, that was your data, and that was online, and your name and address and phone number was attached to it. What *harm* could actually, in reality, happen?

So we said: "Let's assume your anonymous DNA information *was* publically available so anybody could access it." It's anonymous. But let's say someone then *linked* your name and address and phone number *to* it, do you think you could be harmed? If yes, what concerns you most?

A lot of people said "I really think there is too much harm from this." But some said "yeah, I'd be worried about this." And the biggest concerns were: *being targeted for marketing*. Which I think is actually really interesting. I mean, the other answers are insurance discrimination, employer discrimination, being stigmatized in some way. And these come up *all* the time when you talk to anybody about genetics. Oh, "will I be discriminated against on the basis of my genetics?"

And given that we have a moratorium in the UK, you should *not* be. Insurance companies *aren't* allowed to ask you the results of a genetic test, they *are* allowed to ask you about your family history. But, certainly, targeting for marketing is not something that I would have, particularly, labeled as the *biggest* concern. And we're being targeted for marketing *all the time*, aren't we? So, interesting that this came up in this context.

So these are just some – there's *loads* of results I could have shared with you, I've just picked a few because I *know* we've only got ten minutes.

So, some preliminary conclusions: Some people are OK about donating their data for use by others, but *less* keen on its use for profit research.

Believing there is a potential for exploitation affects interest in donating, however, *actual* exploitation, in reality, might be very different. And that's up for debate.

And we're going to be keeping the survey open for the next couple of years and it's going to be translated into *lots* of different languages so we'll be collecting *loads* of data about this for the Global Alliance for Genomics and Health.

And this is the website at the bottom **[refers to slide]**, yourdnayoursay.org

Thank you.