

Andrew Morris
[MALE RESPONDENT]
[Other comments:]

JULY 2018

CRI - PRESENTATION - 12 MINS

My pleasure next to welcome my good friend and colleague, Professor Andrew Morris, the director of Health Data Research UK. Andrew is going to talk to us this morning about cutting-edge data science to address the most pressing health research challenges, Andrew.

Thank you, Alex. What a great pleasure, so thank you, Naomi, Rory and Cathy. It's a great privilege to be here, and this is, I think, a great celebration. I would echo Alex's words. I was reflecting that I think arguably this is the best example I know of open international biomedical science at scale, so let's be proud of it for the UK. You're the SERN for biomedicine, Rory, and I think it's appropriate that we're celebrating this on the Summer Solstice, because I think there is a lot of scientific sunshine ahead of us.

I want to talk about HDR UK. I just want to briefly introduce who we are, secondly, talk about how we are forging a very strong partnership with UK Biobank to really look at driving data science into health at scale, and an exam question for me is this: can we scale this to 65 million people to make the UK people truly world-leading? Definitions are important, so my definition of data science is the interface between mathematics, statistics, computational science across all its strata - Cathy touched upon some of these disciplines and domain - and our domain is health and social care delivery in biomedical science. That's the script, and this is relevant to the fourth industrial revolution. If you go to Davos, this is Davos in Switzerland - I've never been invited I should say - they talk about the fourth industrial revolution because we're seeing a fusion of biology with physical sciences and computational sciences, and it's disrupting every industry in every country. Our opportunity is to bring these tools into health and biomedical research in a trustworthy way, and why now? Here is old Leroy who talks about pre-emptive predictive personalised participatory medicine enabled to genomics, phenotyping, informatics, analytics and the derivation of a new social contract, so that the public, or the publics are behind our endeavour.

I was giving this talk to some finance director colleagues in the NHS, and a guy put his hand up in the back and said, 'You've missed out a fifth P', and I said, 'What's that?' I said, 'It's parsimonious', but that is important because we've got to look at value. As we use data, how do we drive value through the health system? The opportunity, so this is a slide of domains in computer science, is to bring this very strong toolkit into health and biomedical research, so yes, it's about data linkage, but it's far beyond that. It's about NLP, speech recognition, even robotics and sensors. How do we bring these tools in a meaningful way into our research and health systems?

That's what we're trying to do with HDR UK. We want to make this UK-wide. Our vision is to create a UK-wide network of research excellence that drives new scientific discovery by harnessing these new technologies in a population of 65 million people. That's the prize for the next 10/15 years. We are very grateful to our support at the moment from an initial investment from the MRC, and I think partnership with Biobank makes sense. We share six of our funders. We now have a funding pot of about £120 million, which

we can do something meaningful with, detail and scale across the UK. Our model is to be institutionally agnostic, so we actually sit in the Wellcome Trust. We're a separate entity. We're very liked by a bank, a company limited by guarantee with charitable status. We don't want to be a data controller, but we want to really provide the tools, the governance, the public engagement that allows us to unleash the value, particularly from NHS data at scale. We're doing this establishing substantive scientific sites across the country. The key is that not one institution or individual can do this themselves. This requires a scale of partnership and collaboration arguably never seen before, so our institutions, all 21 of them, subscribe to a single set of terms and conditions, as I said £150 million, but it's early days. I like data, so we are 323 days old. Rory, you're 4,389 days old. You don't look it, so I won't ask you to engage.

We have a triple aim, and we're sort of modelling ourselves like the EBI for health data, so we're doing three things. We're doing great science, we're training the next generation, but I think the biggest prize for the UK is to create a UK-wide set of expert research data services, unleashing value, phenomics particularly from NHS data. All of this is predicated on trust. If you are trustworthy, you're reliable, honest and competent, and I think that's a key facet. Our desired impact will be purely potential, but we've very keen on cohort enhancement because the lessons from Biobank, how do we scale that? Three million people in the UK have consented for participation in cohort studies. Every one of those individuals should be followed-up using valid phenotypic data derived from NHS sources. We're partnering at scale after an initial competition. We have six substantive sites, as illustrated here, 21 universities. We do have a gap in the north and a gap in the south-west, which we are determined to address in time, and key to this is breathing the same air.

If we are going to get computational science into medicine and healthcare, it's about bringing disciplines together. It's great to see emergent locations around the country like the BDI in Oxford, like the Big Data building in Wales where we're actually driving this science. Key to this is leadership, and I'm delighted we've got a very strong leadership team across the country, many of them very familiar with the UK Biobank community.

This is our initial research landscape in terms of actionable analytics, precision medicine and clinical trials and public health, and within these research activities there are many which are targeting the derivation of new phenotypes, which will be relevant to the UK Biobank follow-up. We want to develop this further, so we've currently put about almost £4 million aside for so called implementation projects, which actually allows us to scale the science, so this will go live very soon, and I think this is an opportunity to actually pump prime partnerships for buyback.

The second thing is training. We've got a very strong training team in place: George, Gabriella and Professor Peter Diggle. We're looking at how do we provide a career track for these key people, the technologists, the folk who look after the curation of data, the ontologists who often don't get the recognition they need? In addition, our vision is to train at least a thousand health data scientists who will all access UK Biobank, Rory, by 2021, so we've partnered with INHR - this is Dave Jones - and we've got an initial cohort of 250 who we hope to support through health data science training, education, summer schools et cetera.

The third thing is data hubs as part of the life science industrial strategy. I think John Bell would have talked about this today; three main themes in this phase: genomics, AI and machine learning as applied to medical imaging and digital pathology and the creation of digital innovation hubs. We've been invited to look at creating the EBI for health data. This was as yesterday. How many computer scientists can you get into a very small room? We, in early '29, will be going out to commission three to five digital innovation hubs, which will be about driving phenomics and reliable phenotypes from routine NHS data. It is shown schematically here.

Can I applaud, as Alex did, NHS Digital who actually have agreed that we should have a national research data services platform on their data. We've got the NHS coming together with the local health and care records, and we will provide an abstracted functionality upon that, so we've got quite a lot of money to actually demonstrate confidence in concept.

This is all about partnership, right? Rory emphasised that; international and UK, so on a UK-wide we're working with the BHF to look at creating a BHF centre in cardiovascular data science at scale across the UK. We're working with NHS Digital. Sarah Wilkinson, Tom and Noel Gordon, yes. I usually say Noel Coward. It's this sort of mental block I've got, so Noel. They have been terrific, and they are absolutely committed to putting research on their national data sets, and that is a sea change. Finally, of course, working with Turin, Adrian Smith and Chris Holmes from Oxford to look at this key interface with computer science. There is an example of that. You can never get Cathy away from the microphone, as you can see, working with Turin excerpt talking about how we derive phenotypes and emerging data, detail and scale, getting the expertise around the table.

This is moving very quick. If we don't do it, the Chinese will, so there is a race on internationally. The artificial intelligence eco-system is, I think it is misplaced optimism at times, but I think we have the expertise to do that. There is a big thing around policymakers and regulation, and, lastly, convergence of technologies, but also convergence of research initiatives, so the whole is greater than the sum of the parts.

In summary, I think there is a huge opportunity for the UK to develop what I call a deep informatics workbench to derive new phenotypes, especially from the wealth of NHS data we have, but it's challenging because of its quality, its completeness et cetera. We propose to focus on so called precision phenotyping, developing the algorithms, the tools to actually derive new reliable phenotypes, and I would ask, on the Summer Solstice, to say is it time to focus on metrology in healthcare, the science of measurement to reduce uncertainty? That's what deep phenotyping is, because I'm sure you are all aware that today the earth's maximum axial tilt towards the sun is 23.44 degrees. The physicists have been doing this for years, so it's about time we were precise in our phenotyping.

I'll finish there. To say it was going not badly, but then my director of communications phoned me up and said, 'Andrew, we've got a major problem. HDR UK has been taken by Haverling Demolition and Recycling', and actually I prefer that as a name for a national institute, so that's now our new title, so thank you, Rory, and team and look forward to working with you.

[END OF TRANSCRIPT]