M1: Thanks very much Paul, that was a great talk. Now I'd like to welcome Professor Paul Matthews, our second Paul of this session - I think that's why the AV got a bit confused - who is at Imperial College London. Paul has been one of the leaders of the UK Biobank Imaging Consortium, which I think is a tribute to consortium collaborative working. Many people inputted into that as I'm sure he'll say, and to also the vision of the funders who committed to, what at the time looked like a very ambitious study, but as I think we heard earlier today the 40,000th participant has already been imaged. So, Paul, looking forward to hearing more.

M2: Thanks Bernard. Well, you've already seen these slides, but I just want to emphasise what Bernard said at the beginning. What I'm going to say now is a tribute to an incredible team effort. The UK Biobank team is working seven days a week every month of the year and they've been doing so for years and they will continue to do so to deliver this study. The funders took an enormous leap of faith. It was a near impossible looking task at the beginning and the imaging working group consortium, along with many other people across the world has contributed to this sort of advice that helps to keep this on the road. It's an amazing effort. I'm going to take you on a quick run through some of the things that are beginning to be discovered, but I'm going to start by telling you a little bit about the UK Biobank imaging enhancement. I'm just going to lead you through in numbers.

So there are five clinical imaging modalities that are being performed on each of the subjects in the study. There are four imaging centres that have been set up. You see here the target by May of 2023 is to image 100,000 of the 500,000 in UK Biobank and the progress that's been made is shown in the left hand part of the slide, where you can clearly see the target is well within sight given the fantastic progress that's been made. It's being conducted across three countries within the United Kingdom. Within the examinations across all of these sites, more than 90 per cent of the planned protocol is being achieved despite the complexity of this degree of clinical imaging within the tight time constraints needed to get through 100,000 people in this timeframe. There are a thousand approved projects within UK Biobank at this point. Sixty per cent of them are asking for this imaging phenotype data. Twenty three per cent of them are asking for the raw data. This is beginning to make an enormous contribution to the science that people are doing. This is out of date. It was Monday. Today is Wednesday and there have been many more. Forty plus thousand people have gone through at this point. It's an incredible effort.

There are lots of words that can describe the uniqueness of this effort. These are just a few that I was thinking of last night. The size is impressive. This by far is the largest prospective imaging study for research in the world, but it's also multimodal. The data is all rigorously quality controlled with these dedicated sites and on we can go about the benefits of embedding this within UK Biobank. There are some extraordinary things that have been demonstrated as real tour de force within the field. This was from Carla Miller and a number of colleagues within UK Biobank demonstrating the relationship between brain imaging phenotype measures. That is to say measures of shape and size of different parts of the brain and 11,000 individual characteristics for UK Biobank participants being imaged. What I want to just highlight here is this is the
negative LogP scale, that's negative LogP 50. This is the Bonferroni correction line for this 2500 by 11,000 contrast and here you see all of this incredibly rich univariate correlation.

Of course, this is multivariate and we're not even scratching the surface yet. So this is work that Wenjia Bai at Imperial has recently done. This is looking now slightly differently. This is looking at the 2500 brain imaging phenotype measures, negative LogP and each of these colours represents a different cardiac imaging phenotype extracted using machine learning methods to segment the heart and large vessels in the chest in order to begin to understand this. Why is this important? By looking across imaging precision phenotypes, we can begin to understand disease, disease risk and comorbid disease patterns in much more detail than we've ever been able to do before. What's happening with UK Biobank? What's the value of it or at least the imaging component of it?

Well, the first thing is if one looks at the 50 plus papers that have been generated and published to date, since the data started to be released in late 2014, what you'll see is that there has been an enormous growth in artificial intelligence and machine learning tools for deconvolving the complex clinical imaging data into metrics that have potential clinical meaning. This is really important in itself, because these are the tools for the NHS tomorrow. By delivering these sorts of precision phenotypes there are beginning to be examples just like those we just saw from Paul in the eye, looking across the genotypes to begin to identify the sources of population variance in these important characteristics. Steve Smith published an impressive paper from his group and the big data group in Oxford last year for the brain. New data is beginning to emerge for the heart, variation in cardiac phenotype across the population. Really exciting finding is being shown to be driven just as we heard earlier Gerard described by variance of genes that in their extreme forms drive cardiomyopathy or disease. It's an exciting new range of development.

Now, Paul has highlighted a study which I just want to come back to. Imaging is as important for the eye. We haven't married these yet, but there is an opportunity in the plug I'll come to at the end to explicitly marry imaging of the eye with that of the brain. Remember the eye is the one place in the body where we can observe the central nervous system directly. It's also one of the few places in the body, as Paul showed us, where we can get very precise measures of microvasculature to complement these large vessel studies I was talking about. This is the study from Google and Deep Mind that Paul was talking about. If they used an AI trained on the eye image alone rapidly acquired literally tens of seconds at most, they could predict 70 per cent of the risk of major cardiovascular events. The important thing that I want to highlight here is this 70 per cent is rapidly, inexpensively achieved. Here if one looks across a standard method for clinical assessment of future cardiovascular risk, which uses for example a cholesterol as well as a variety of phenotypic measures, the predictive properties are certainly in exactly the same range. This actually begins to herald an age where imaging can be made affordable for the NHS at volume and can be significant for the management of patients.

Let me tell you another example which is actually one of my favourites. This is from Jennifer Linge and colleagues at AMRA and the group at Westminster University led by Jimmy Bell. Here are all of us somewhere along this spectrum of distribution of body fat amongst our organ systems. Everything in blue here
is subcutaneous fat. Everything in red is visceral fat. This was picked up and mapped in a ten minute abdominal MRI examination as part of the UK Biobank programme. Here we see the upper thighs from which one can also extract ratios of fat to lean mass and of course we have the liver lurking in there from which one can also derive fat measure. The distribution of fat in the body between the liver and the visceral and the subcutaneous fat and in the muscles in the top of the thigh tells us a great deal about the risk of disease and they've begun to decode this.

Again, a ten minute imaging examination. These are the risks. This person looks like a very well-proportioned woman and has an incredibly, despite the subcutaneous fat, a very low likelihood of cardiovascular disease or type 2 diabetes. This rather large male, again well-padded in a similar way, but now with a large deposition of visceral fat has a massively higher risk. This is a ten minute examination. It's beginning to show how UK Biobank data can be applied to solve real clinical problems that are important today. What's the future? I think, as I said at the beginning, May 2023 when we'll be through not just the original 100,000 target, but also some more as I'll describe in a moment. A cross sectional image at one timepoint is a really powerful tool as I've just described, but in fact, disease trajectories are complex. This is using a very simple set of phenotyping. Each of the lines represents two timepoint measures with roughly a five year period between them.

What you can see is that from these lines, particularly when one gets to vulnerable periods in later middle age, one can predict a trajectory of health for the long-term future. Some people do very poorly. Some people do very well and the slopes of those two point measures are incredibly informative. The one point measures that are being taken now are adding to our phenotypic data, adding to predictive capacity, but in fact, a single timepoint measure can map across all of those trajectories together. If, what we do instead, as we start taking that single timepoint measure and add a great deal more data that we can grasp from UK Biobank and clinical records, we clearly can begin to narrow that down, but making fine discriminations that are very important for planning and for the patient is more difficult. If we start to add repeat imaging so we can determine trajectories of precision phenotypes based on the same sort of two point measure that I've described there, we can begin to map people on to individual trajectories and with short-term follow-up we can begin to make moderately precise discriminants.

This is happening as part of a dementia platform UK pilot of this approach which will add 10,000 reimaging visits on with a two year interval before the 2023 date. Just think what we could do if we added more. So, we believe that we can get back an additional 60,000 of the original 100,000 group. One could add on at relatively modest cost given the fact that the centres are in place, the reimaging for another 60,000 making 70,000 in total and that could be done before a key date 2026. By 2026, 6000 of these people will be developing dementias, 1500 Parkinson's disease, 3000 will have Mis. These measures will allow us to develop the predictive algorithms that actually not only can tell us a great deal about the disease mechanisms themselves, but also begin to be mapped translationally back into the health service. This begins to describe the increased precision that one could begin to achieve.
So let me close by going back to the major point that I started with at the beginning. This is a team effort and I want to acknowledge the enormous work of the UK Biobank team for this, the efforts of the imaging working group who have given so much of their time to help advise UK Biobank and I just want to emphasise that they are backed up in turn by over 125 researchers across the world who helped establish the robust protocols that have set this on to such a successful interim outcome. Thank you.

[Applause]

[END OF TRANSCRIPT]