Andrew Dolan 0:01
The next speaker is Dr. Nathaniel Hupert. And I'm going to break with convention and ask Nathaniel to give a bit of background to his area of interest prior to starting, because this man has more jobs than he's going to see Donald Trump. But Nathaniel is a very, very busy doctor in a very, very busy Hospital in New York dealing with COVID. But he's also an educator within the Medical University system. He has also been an is an expert associated with the US Centers for Disease Control and Prevention. But there are so many other sides to him, including one of the things I know he's going to speak about, which is pandemic and pandemic modelling, and how artificial intelligence is influencing many of these developments. So on that basis, Nathaniel, the floor is yours. Thank you.

Nathaniel Hupert 1:07
Thank you very much, Andrew. And I'm going to apologise right off the bat. I'm in an office where there are several zoom meetings going on at the same time. And in addition, they're, they're digging the foundation for a new skyscraper next door. So you'll hear some pile driving in the background. I want to thank Dr. Woolen, for his very kind words and his very fascinating talk. And I want to apologise also to the prior speaker, I was actually giving paediatric grand rounds for my institution and only came in at the very end of that. Let me pull up my slides, if I may. So by way of introduction, I am a practising physician actually spent the last three nights on call at Lower Manhattan Hospital where we had, you know, we're seeing an uptick in in cases, with some COVID cases. I'm a hospitalist so I only practice in the hospital now. In the past, I've spent 10 years up until last year, as a senior medical advisor for the Division of Preparedness in Emerging Infections at the Centres for Disease Control in Atlanta. I also previously had a stint as an advisor for the National Hospital Preparedness programme. Currently, I'm at Weill Cornell Medicine, which is Cornell's Medical School in New York City. I'm in the Departments of Population Health Sciences and of Medicine. And I'm the co-director with wonderful emeritus professor of engineering Jack Muckstadt of the Cornell Institute for Disease and Disaster Preparedness. I work with many individuals who are true experts in AI, including some who have come over from IBM Watson. I am a modeller, I've been creating computer models for bioterrorism and public health preparedness for the last two decades. I would not consider myself an expert on making them, but I would perhaps consider myself someone who's knowledgeable about using them. And that's what I'm going to really focus on here. Let me start off with this interesting picture. This is what's happening in the United States, in terms of cases in green, hospitalizations and deaths in orange. What you can see is that we've had three waves in the United States, each one of which is dramatically defect. The only classic one is really the middle wave where you have the rise in cases that's the green, followed by the rise in hospitalizations, followed, sadly, by the rise in mortality. If you look at the first wave of hospitalisation, and certainly the mortality curves, and I should say the mortality axis is on the right, and the case and hospitalisation axis is on the left, what you see in the beginning is a dramatic divergence between the cases which are low and the hospitalizations and mortality which are high. You can see that there's the opposite of which is the cases are far higher than hospitalisations and the mortality is following the hospitalizations. What seems like a slave so this is very confusing. What's great is that we have AI solutions to fix this. And what I mean is you can simply using Excel, dial up a sixth order polynomial and you have a perfect fit for your cases with an amazingly high R squared, which is a measure of the fit, and you could sort of say, well, we're done. But the question is, is this and I'm putting "AI" in

quotation marks, of course, because this isn't really AI, I mean, it is it's artificial, it's smart, because it figured this out in a fraction of a second.

But this isn't really AI and this isn't really an application, that's useful, because we have no biological plausibility for this sixth order polynomial. And we have no confidence that this sixth order polynomial will for any reason, predict anything in the future, that's helpful to us. And yet, it seems like an application of technology to a very important pressing problem.

So what's the alternative? Well, the alternative is something that I've been involved in now for the last several months, which is what we call a participatory modelling engagement, where we've created and I say we, it's myself and individuals at the University of Oxford, where the project is centred, and groups around the world, mainly in Low and Middle Income countries (LMIC). We have created this “thing” called the COVID-19 International Modelling Consortium, now on version 16. And what's interesting about it is that it's very un-AI. I should say, before getting into the details, that we are now 46 to 48 country members in this Consortium, including the entire WHO Eastern Mediterranean Region. You can see New York State in green, because I was part of Governor Cuomo's COVID Task Force briefly in the Spring of 2020. I used this model for some activities back in April and May.

This is a modelling environment in which you can visually fit cases, mortality, and hospitalizations and use that fit to explore the implications of various non pharmaceutical and pharmaceutical interventions. In that sense it's very non-AI, in that this involves you, the user, with your knowledge of biology and public health interventions, tinkering with information to recreate a pandemic, in a computer environment. I should say that more recently, we've actually begun to use more standard AI techniques to have automated fits. And so now you can create a template, literally, it's an Excel template of your best guess of what's going on. And you take that template and you put it into Dropbox and the technical team at the University of Oxford, now has a very sophisticated automated fitting programme, that in the morning, you go back to the Dropbox and you see what they've done with your with your attempted fitting, but you can actually get pretty far just with your own brain. And so this is an example, unpublished of an estimate that I made in late October of what might happen in the United States.

It runs out through the middle of next year. It obviously has a very high mortality compared to what we have now, which is a little bit north of 550,000 deaths . But what's interesting about it, is that I had to adjust for some things personally without the use of AI. I think AI could probably do a better job. And so one of the things that I'd like to do the next couple of minutes is just go through some of the things that I'm pointing out where I think there are perhaps not low hanging fruit, because this is going to be difficult, but certainly valuable areas of potential intervention for better and more comprehensive automated fitting of models like this. One thing that you'll notice is that to get the mortality curve fit correct, that's the lower red curve. I had to artificially inflate the apparent case curve in the first wave that is the green line above the red line. These two curves, the green and the red, meet at a certain point in June. Now I did this simply by visually and manually fitting the upper green curve to make the load green curve match the mortality curve. It turns out, that that was precisely the moment when testing in the United States achieved the WHO recommended level of only 5% COVID-19 PCR test positivity. I think this is strong evidence to support the idea that what the WHO is telling individuals around the world that you need to have a robust enough testing capability to only have 5% positivity. But certainly that is a number that is ripe for better fitting. I mean, 5% is a nice round number. But I'm sure that that's not a biologically sound number. It's a target. And it may be the best policy target. But it's certainly something that could be supported better. So what I've done, here is one type of evidence to support them. We actually have interesting data on hospitalizations now from this great enterprise called the COVID Tracking Project. And one thing that we can see is that this fit that I made by myself with my own hands without an AI actually tracks hospitalizations pretty well. I should add that in October, it predicted that we were going to reach almost 100,000 simultaneous hospitalizations. Now, at the time, we were down at 25,000. And I'd looked at this and said, I'm not sure that I believe that yet. And yet this [120,000] is where we are in the United States. And so again, it's the type of projection that I think would be really improved if we could have a more scientific approach to to making sure that the fit is accurate. Let me just quickly go on to another example. This is New York City as of yesterday, in terms of the trailing seven day percent positivity, you can see that the dark red, you're up to 9% positive test results. And it's scattered throughout the city. It's very interesting. There's an island south of the main part of New York called Staten Island, where people don't like wearing masks. And they're having a terrible outbreak there. So there's a lot of what's going on, generated from that area. Let me show you. This is a some very interesting information. This is syndromic surveillance data from the 53 emergency departments in New York City, you can see that this influenza like illness signal picked up the spring wave of COVID very dramatically. And that right now, there seems to be no wave. If you ask New Yorkers on the street, what's going on with COVID, they'll say there's a fall wave happening. But according to this measure, which is accurate as of yesterday, it doesn't seem to exist. Which is quite interesting. If you combine that with data that the New York City Department of Health and Mental Hygiene has produced the antibody positivity, it looks like, again, things are pretty flat. This is the percent positive antibodies, and there's a lot of dynamics of antibody generation and trailing off of antibody levels that probably underlies this deceptively flat curve. Since July, again, a place where AI is possibly can be very useful.

But this is the information about cases, hospitalizations and deaths in New York. And clearly there's something going on because since October, we have had a rise in all of these. And there have been little micro clusters. These are the five boroughs of New York City, you can see little clusters of outbreaks, the various boroughs, some sequential contemporaneous at the same time and some not. But if you look at the emergency department signal, this is broken up by age groups. And the dark line is the overall number of emergency department admissions for COVID like illness from September until yesterday, you can see that it's been fairly flat, around four per 100,000. And certainly very different than what happened in the spring. But the question is, this is just me looking at this is there is there a signal here that I'm missing? Could I be warning people about what's going on? So here again, with great opportunity for signal investigation using artificial intelligence, I started, I did a very cursory trial of looking at the signal by simply measuring the coefficient of variation which is that the standard deviation over certain period of Time of the numbers coming into the emergency room divided by the mean, for that same period of time. And you can see very clearly, for older individuals, there is a signal there, from the 11th, to the 27th of March, our big wave started just about on the 11th. And it extended into April. Of course, there was COVID circulating in the city before then we closed schools, for example, on the 16th. And you can cut this in different ways, this is the seven day trailing coefficient of variation with a higher cut off, versus the 14 day trailing coefficient variation with a slightly lower cutoff. What's interesting is that the younger age groups have excursions above this red line, but they're not associated with any outbreaks only in the older individuals 25 and above, does this seem to be a true signal. So these are the types of things that I think AI could profitably investigate, in the interest of time, I'm not going to go over this slide. But let me just skip to some, some modelling. So this is my fit, using our CoMo collaborative model of New York City. And when I modelled New York City as the central, geographically central borough, that's Manhattan, I found that there was a possibility for a second wave. This is a model that I made again, back in the end of September, early October. And so this is a prediction, I did not publish this because I wasn't convinced that this was something that was of publication or the quality because I didn't have AI or other types of scientific aids to confirm the fit that I was able to produce. But that's the sort of thing that that I think, again, would be helped with a more quantitative approach. I could mention just briefly, we're doing some other investigations of a different sort that I think there are many quantitative methods in data analysis that will be useful for this is one example, looking at overcrowding of housing, versus the overall prevalence of infection, you can see that there's a very clear overlap of these two maps on the left hand, that's the cumulative COVID cases, by neighbourhood in New York City. And on the right hand is percentage of housing that involves more than two people living in a room. And then you can look, again, at the larger country, this is simply who wears masks, and who doesn't in the United States, and, this matches up visually very nicely with the hotspots in the United States right now. And also with the cumulative caseload per capita in the United States. Again, you know, GIS and AI, I think go very well hand in hand. So there are many, sciences that that should be brought to bear on we're making public health policy about, about outbreaks like this and public health in general. There are also non-scientific fields like ethics and economics that should be brought to bear. And I'm just going to close very briefly with an example that I find fascinating right now, that just tells us how much we don't know and how much room there is for investigation. Let's look at Haiti and the Dominican Republic, they share an island, Hispaniola, and if we do some simple demographic based projections, this is using very, very well studied data from France, a group at the Pasteur Institute, if we enter information for Haiti’s demographic structure, and assume a 40% final attack rate, it estimates that there would be about 10,000 deaths and about 75,000 severe cases. If we do the same for the Dominican Republic on the other side of the island using the same attack rate, it estimates about 14,000 deaths and almost 100,000 severe cases. And that's because the demographics are a little bit older population.

But if you look at the actual data, it tells a completely different story. Here we have cumulative confirmed COVID cases with the Dominican Republic above 140,000 and Haiti hovering at less than 10,000. And the cumulative confirmed deaths were at Dominican Republic is above 2000. And Haiti is hovering around 200. If you look at a larger spectrum of countries, it's clear that the Dominican Republic is following a trajectory of Sweden. And Haiti is following a trajectory of Norway. Which raises the question, what makes Haiti act like Norway and the DR acts like Sweden, in relation to COVID. I should point out that the cluster above the DR/Sweden cluster is Europe. And the lowest point out there in space is the United States.

One slide for my last example: vaccine supply logistics. We have many, many challenges ahead of us. beyond simply predicting what the disease is going to do, these relate to human systems, communication, integrated with command and control, I think that there are going to be many opportunities for industrious and innovative uses of AI in the coming year as we try to vaccinate the world against this disease. And this is just at the start of that. So to wrap up, modelling, the type of modelling that I do, and that others can help greatly with defining what the problem is, what we should do about it, and what we can logistically do about it, it can help us clarify knowns and unknowns. And it can help us quantify things, the risks and requirements to it that are associated with our actions, but only if we have appropriate modelling frameworks. And that's where I think AI can really help appropriate data elements. Again, AI can help us tease out what is true, and what is the signal and what is not using sophisticated methods to validate what we do. And then having appropriate forum for application and discussion of the results of this modelling. Let me just acknowledge a whole bunch of people who have helped with this work. First, the group that Oxford headed by Lisa Wright and Ricardo Aguas and the COVID-19 International Modeling Consortium, colleagues at Weill Cornell including the immunologist Doug Nixon, and his postdoc, Daniela Hernandez, and then my colleagues in engineering, Peter Jackson and Jack Muckstadt and others including Robert Gougelet, and then the mathematical biologist Alex Washburne.

Andrew Dolan 22:43
Nathaniel thank you very much indeed. For rounding us off today.