#414_ Will Machine Learning Overtake Traditional Nutrition Research Methods_

DANNY LENNON: Today, we're going to be talking about the topic of machine learning and data science within nutrition research going forward. And essentially, we're going to be discussing a claim around whether machine learning and data science can overtake traditional research methods within nutrition. Before we get to the setup of the episode, let me introduce the latest member of the Sigma nutrition team. As some of you may have seen through our recent posts on the website, as well as on social media, we have a new data and research analyst on the team, Dr. Niamh Aspell. So Niamh, welcome to the show, and maybe to start, can you give people an introduction to yourself, your kind of research background and anything else you might think they'll find interesting.

NIAMH ASPELL: Okay. Yeah, thank you so much for the introduction. And, yeah, I'm really excited and happy to start joining your conversations that I'm working with you both. So just a little background for myself, I suppose maybe it's not so linear, but I have a degree in Human Nutrition, and during my undergrad, I worked on various different research projects. So they kind of focused on exploring different nutrients data, so vitamin D status, B vitamins, and different health outcomes in older adults. So

we specifically focused on cognitive and physical function, and this is an area that I became really interested in after my PhD then as well. So I spent some time working in the Institute of Neuroscience in Trinity College in Dublin, and I decided that I wanted to kind of merge the fields of nutrition and neuropsychology and psychiatry into kind of the basis of my PhD. So I worked with a professor of old age psychiatry in Trinity College, that was Professor Brian Lawlor, and Professor Maria O'Sullivan from the Institute of Technology in Dublin as well, so her background is more clinical nutrition. So we kind of merged both of those disciplines to create a PhD project on vitamin D supplementation, and different outcomes in older adults. So we focused on cognitive health, and did a small bit on physical health as well. So my PhD research, I suppose, it was divided into kind of two streams of analysis, those are the kind of predictive based analysis and exploratory observational analysis looking at different determinants of vitamin D deficiency, and we observed kind of some new contribution to that area of work. And then, we also looked at physical performance in vitamin D, and I suppose, all of that work then was fed in or some of the findings were fed into the design of the kind of first double blind placebo controlled trial for vitamin D supplementation, and cognitive health in community dwelling older adults. So I designed that intervention, ran that intervention, and got some really good kind of conclusions from that work. I also did a post grad in applied statistics, so I have a bit of experience in that area, and I kind of brought that into my next bit of work, which was in adult social care, where I kind of looked at analyzing risks of dependent older adults who are receiving home support, and their likelihood of transitioning into long term care, so needing more support. And I brought in some nutrition there, focusing kind of on malnutrition, but most of that was kind of cognitive health. And then from there, and the next thing kind of was working with a really

innovative food clinical trial company in Ireland. So I was a scientific manager for that company, and that kind of involved me working with different nutrition companies or scientists to design and obtain ethical approval for their clinical trials. So we tested lots of different types of products, and for lots of different health outcomes, and then for lots of different kinds of reasons, so it could have been safety testing or to make applications to regulatory bodies for kind of health [inaudible 00:04:24]. But at the moment, kind of just a wrap up, at the moment, what I'm kind of working on is research projects, kind of focusing more on tech solutions. So this is like AI enabled clinical decision support tools in healthcare. So my work kind of focuses on evaluating and ensuring that that tech is developed in an ethical and socially responsible way. So yeah, that's, I suppose, my background.

DANNY LENNON: Awesome. And that will, of course, be very relevant for much of what we're going to discuss today. And, of course, also here with us is, as usual, Mr. Alan Flanagan, so welcome, Alan.

ALAN FLANAGAN: Thank you, and welcome Niamh, I'm excited now to have you part of the team and our future discussions and content creation. So it's going to be a nice ability now to kind of really tease things apart, especially today I hadn't particularly looked that much into – I've come across it in different papers, and I've been familiar just from people asking me about it a little bit with the PREDICT study as an example, but it's not something I had delved into a huge amount, the whole area of data science, generally, independent of nutrition, and then in relation to its potential application for nutrition. So the last few days have been interesting, to say the least. So yeah, I think this is going to be, certainly for listeners, I hope that, if anything, it's going to be something that they find novel as a discussion, because I can't see that anywhere else has kind of delved into this.

DANNY LENNON: Yeah. And so, with that, I think, before we get specifically into discussing some of the data science machine learning that I'm going to get Niamh to kind of introduce as an area. I suppose the reason why we may even be discussing this is that if we look at the way nutrition research is currently conducted, and the various different methodologies that are used, there are clearly limitations in different ways with different types of methodologies, and we've discussed this on some of our previous episodes when we've discussed nutritional epidemiology, or where we've looked at randomized controlled trials, etc., etc. So I think without delving too deeply into that, again, maybe a good point, just to kind of set the stage here, and maybe I'll come to you on this first, Alan, would be to look at, as a refresher for people, kind of the primary methods that we tend to see used within most nutrition science nowadays. And then as a kind of basis, how you would put forth what are the kind of acknowledged limitations we have, maybe of some of those, and again, we don't go too deep for now, but what is a good kind of entry point for people as a refresher?

ALAN FLANAGAN: Yeah, I think the entry point is to recognize that what we are interested in from the perspective of diet and human health in a modern society is primarily conditions that are chronic diseases, lifestyle diseases. I don't like using the term lifestyle disease, but that's how those are described. Cardiovascular disease remains the leading cause of mortality type 2 diabetes, fatty liver, nonalcoholic fatty liver, and other components of the metabolic syndrome, etc. And then now we have this exponential increase in prevalence of a neurodegenerative disease as well to mention Alzheimer's. And these are conditions that are characterized by long latency periods they develop over the course of decades – as an example, with atherosclerosis can begin to develop from our teens from the second decade of life. And so, diet, in addition to other

lifestyle factors will influence the disease process, long before there's the manifestation of symptoms, and this is quite a challenge to then investigate. Because if someone presents them with a heart attack, for example, some sort of cardiovascular event at 55, we may have only started studying them prospectively at the age of 40, or even 50, and so, we have this challenge of trying to tease this stuff out. It means that for the most part, the mainstay of the field has been prospective cohort studies, generally considered of observational designs, one of the better or if not the best for a number of reasons you investigate, particularly, the dietary exposure and the related variables that you would gather data on, before the onset of disease. So it minimizes to a degree, some biases that would occur with retrospective studies, or with cross sectional or case control studies. And you follow this at the outset, healthy population over time, and there are limitations to all observational research; as we've discussed in some previous podcasts, some of these limitations are over exaggerated in some respects, particularly, the alleged disconnect between nutritional epidemiological findings and findings from randomized controlled trials.

There is a lot stronger concordance between those findings than people would otherwise acknowledge, and the recent Lucas [inaudible 00:10:01] paper in the BMJ corroborated a previous study that had looked at diet disease outcomes. And so, we move from epidemiology then we're often looking at hard endpoints like mortality, and then we try and take observations from epidemiology and look at them in the context of randomized controlled trials. But the model of randomized controlled trials, that is typically emphasized in health sciences research generally is the biomedical randomized controlled trials, high internal validity, double blind and placebo control. And as we've discussed before, a lot of these desired methodological hallmarks are simply untenable for a lot of nutrition related exposures,

particularly, if you want it to be a food, that is your exposure of interest, rather than a capsule pill or supplement form. And this has created its own kind of disconnected methodological challenge. And so, you have a lot of inconsistency in an evidence base, but an inconsistency that can often be explained if we really start to scrutinize the kind of methodological friction between what's desired as a methodological quality versus what's achievable. So as one example of that, in relation to randomized controlled trials that we've discussed before is the biomedical model will assume that you have a clearly defined exposure and a placebo, that's possible when you're investigating a drug. I don't happen to have a bit of statin floating around in my body, I can give a group of people a statin, they have statin, and we compare it to no statin. Nutrients exist on a bell curve of action, and so, you can have a randomized controlled trial, and you're comparing vitamin E supplementation, for example, on cognitive outcomes; but both your placebo and your intervention group at baseline already have sufficient vitamin E levels. You're comparing whether more of an off is better than more than enough, and many null outcomes come, and then people then assume that there is the RCT being correct and the observational finding being correct. And we have this supposed disconnect, but we can actually reconcile that.

So they're the two and then at the top you have meta-analysis and systematic review. Meta-analysis is as much responsible for apparent confusion in the evidence base for nutrition as any cohort study. It is prone to distortive lumping, the line Deirdre Tobias had in her recent review about grading nutrition evidence was quite scathing, but also to the point. Nutrition meta-analysis are conducted by groups lacking subject specific matter knowledge, lumping different studies together that aren't really combinable, often getting a weak or a null effect estimate and kind of a weak effect size overall. But it's a conclusion

that may not necessarily reflect the evidence base. So much of this comes from the fact that the biomedical model, and the hierarchy of evidence has been kind of superimposed onto nutrition science, and everyone has expected that will yield the same body of evidence that we would, if we were studying drugs. And a lot of the supposed limitations of nutrition science come from these methodological frictions at different levels of the hierarchy, whether meta-analysis, randomized controlled trials, or observational research. And there are absolutely limitations to each of these methods, but those limitations themselves, if we recognize them for what they are, and actually have a discussion about them, there's a lot of low hanging fruit to pick to improve methodology at each of these levels. And so, that's a point of discussion that's becoming a bit louder, particularly, as people call for nutrition to have its own grading systems, for example, and not simply take biomedical standards criteria assessments for either trial quality or strength of evidence, and just again, apply it to nutrition studies and expect to get the same results. So that's generally where we focus our attention is prospective cohort studies, randomized controlled trials and meta-analysis. But the idea that there's just a clear cut delineation in quality between the three of them is not as stark as it may be made to make out by the pyramid of evidence; and each of those trial designs comes with particular challenges when applied to nutrition, and indeed one of the most important developments that the fields can and is working in many respects towards improving is how to do each of these better. And so, I think that's as much Niamh is there anything you'd add there?

NIAMH ASPELL: There's so much, isn't there, to kind of talk about when we think of it, I was just scribbling down some points there and some tops, but yeah, I totally agree, I think observational perspectives and epidemiological evidence gets a real bashing sometimes. And I think it's usually because there's an overabundance of

publications that are taken, you know, we might have a researcher who maybe doesn't have domain expertise or doesn't have the statistical expertise, but they have access to an open dataset, and they want to publish papers. So if we pull down an open data source, anybody can find associations with anything without really considering it deeply, and sometimes the peer review process in certain journals isn't as rigorous where they would meet, maybe pull the statistics apart. So when we go to then meta-analysis, sometimes we can just throw everything in, and you might have two really strong well-designed epidemiological studies that are thrown in with 20 poorly designed end studies, and then you're getting really bad meta-analysis then as well. So I think, yeah, there's definitely, I suppose, negative aspects to some traditional research methods if we think of epidemiological studies, but good researchers would avoid a lot of those kinds of poor practices in their work. And as researchers or scientists, you can look at those papers and say, oh that's quite poorly done, or, that's really well done. But mass media would just look at it and say, well, that's a really good headline, so I'm going to take that, and that kind of confuses everything. So I think yeah, there's lots to kind of capture there.

And then in terms of the meta-analysis, again, it's just something that I noticed when I was doing my PhD, and I did a literature review of vitamin D status of all the epidemiological evidence, but you couldn't put it all together, because the way we put different cutoffs and criteria around what's deficiency in vitamin D, what's sufficiency, a lot of the meta-analysis that's there has just thrown that all together, where I might have been saying this person is sufficient, and the next person would call it deficient. So it just doesn't make any sense. And it confuses, like you say, some of those effect sizes then, and you just have null findings in your meta-analysis. So yeah, I think there's lots of different things that could maybe be looked at, and some practices from trials

that could be maybe incorporated more in observational studies, just in terms of registration of what's going to be done in the observational data. But yeah, no, there's so many different points, but I think that's kind of my take on it.

DANNY LENNON: Yeah, and I think we'll circle back to many of those that have been raised, and they'll pop up at different points throughout this conversation. I think for now, let's maybe introduce the concept of machine learning, and essentially, why are we even talking about this in relation to nutrition science is obviously to try and have some benefit for the quality of data we can get out of studies going forward. So before we talk about the kind of its role and within the context of nutrition science, maybe how do we even define that for people who are unsure, and how would you introduce them to the concept of machine learning, and then how does that fit into a kind of a research setting, and being able to hopefully improve some of the outcomes that we're going to get?

NIAMH ASPELL: Yeah, of course. So I'll probably start with kind of acknowledging that machine learning isn't – it's not a magic wand, it's not going to kind of change everything, but there's aspects of it that could be incorporated to kind of offset some of the issues that we already spoke about. But I suppose machine learning and data science applications, they've been applied at different levels so far in nutrition research. But it's not a new areas, so machine learning is seen as quite new or quite novel, but it's been around for a really long time. So I think the first AI program was created in the 50s. You might have heard that the story of how an IBM computer beat the checkers master. So the computer essentially beat this guy who kind of self-proclaimed that he was a master at checkers. For a long time people then kind of thought that the checkers game was solved, but not necessarily, it's just a really big milestone in terms of AI and data science. And then the field has obviously

developed an awful lot over the last half century.

So specifically, if we think about AI, and if we think about nutrition research and some of the research that we'll talk about today, they focus on that machine learning branch of AI, and that's been able to kind of change quite a lot over the last number of years, because we know kind of a couple of different factors that we've obviously got access to a lot more data. So we're all generating copious amounts of data all of the time. But in terms of the actual hardware and software development, we also now have a greater storage capacity. So we can hold all of this data, our computers are much stronger so we've got much better processing power, and we've got cloud computing. So if you think about kind of all of the data that you generate over the course of a day, and particularly over the last kind of two years, where we have much more kind of mass surveillance or public surveillance during the pandemic, we're collecting lots and lots of data. And we can engage a little bit with AI and machine learning and how we can use these technologies if we use that data correctly. But I think if we're to focus more on the technical aspects, so the machine learning, it's just a sub area of artificial intelligence. So the goal of machine learning is to create these learning algorithms to do the learning automatically without any kind of human intervention.

So it's broken into kind of two different fields or two different categories, so we've got supervised machine learning, and we've got unsupervised machine learning as well. Most people use supervised learning in a lot of the applications that are kind of currently available, and we've seen it in lots of other industries like banking and ecommerce and much more in healthcare now as well at the moment. But supervised learning, it involves kind of mapping a series of inputs to an output, and these are based on an example of labeled input. So if you expose a model to these

variables, say, its age, and you tag the age with bone mineral density, when the model is exposed enough for that data, it should be able to get that input data of age and predict that, okay, that bone mineral density for that age is around about here. So the machine learning models are usually then, they incorporate some statistical approaches, so they are basically they're on regression or classification. So for supervised learning, you would initially have two datasets, so you have your training dataset, where you kind of guide the training of the model, and then you have a test dataset which you kind of confirm the accuracy then of the model. So you're kind of best practice would be to expose the model to lots of data points, and then to have a separate dataset on the side that you then test the accuracy of your model to be able to predict those kind of data points then to make sure it's generalizable. So you've kind of created a model to see how someone might say, respond to a certain meal, but you need to understand what group of people are you looking at there in terms of their diet, because you might be able to compare a dataset from maybe a Japanese cohort to an American cohort, they might [inaudible 00:21:58] be able to predict that though.

For a regression type system, so regression, very same principle of really a traditional statistical method, and you're just wanting to find the target value based on an independent predictor, but they become much more complex methods in regression. So you might have seen things in some papers where they've published that their machine learning methods focused on decision trees. Essentially, this is just a hierarchy of information or hierarchy of – they refer to them as nodes or decision points. So you could say, at the top of your decision tree, say, does this person have elevated glucose, yes or no. Do they have visceral adiposity, yes or no. Male or female, age, lots of different information points. Obviously, it gets more complex than that. And the more nodes or information points that you

have, in general, the more accurate your decision and model is. These get much more complicated. Then you've got Random Forests, which is essentially, they're called ensemble learning techniques. So they combine decisions from multiple models, so they incorporate multiple decision trees. And that gets more confusing again. So then you've got neural networks. It's not really more confusing, it's just more and more layers of information, I suppose. The next is then your neural networks, and this is really popular, and I think in most of the papers that I've read, that's what they did the type of machine learning, and supervised machine learning techniques that they use. This is a multilayer model. So each layer of nodes or information has a different function. So you have an input layer, lots of hidden layers, and then you've got an output layer, and I think that that's kind of where a lot of the work is kind of focused. And then, again, so this is a regression with classification, pretty much the same models available except you're just using discrete data. So it's logistic type regression, and you might have like probability, much more kind of research questions that are based on probabilities of events occurring.

So they're the main kinds of techniques in supervised. And then unsupervised, it's probably not so important to this conversation. I haven't seen it used in nutrition research, but essentially, it just draws inference from different input data that hasn't been labeled. So you'll just have lots of information around a participant, but it's not labeled at what it is. And the computer, the more it's exposed to, will start to see clusters or patterns in the data that you're sending in. But it won't be tagged or labeled. Yeah, so that kind of – does that make sense or do you have questions on that?

DANNY LENNON: Makes sense on my end, and I think maybe that kind of sets us up to maybe start discussing some of the problems that we identified with, say, traditional research methods earlier, and then the promise for machine learning to be

able to either solve these or at least to avoid some of those pitfalls, or, at least, that's what we can look forward to. And there's maybe a few different ways this has been applied, and I know some of the stuff that you've sent me is highlighted a number of different ways in which it may help. So maybe let's start working through some of these one by one, and we can discuss them. So the first one: what is one particular problem that tends to arise currently within research that you see a clear benefit or a potential for the use of machine learning?

NIAMH ASPELL: Yeah, and so, I think one that stands out for me, and it's probably based on kind of practical experience, but then also kind of been involved in wider conversations, other nutrition scientists is the conduct and the some of the pitfalls around dietary assessments. So like data collection with dietary assessment, so I think we probably know a lot of the limitations, so whether it's in your food frequency questionnaires or 24-hour dietary calls, or food diaries, being kind of the more traditional approaches. We know that they're really, really dependent on the participants' motivation. There's a lot of participant burden or responder burden in conducting some of those assessments, particularly, when it's over a long period of time. And we've seen that people, and I'd probably be the same myself, people act differently when they know, they're kind of being watched, and we're kind of measuring their diet in some way. And I think going back to the epidemiological type research as well, where we're looking at, like you said, kind of showing those relationships with dietary patterns. I'm looking at disease, presenting over a long period of time, in fact, five or 10 years of time passing. Sometimes with those studies we'll only have a food frequency questionnaire at one or two time points. But we know people's dietary patterns change quite frequently, and I think the reasons why we don't collect that information more frequently is down to, I suppose, participant burden, time, cost, resources. And having worked in kind of

the more operational side as well of these food clinical trials, I know it's a big component of the on the ground kind of research team is to analyze these diaries or to conduct the assessment. And I think, if we removed maybe the hour of time that's needed, you might have more time to do other assessments, or kind of keep the participants engaged as well.

So this is one area of nutrition research where the machine learning or computational aspects have started to already come in, and I've seen it being applied in a couple of different tools that are currently being developed based on image detection and different techniques around image detection. So specifically, they look at using and incorporating a deep learning technique that's a part of the machine learning side of things, the supervised side of it. They use these things called convolutional neural networks, and this technique is really commonly used for image based analysis. So we've seen it in lots of other industries already. If we think about healthcare, it's starting to be used in different devices around radiology or pathology, where their role as a clinician is very much based on looking at patterns and patterns of change. But they started to use this a little bit more in research [inaudible 00:28:23] as well. So there's an example of Liu 2016, we can probably link the study, but the project that they were working on is called DeepFood. So this is this deep learning based food image recognition using a computer aided dietary assessment. So essentially, that they've also incorporated this convolutional neural network technique, where essentially, they show, they present lots of images that are labeled because it's supervised, so they label these images, this is a broccoli, this is a broccoli, keep telling it, it's a broccoli, and then eventually, you show [inaudible 00:29:00] hasn't seen, and hopefully, it will say this is a bit of broccoli. And if it says it's a cauliflower, you show more broccoli, and you keep training and keep trying to make it better. So they've assessed this, and tried to improve the accuracy

in reporting by analyzing foods that are captured on a smartphone.

So if you were to have an intervention study, and you wanted to conduct remote analysis, so the participants don't need to come in every six months or every year or however long, but you can send them a notification to smartphone, and say we want to check your diet now again for the next week. That decreases the burden on both people, on the researcher and on the participant, and you're getting more of an understanding of their dietary patterns, because we know they're obviously going to change more frequently. But in terms of the actual, of how this group applied this type of machine learning, they had access to two really large repositories of food image datasets. There's one Japanese dataset, and then there's also one in the US as well. So they developed this system where they were able to detect the food from these images. So they pre-trained a model using over a million, with 1.2 million images from the Japanese model, they trained the dataset on that, and then they tested it with an additional dataset of 100,000 test images that were never exposed in the model previously. It was much quicker, years ago it would take a really long time to train a dataset in terms of actually just the processing power. But that's all developed now, and it took about two or three days to train the model, and then the model was able to classify an image once it was trained in less than a minute. So it's gotten much quicker in terms of performance.

And what they demonstrated, once they trained two separate models, the American model and the Japanese model was that they were able to accurately predict or accurately define or classify the food types to 93 and 94% accuracy for each of those models. That was an improvement a couple of years ago, another group tried to do it and they got about 64% accuracy. So you can see how the field is constantly still learning. But if we look at accuracy rates of 94-95%, and they would be

considered to be quite good scores, there obviously is kind of flaws in their method, in that they use two different cohorts so the Western food group and the Japanese one. So it needs obviously this – their algorithm needs to be generalized to other populations, and there was just certain foods represented in that. So it still needs to be kind of expanded, you know, they expanded [inaudible 00:31:44] developed a little bit, but if you kind of think about the potential of that, in terms of the conduct of nutritional studies and epidemiological and intervention studies taken on this new type of conduct, and where you don't need to go to a clinical site for visits, you can start to do everything a little bit remotely, or at least report these other factors more frequently, what kind of less burden. So I think that offers a really good opportunity. I think it'll be exciting to see where that goes.

DANNY LENNON: Yeah, with the accuracy you mentioned, is that in terms of accurately identifying a certain food, and then thinking into the future where's the ceiling for this in terms of identify not only foods with close to 100% accuracy, but then portion size, etc., that's all done automatically, like, how far away is that or not?

NIAMH ASPELL: Yeah, it's really interesting. You said they looked at food components. So I think one really important thing is they don't go into any detail around, okay, well, there's a piece of chicken on a plate, but how did you cook the chicken, did you fry the chicken, or was it steamed or whatever it might be. So obviously, there's important interactions there, or, additional ingredients that could have been added that you can't see on that image, so that obviously needs to improve. And then, there's another, how it started to develop a little bit more with other groups, there's other evidence, I can't remember the word where this was conducted, sorry, but they had started to look at the fiber content of foods where they don't report on the fiber content of the food, where they're not required to do that. And they, at the

moment, you can manually look at a nutrition label and try and figure out the fiber content of the food, whereas these systems are able to calculate that for you. And so, I think they're getting a bit more specific in terms of components of foods, but at the moment, this was very much a classification. So the way it kind of works is it will break down the characteristics of the food, so that the model will see an image of say broccoli again, and it will pick out what features are important here, and you can train the model based on the features that you show. But I think it's been a very kind of basic layer, and that is just picking up these physical determinants of that particular food. Whereas we need to ask a lot more questions around the kind of preparation and the quality of the food as well, which obviously isn't considered in some of the models that have been created so far. So I think there's a lot more that can be done, and we could incorporate other supervised learning techniques, such as natural language processing, which can take some other information around the text that's available on foods that they're packaged foods, and make sense of that to give you a better kind of conclusion or feedback on what foods you're eating.

DANNY LENNON: Cool. Yeah, if I can just derail the whole podcast for a second, I don't know if you guys are familiar with the TV show Silicon Valley, if you've seen that before. It's a comedy show. One of the guys develops a kind of food identification app, and they think it's going to be amazing because he pulls it out, and then he puts it in front of hot dog and it correctly identifies, and they think they're onto something huge. And then they put it in front of another meal, and then it just comes up, not a hot dog, and then they realize that this guy just made something completely useless. But then it turns out, like the twist is that eventually it gets picked up by some tech company as a way to identify basically sexual harassment on social media, because it can

identify people sending nudes and things like that. But yeah, that's my inappropriate anecdote.

NIAMH ASPELL: There's lots of examples of how – I think there was one, I can't remember the full story, but some tech developers were developing a system where they wanted to be able to detect – I can't even remember it was a horse or if it was a man, but it was one or the other, and the image was able to detect, oh there's a horse – there's a man and a horse in this image, but it was a statute, it wasn't actually a man and a horse, it was a statue in a building. So yeah, they're still, in terms of motion and all of these other, and if we think of motion in what can be incorporated with natural language processing, we can look at video instead of just looking at image. But the tech isn't quite there yet to be that distinguishable, and it takes a lot of, if you think of all of the food types, and all of the different, you know, if we break down all the different food groups, you have to singularly take each of those foods and train a model to do it. So there's a lot of work involved. And so, yeah, it's still a good bit away.

DANNY LENNON: So then if we kind of maybe counter that on one side with, if we think of traditional research methods, specifically, when we're looking at dietary assessment, and this is something we've touched on a couple of times before Alan, of there are, as we've noted with other things within nutrition, there are better and worse ways to go about doing that. So in lieu of advances that may be coming down the line in machine learning for right now, how would we kind of summarize or think about dietary assessment methods the best way potentially to go about them, and where, yeah, how people should just think of where that sits within how we currently do research?

ALAN FLANAGAN: Yeah, and thinking of when Niamh describing this potential for the ability to even be more granular with artificial intelligence, identifying, say, weights of foods and portion sizes, I'm just

kind of thinking in my head that the current study that I have – well, it's finished now, but I had people use a food, a photographic food app to take photographs of their food, and it would timestamp the meal whenever it was taken. So it was great, so I'm getting that time component. But I'm still on the backend doing manual input, and I'm reliant, again, on the individual participants to – and thankfully, I had a highly motivated cohort, so I was getting a lot of detail. But I can see how there is some fantastic potential for some of these technologies to provide a means to really get accurate analysis of food that doesn't rely on the individual. For now, we're reliant on still a number of methods that do continue to be improved.

So for nutritional epidemiology, the main assessment methods for diet would be a food frequency questionnaire with the preference for it to be semi-quantitative, so a semi-quantitative FFQ. Semi-quantitative means that there are nudges as far as portion size. So as far, it doesn't just say milk, it could say a glass of milk, 150 mil, 250 mil, this kind of thing. So there are some elements of quantity involved in it, and there's evidence that that does improve the granularity of the responses that an individual would fill out. There are other assessment methods including 24-hour recalls. Niamh raised a really important point, which is participant burden. 24-hour recalls have a particular utility, because food frequency questionnaires do have a lot of burden, they're not particularly appropriate or useful in, for example, low literacy settings. And so, 24-hour recalls can often be used in those kinds of contexts, and can be worked through with a trained investigator. And again, this is more burden and more cost in the actual conduct of the research because you need a part of the research team, ideally trained in what's called the multiple pass method. So that means, I say to you, what did you have for breakfast, you say cereal. Okay, what type, did you have milk with it? Yes. What type of milk

did you have... And so you work through trained investigators, work through this, it's called a multiple pass method, and you'll keep going through to add, did you have sugar and all of this kind of stuff. That's burdensome on the investigators, it's slightly easier on the participants, and again, it can be used in certain settings.

With both of these methods, and this is a point that I think will be important for machine learning, it's a note I made when I was looking on through some of the links in relation to this food recognition technology. With both FFQ or 24-hour recalls, a really important part of the process is validation and reproducibility. So there's because food frequency questionnaires give us better data, most of the large cohorts that we will have either in Europe or the US, or really anywhere where a good well conducted prospective cohort study has been established, the food frequency questionnaire will typically be targeted to the population of interest, as Niamh highlighted the difference between say Japanese food image database versus the US. We wouldn't take an FFQ validated for use in Japan and apply it in Ireland. So population specific specificity is important, and then you want to validate that instrument, and typically, that has been done against a subset of your actual study cohort, usually give or take about 8%, and you take those people, and you get them to do a weighed and measured food diary for seven days, so they're weighing and measuring food. And currently, a weighed and measured food diary is the most accurate way that we can assess, quote-unquote, real intake in a given individual. And that's probably something people, many people listening would have more familiarity with now, because so many people do track macronutrients and weigh and measure food out. And really, that's what you're asking people to do for a week. And then based off the data that you get from that, that's considered your true measurements – you look to see how well your food frequency questionnaire correlates with the object, I

wouldn't even call it objective, but the real measure of dietary intake, for example. And what you get from that is what's known as the correlation coefficient, and that's anything between minus one to one, and minus one is a negative correlation, zero is no correlation, and one is a perfect positive correlation. And the correlation coefficients depends for different nutrients, depends on the nutrient of interest, the exposure of interest. For most major macronutrients in well validated cohort studies, there's actually decent correlation coefficient for total fat, carbohydrate, protein slightly less, and specific fat subtypes. Some particular nutrients would have good correlation coefficients, others would have weaker. And so by good, I generally mean say, 0.6 to 0.7, sometimes total fat, and then specific foods can have good correlation coefficients.

Now, one of the big knocks against nutritional epidemiology has been these methods, these correlation coefficients or the method of investigation of diet is inaccurate, and indeed, some have argued it's just inadmissible, Edward Archer in particular. However, these correlation coefficients are as comparable, sometimes better than the correlation coefficient that we would have for blood pressure, which is 0.55, or blood lipids, which is 0.65, or blood glucose. And yet, these are considered totally appropriate and uncontroversial and unquestioned prognostic biomarkers of risk for various diseases. So that epistemic inconsistency in the critique of nutritional epidemiology, I've never really seen a good enough answer given for that. That's not to say that you would use an FFQ, for example for establishing maybe vitamin B6 content, right, because it's really poorly correlated, or otherwise B vitamins generally along with the water soluble vitamins. So the whole blanket criticism of it is inaccurate. There is no blanket criticism. Right? The accuracy depends on the nutrient that is our exposure of interest. And with many of the major macronutrients that

we're interested in, and some micronutrients, there's sufficient granularity in a semi quantitative FFQ to allow us to derive meaningful conclusions from when an appropriate research question has been asked.

Now, more up-to-date evidence would suggest that you don't even need a full seven days of wage and measured food intake. Victor Kipnis' research in particular, they've shown that actually if you use multiple 24-hour collections, 24-hour recalls, you can obtain fairly similar correlation coefficients than you would if you used a full seven days. So that somewhat reduces the burden. One other particular thing that I think is worth mentioning, because it was a point in the discussion was this idea of reproducibility, and we may often have a cohort study that only measures diet at baseline, for example, or maybe has one follow-up period. Whether that is a positive or a negative depends on the population in which the cohort study is being conducted. Diet does change over time, but it's often incremental, and in particularly Western developed countries, it doesn't change that much, and this has been shown in cohort studies that have multiple follow-up periods, like the Nurses' Health Study, and some of the subcomponents of the European Prospective Investigation into Cancer. So you can have decent reproducibility over time in a population with a relatively stable diet, which many western industrialized countries have. That is still a limitation. Ideally, you would have multiple. But it's not as much of a limitation, as it would be in say, for example, a country like China or an Asian country undergoing what they call the nutrition transition. And as an example of that, the Chinese national nutrition survey data, if you look at the mid-90s, the macronutrient composition of the diet reflected what would have been the traditional Chinese diet, i.e., it's about 20% total fat, 65% carbohydrate, and their remainder to protein. And you have a study like the PURE study in which diet was measured during this transition period, and

they've come to conclusions in relation to both fat and carbohydrate intake that are quite contradictory based on our current knowledge, yet, you look at the Chinese national survey data now, and their average fat intake is about 33%, and their carbohydrate intake is split 55%.

So the fact that there was only a baseline assessment in these studies, in this case, is a major almost write-off, because there's been such a dramatic shift in their diet. So it's an important question to ask in relation to reproducibility is, is that a population in which diet is relatively stable. Of course, the more that we remeasure diet over the course of a follow-up period, the stronger the results get. But it's not always as fatal as it would be in the example of the Chinese changed nutrition transition in diet. So just to sum that up, the best available method for prospectively investigating diet is a semi-quantitative food frequency questionnaire, and it outperforms other tools like a 24-hour recall. 24-hour recalls are useful for obtaining population wide kind of snapshots and statistics, so they're often used for national diet survey collection, they're often useful as well then in low literacy contexts. There is a burden on them that's both an investigator burden, and there's also a burden in relation to completion for food frequency questionnaires. But as I've thought about some of the application of machine learning, and the potential progress of these available techniques, one thing that occurred to me is that we're still going to need to validate these tools, and we're still going to have a situation unless the artificial intelligence can do literally everything, i.e., the only thing that we're asking of an individual is to take a photograph of food. We're going to introduce – we're going to be dealing with different sources of measurement error, inevitably, because there is still going to be a level of participant compliance, so to speak, involved. Now it might be minimized. As an example, using a photographic app, you could have participants

that forget to take a photograph of breakfast, and then forget to take a photograph of the snack they had at 11 and 4.

So the idea that these are going to be assumption free, validation free and measurement error free, assessments of diet, I think is potentially untenable, because ultimately science is human endeavor and we're still going to be asking people to be complying at some level, although we may not need them to enter all of the extra detail about whether the chicken was fried or baked. But we will still need them to do something as far as getting the data, and in that will be measurement error. And so, while this is really exciting technology, we're still going to have to do due diligence to validate the tools, to account for measurement error, to know that measurement error, as has been done with food frequency questionnaires, there's 50 years of research trying to really get to the crux of where the source of measurement error is, and how it can be attenuated, and there's various ways of doing that that I think is adding a layer of complexity to the present discussion that we don't need to go into. But it occurs to me that we'll still need to do that with these new technologies, we'll still need to validate that, the output that we're getting from the algorithm based analysis of a photograph of a meal actually reflects what that meal contained, and we will need to account for the inevitable measurement error that comes with asking human beings to do something in a free living context.

DANNY LENNON: Niamh, any additional thoughts off the back of any of that?

NIAMH ASPELL: It reminded me of a couple of things as well, but I think, yeah, there's so much progress that needs to be made, and I think one important thing when you touched on different populations was digital access and digital literacy as well. So if we do conduct trials or studies, it requires an ability to either have

access to a smartphone or to have access to internet, and then also to be able to use it. You're a target in a certain population there, even though we might think, oh sure, everyone has a phone and everyone can use their phone easily and access it or whatever, but there's some subgroups of the population where that wouldn't be as straightforward. So you have to think about who you're excluding if we move towards these kind of smartphone technologies. And then, again, just around kind of participant engagement, I think if we move entirely remotely, some of the responsibility that the participant feels kind of lessons, so they might feel – I think there's a big part of the participant and researcher engagement that kind of makes the researcher, the participant feel like they're involved in something that they're contributing to science a bit more, and they feel like, well, I want to commit to this, I want to, you know, I've met this researcher, I'm really interested in their study, I want to help them out. So they are a bit more compliant, and they may be would stay on the study for longer, so you've got better kind of retention rates of participants if you kind of balance the time that you need with them. And I think that relationship is really important, especially in clinical trials, so that they're not just talking to save these smartphone apps to start then having chat bots where they never get to talk to a researcher, there's just told do this, do that, or whatever. They'll be more likely to drop by, and they'll be more likely not to feel like they're engaging or contributing to research as well. So yeah, I think there's lots of things to consider in terms of moving in that direction. And again, with the validation, yeah, it's back to square one, essentially. I suppose, all of these things need to be validated in terms of their tech development, but then also their application in real life.

ALAN FLANAGAN: I think I'll point about participants, and it's something that I became really acutely aware of with our lab study that we did in 2019, which Danny came in for a day of, and the

experimental measures, it was quite an intensive study day, and there's an incredibly human aspect to the conduct of science that can get lost a bit. And no patient or participant wants to feel like a number and a sack of blood that's there for the milking, excuse the crude expression. And I think we need to, with so much research, particularly for nutrition as such a behavioral intervention, and you look at some of the qualitative research that's started to come out in the last few years. And it's just, it's so obvious that the human voice is absent – nutrition research a lot of the time, high attrition rates and a lot of interventions, and we've often just not bothered asking why, and I think retaining that element of humanity and science, as we become a much more technologically advanced society is going to be quite important to the conduct of science. So I think that's such an important point in this discussion.

DANNY LENNON: Yeah, I remember that study you were doing and, I mean, one of the things you said to me at that time was the role of the nurses that were on the ward throughout that study and the help they've been not only to you but to participants, and given, especially like a study that was being done with this kind of circadian phase shift and they're stuck inside and people are probably on the verge of cracking, and...

ALAN FLANAGAN: They didn't see a window for eight days.

DANNY LENNON: Right, and so then if you think, okay, if we took them away from not only interaction with the researcher, but also the nursing team as well, who obviously have a clear expertise in how to deal with people as well and are typically good at that, it might have been very different.

NIAMH ASPELL: Yeah, I think it touches on that public and patient engagement in trials as well, when do you get that feedback, you'll design a study, but we only know it from the side of the researcher, if we don't get the participant experience, particularly if it's in a disease cohort, where

they have lived experience of what we're trying to understand, and we need to understand it from either the domain expertise, but also the lived experience too, because that can really alter how things are conducted.

DANNY LENNON: Yeah, one, if we can shift now from the dietary collection into one of the areas I was most interested to talk to you both about, and this is in the area related to individual response to eating or, more broadly, personalized diets, and this is an area where I think, in large part, what most people see in relation to this is often hype, and maybe not as much data a lot of the time, although, depending on what we look at, there's obviously studies being looked at this area, and there's maybe different types that we'll have to distinguish between. Some of the studies, for example, that you had brought up, Niamh was a couple of Eran Segal's lab. David Zeevi was lead author on that. People can hear him in Episode 298 of this podcast. Alan earlier mentioned, the PREDICT study, which again, is looking at kind of this individualized response. And so, this is one of the areas that is often touted as one of these primary areas where we can take huge amounts of data, and then use that to be very precise with not only how people respond differently, but then going beyond that and being able to prescribe individualized recommendations. So with that, maybe to open up, where do you see this current area of research Niamh, of the trials that have been done to date, and maybe even in comparison to what typically maybe gets reported in the mainstream media, what is the kind of current state of things? And then maybe to explain for people a bit more, what type of data science is working in the background of a lot of these major trials that we might mention?

NIAMH ASPELL: Yeah, I think this for me as well was one of the more interesting applications, I suppose, of machine learning. I say the same, kind of, go back and listen to that episode with David Zeevi, and that group in the Weizmann

Institute in Israel, because their work was really, it's really exciting, and there's a lot of opportunity with it. So, I suppose, they were initially kind of interested in kind of addressing what most nutritionists are trying to address with this kind of the growing incidence of metabolic disease, specifically, kind of focusing on diabetes and obesity, and they were interested in trying to understand why things were kind of changing, so this relationship between changing dietary patterns over the last couple of decades and also increased rates of obesity and diabetes. They conducted, I suppose, a couple of different research projects, and they all built on each other, but they developed different machine learning algorithms where they incorporated more than kind of just kind of basic lifestyle and dietary and health information.

So they integrated vast amounts of data, so they blood parameters, anthropometric measures, and they also included – and this is an exciting area of research as well, we are in the gut microbiome and so, so complex, and there's still so much to learn around different species, and the gene itself within the microbiome or more and they're trying to explore that a little bit further. They also included diet and lifestyle information in order to predict an individual, so instead of a group level, they said we need to start looking at this at a population level, we need to see how things differ from one person to the next. So they wanted to predict if an individual's and post meal glycemic response, postprandial and glycemic response was different. Again, they used some of these predictive kind of algorithms, and they collected data on, I think there was 1.5 million glucose measurements obtained. So they use the CGM or the continuous and blood glucose monitoring to be able to detect, typically it picks up a reading of your subcutaneous glucose every 15 minutes. So if you're wearing that, it's just massive and massive amounts of data and it's been investigated quite a bit now using probiotics,

but they collected a lot of data and glucose measurements, they also collected all of the information on the meals that were consumed.

So I think they're close to 50,000 meals, and they collected the data for more than 5000 days, if you incorporate all of the participants. So they measured, it's been major determinants of the variable, they did a lot of exploratory kind of work, some of the more basic kind of stats applications that we would use in nutrition research, when we've got smaller amounts of data to kind of pick up some early correlations, without maybe looking at all of the collinearity between all of these different factors as well. So we know that lots of different things can be independent and dependent predictors. So they did a kind of exploratory analysis of this, and then developed some more stronger algorithms essentially. So they moved then that system of supervised learning to be able to then predict each person's unique glycemic response to food. They had 800 people to take part in the study, and what they what they showed from that was that they could really, really see, really distinguish differences, if a person, if two different people ate the same food, they could have totally different glycemic response. So there's a lot of variability there, and they use decision trees, decision tree and machine learning model. So they ingested all of those data points, which we wouldn't be able to do with our traditional research approaches using traditional statistics. So they identified 137 different factors that were able to predict glycemic response to a specific food for each person.

So what they kind of said, I suppose, [inaudible 01:01:29] was potentially over a hundred different things that could determine why you're responding that way, and to certain foods. And then, once they have that algorithm developed, again, going back to the importance of kind of validation, they then – so this was their training essentially, and then they went

out and recruited a test at another hundred people, and they validated their algorithm against this test group. And then to take the validation one step further, they then conducted a randomized controlled trial. So it's a smaller group, this is very much the proof of concept study, where they use these personalized diets based on the algorithm findings, and these findings showed a significant improvement for postprandial glucose response derived from the group who were given the diet that was predicted by this machine learning model compared to a controlled group.

So the algorithm was accurate at predicting the glycemic response, and it outperformed the predictions of what we would be able to get from a traditional statistics type of analysis. So I think, yeah, the study is really, it's an amazing paper. I think everyone should go and have a look at it. It’s quite extensive. They've done – they've been so thorough, but the study just essentially highlighted the significant variance in those individual responses, and how we currently provide guidelines around diet and different diseases, might not necessarily meet the requirements for everybody, because for lots of different factors, I think, particularly, they focus a lot on the gut microbiome. But this was like a key driver of how somebody might respond to a certain meal, and how it differed so greatly between people. So yeah, I think it's a really exciting application. I know he spoke as well on the podcast that in terms of actually translating that into people's ability to use this kind of software that there's a company developed afterwards where you could download this health app, and you could use, based off their algorithm, I think you could send some samples like a stool sample for gut microbiome analysis, and you potentially have to wear a CGM monitor just to look at your glucose response. And then once they have that baseline information, they can use their algorithm to decide which foods you'd respond to or not. But then, that hasn't been tested, it's

not a validated app, and it sort of hasn't been tested, if it's any better than if you just followed the normal recommendations of what a good diet might be.

DANNY LENNON: Right. And I think this is kind of one of the big things of like how this translates to actual end health outcomes versus, say, just glycemic response and so on. And I know this is almost not only kind of a pragmatic question, but maybe even a philosophical question to some extent, and I think, Alan, you and I kind of talked about this, maybe in relation to PREDICT more specifically, but the question of when it comes to personalized nutrition and this idea of looking at individual responses, and thinking about then what way the diet should look, and kind of maybe a question without a direct answer, but like, how personalized does personalize need to be, as in, like, what is the net benefit going to be beyond recommended dietary guidelines if someone's following those versus going this step further, and getting something that's personalized in terms of meal response. And yeah, at what point is personalized enough, I guess, is the kind of philosophical question to consider here.

ALAN FLANAGAN: Yeah, so I think that there's a number of potential issues that arise, some kind of philosophical, some ethical, potentially, and also, some just practical. And then, one that I have actually that is a question that I have marked here that I'll ask Niamh then at the end, because it relates to this idea of the use of machine learning for predictions, but what some of what I did wider reading around ML, unrelated to nutrition and coming across some critiques from AI researchers and stuff like that. But I think at the high level, the personalized nutrition question is, is nothing new, necessarily, although ML now adds in a different dimension to the potential for it to go somewhere. And I've got incredible research from that group, I mean, that paper is so rich in detail and so thorough in their comparisons with the data that they had with actual or prior

knowledge. So they weren't just saying, look, this outcome, this prediction happened, and the prediction's right, but, you know, as an example, their algorithm showed that as the total fat content or the fat to carbohydrate, as the fat content of a meal and carbohydrate content of a meal decreased, then you would get less of a postprandial glucose response, and that they had kind of checked that against wider research, it's an incredibly thorough paper.

But the concept of personalized nutrition is not necessarily new, it started to emerge in the late 2000s, particularly as single nucleotide polymorphisms or SNPs for type 2 diabetes became recognized. Its genesis probably, arguably, could go back to the 90s with the discovery of the leptin gene, and the identification of a congenital leptin deficiency; and then we also have other examples of the application of personalized nutrition with, for example, the MTHFR SNP, which influences folate requirements. I think there's a couple of levels that we could think about this getting sticky. One is that, for the most part, when you look at how these things tend to go, what I noticed, whether it's that study, the Israeli group, and immediately, there's a company being set up, and you can send them your poo – predict, oh hey, here's this app, and we'll, you know, all the problems are solved. And that's a feature of tech. That's not anything to do with nutrition. This is a characteristic of the tech revolution, where immediately it is capitalized upon in a kind of commercial context, often racing ahead of where we're actually at with that data.

There were companies in San Francisco, for example, that started doing genetic testing, and then providing macronutrient tailored meals based on apparently your genetic profile. This clearly is something that is only ever going to be accessed by the wealthy and healthy seven-figure Silicon Valley crew. And I think there's a couple of levels of personalized nutrition that

creates some potential, just barriers, or the fact that it may be just as putting the cart before the horse. One is that nutrition science to date, has gotten some really actionable answers, independent of any criticisms that want to be leveled at any level of research, whether it's epidemiology or interventions. Replacing saturated fat with polyunsaturated and unsaturated fats generally is likely to reduce cardiovascular risk. Most people who do that are likely to improve their overall cardiovascular and metabolic health. Increasing dietary fiber is a good idea across the board. More fruits and vegetables is better than less. We've got these answers. And these answers aren't implemented. We have guidelines, they've never really been followed.

Recent national diet nutrition survey data in the UK puts average fruit and vegetable intake at 261 grams a day. When people are so far off targets, the idea that we're worrying about their blood glucose response to a banana when we need to get them eating a banana in the first place, is, in my opinion, getting really ahead of the game as far as population health goes. I think there's incredible potential clinical utility for this in the management of a metabolic disease, particularly with CGMs, which can feed into a team of dieticians in a hospital and they can monitor this, and this can be really tailored. The idea that this is the future of improving the burden of lifestyle disease for which diet is a major driver at a population level, I've never seen it, and I still don't see it. This is something that I think has tremendous potential individual clinical application, but for the most part, we know enough – there is sufficient current knowledge to improve and dramatically shift the burden of disease in the population. The reason that that hasn't happened is not a knowledge problem, we're not in a knowledge deficit with nutrition and the application at a population level. We're facing political barriers, social and economic ideologies that feed into political, particularly

in the UK, where you have a particular narrative of personal responsibility.

My worry about personalized nutrition is the idea that essentially it inadvertently and through no fault of anyone in that research area, bolsters the narrative that this is all down to the individual. Well, if I can tell you what's in your poo, and how you respond to an apple, well, I can give you all this information, and you have to act on it; and if you don't, it's your fault. That would feed right in and fit right at home with, for example, Conservative Party policy in relation to public health since 2010. And so for me, they're kind of ethical considerations with the application of it. And do we really need this to actually achieve those health outcomes? So, as an example, one of the studies that I find really interesting in this is the Food4Me study, and when you read this study, it reads like following personalized nutrition advice leads to all of these outcomes compared to just your standard advice. But interestingly, the personalized advice was more effective, and the conclusion is that this is better than our standard public health recommendations or standard recommendations. But I wonder whether it's a placebo for those recommendations. When you look at what the personalized recommendations were in the Food4Me study, lower saturated fat, lower salt, lower red meat, more polyunsaturated fats, more fruits and vegetables, every single one of them is a pillar of public health recommendations. And when you look at the actual change, 14% energy, 14.6% energy saturated fat, that was changed to 13.5%. Do we need personalized nutrition to get people to make a 1% change in dietary saturated fat content, if telling them this is super personalized, has some sort of effect on their psyche in relation to their diet, hey, this is your [inaudible 01:12:32] special little, yeah, if it's an adherence thing, this is, you're so special, I need you to reduce red meat intake, or, I need you to eat more vegetables. You're so

special, I need you to eat more vegetables. Okay?

I find the changes in the Food4Me study are so minuscule, and each of them entirely in line with current public health recommendations, that I find it interesting to say that we're superseding current guidelines, all that Food4Me study did was get people to do current guidelines through another means of intervention, so to speak. So I think there is huge potential application for this in a clinical context. It's never been to me, and I remain to be persuaded that it's a solution to population health approaches. We have sufficient knowledge to do that, you know, machine learning could give us incredible predictive power in relation to metabolic and cardiovascular disease, but it's not going to reduce the number of people needing to go to food banks in the UK, it's not going to help people make simple diet swaps like the [inaudible 01:13:42] study did. Yeah, I think personalized nutrition has the potential to be politically hijacked as well in the context of the prevailing personal responsibility narrative that abounds particularly in kind of very neoliberal societies like the US and the UK. And I worry that ultimately, because of the sophistication of the requirements, and because of what's involved, it will ultimately be something that just adds to health disparity, because the only people that are going to have the means to access this are people of means. So adding to some of the kind of ethical hurdles that I see with this, unless it was some sort of kind of nationalized program, in which case you're asking the state to provide subsidies, and again, a lot of governments just aren't going to be interested in that.

So I have some deep reservations from a philosophical and ethical standpoint, I guess, and from a pragmatic standpoint, and just practical standpoint. Like I said, I feel like we have enough current knowledge and there's so much low hanging fruit to pick, like I said, the

net health gain for someone going in the UK population right now with their average fruit and vegetable intake of doubling that and getting their say, for example, saturated fat down from the 13% energy that it is now to 8%, like, if those substitutions were made, the magnitude of risk reduction in cardiometabolic disease in the population would be seismic. And we know that from modeling, so I think there're still those low hanging fruit to be picked, and to pick them requires removing these socioeconomic and political barriers to that change, it's not a knowledge deficit.

DANNY LENNON: Yeah, and, I guess the aspect you mentioned earlier Niamh around just the sheer amount of data that we're all creating right now, and the access that people can have to that, and therefore, I wonder that double edged sword that comes with a, that the speed at which tech moves, and we now have the ability for someone to be able to create a product or a piece of tech or an algorithm that is going to move much faster than what we might even know about certain disease states, let's say, and how they correlate with certain nutrition changes, and there's just like a disconnect between the speed at which research is going to move versus how the tech is going to move to the point where you're getting things, like Alan mentioned, like consumer testing that may tell you, it can prescribe certain diets beyond that. So I think that's one aspect to what you just mentioned, I think. And the second that I was going to bring up as well, I'd like to hear your thoughts is around CGM specifically, so continuous glucose monitoring. And while this is a really useful application, at least, in health science researchers seeing that in the context of people with diabetes, whereas now it's become much more widespread for people to start getting hold of CGM monitors healthy people start using them, and then trying to interpret that to think that by changing my blood glucose response through constantly monitoring this, that's going to have a clear end benefit for my health. And then try to think through that

question of, well, does using a CGM for an otherwise healthy person, does that lead to a net positive health impact of like actual true endpoints that we care about or not, and I think those are interesting questions to consider, because we have so much access to tech and our own individual data from an individual perspective. So that's kind of one aspect, and then the second is, like Alan mentioned then on a population wide level, then it's a different discussion as well. But yeah, I think there's a lot in there, so I don't know if you have any particular points you wanted to revisit Niamh or any notes in particular you wanted to touch on?

NIAMH ASPELL: Yeah, no, it's all really, yeah, and I totally agree with pretty much everything you said, Alan. And I think, to go back to the Food4Me study, I think one really interesting conclusion of that study was it's very much still presented as personalized nutrition is definitely a better kind of approach, and we demonstrated that and it works compared to standard population advice. But then, it also ends with, but phenotype and genotype didn't make any difference whatsoever. So when I think of personalized nutrition, it's what they looked at it was based on weight, physical activity, dietary intake. So it's like, okay, the stuff that, okay...

ALAN FLANAGAN: We know.

NIAMH ASPELL: We know it's the gold standard. The stuff can tick the box. The conclusion of that study is it comes across to me reading it that it's very much so that personalized nutrition is the way forward, but when I think of personalized nutrition at this level, particularly with advanced statistics and machine learning, I'm thinking about right, well, what did my genome or phenotype say about me, and is that why personalized nutrition is going to work, and that's the main conclusion from the study is actually that didn't have any effect whatsoever. So yeah, I think that's a really good study to

kind of consider in that, and then, all of the ethical things so important, there's so many, so many different things to factor, right? Access, I think, is a really, really important one. And then, if these technologies are developed, they are usually developed through the elite, or the people who can afford them, and what's the coverage going to be like for the population in general. But then, you have to think around training, so in terms of, do you understand what your personalized diet is telling you, and your personalized diet is based on, you know, with that previous [inaudible 01:19:25] it's very much so focused on, like you say, the poo sample that you send off, that's on one day, we know how much your microbiome can change. So if that's changing in a month's time, you're not sending another sample, you're constantly working on the algorithm that was based on your gut microbiome on that day when you sent it. So it's not giving you – it's giving you a personalized diet for that moment in time, but is it, you know, are you constantly, so then, do you sign up for this lifetime subscription of posting your biological materials to a company and they tell you what to eat and then it complicates things, and it puts, again, that kind of responsibility on the person, it puts a lot of pressure on the person of, oh I can have this, but I can't have that, and this needs to change, and there's a lot of kind of additional health kind of stress around managing kind of your life, when it doesn't need to be that difficult, but sometimes, people want a solution, and tech developers like to provide what they see is these kind of silver bullet solution.

So I think we need to kind of go back and put a lot of that responsibility back on the tech developers as well, because, like you said, [inaudible 01:20:29] they have the technology, but they're not really so interested in, you know, not all, but in a lot of cases, they might not be interested in how advanced the domain is, and they'll say, oh well, this paper says this, this, this, so we'll put this into an app, and they can use it. But we need to have better

collaboration with tech developers and domain experts, the researchers, to work together. So I don't think machine learning or AI or tech development can go off in one direction by itself, I think it can complement each other; and we could come up with kind of hybrid solutions of, well, if we knew a little bit more about this complex issue, this could maybe help our research and the tech need to be involved and kind of get a better understanding of how we capture disease states, and how we look at demographic variability. And yeah, there's lots of ethical issues like biases and the datasets that we're using; there's lots of issues around open data sources, and if they're incomplete or biased than the decisions that are being made, and it's the big thing in healthcare is they sometimes are inaccurate, or they're based on populations that the clinician is not actually currently seeing. So you might have a lot of data might be collected, and once a subgroup of the population, it could be entirely Caucasian population, and then you've got a doctor who's looking at different ethnic groups, and they're making decisions on people with totally different profiles, and the same decisions aren't applicable. The clinician needs to understand how the decision was made. Most of these systems are built with this black box where the algorithms are created, even the tech developers don't know why it made a decision, they just know it made a decision. And then the clinician or the dietitian doesn't know why the computer made that decision, but it's telling them, you need to tell this patient to reduce whatever or to increase whatever. But they can't also tell the patient how they came to that conclusion. And if their conclusion is different than what they would have came to themselves based on their training or experience, who's then liable for making the right decision? Or do you trust the computer system? Do I trust my judgment?

So there's a lot, and there is a lot of work in that area at the moment in terms of its it's called Explainable AI, so trying to understand

how these systems are making decisions, what features in these machine learning systems are making a decision. And then the interpretability of that, so how much can I understand that, how much does the patient or the person understand it as well. So there's lots of ethical things to consider in that in terms of kind of transparency and reporting, and it's this methodology, ethics by design that tech companies should really engage in where they have ethicist or people who aren't just patient advocates or who work with the tech developers to kind of flag, well, no, you're missing these issues here, and we need to address these before it's marketable. Whereas we know a lot of companies, and particularly with like nutrigenomics, that there's not strong evidence there, but there's lots of places where you can send your specimens or samples and get an idea of your profile and what you should [inaudible 01:23:34] and they're not really, they're not validated. So there's definitely lots of things that still need to be addressed on that side of things. And I can't remember your last question, Danny.

ALAN FLANAGAN: Niamh, there was a couple of notes that I had from kind of not nutrition specific, just wider reading around just machine learning and data science generally, and some interesting critiques, and you mentioned there that idea of the black box and there's a guy I came across, it was an article in science, but I think his name is Ali Rahimi, he's an AI investigating, or an AI researcher in Google, but he basically published a paper pointing out a lot of the major issues with the – and he describes machine learning as alchemy in the paper. But one of the points he was putting was that this idea that the system has this black box. You put in this input, and you get an output, and no one knows what has happened in between, and I believe that him and a group of others are kind of working on trying to actually provide explanations for that. But a couple of things that I thought were quite interesting about how essentially these – the algorithms are trained

first and foremost to actually find something, so an algorithm is never going to come back and say, oh yeah, we searched through this massive genomic database and didn't find anything. And one of the critiques that I came across, that was quite forceful was that if we're assuming that this will solve a reproducibility crisis in science from current methods, there is its own reproducibility issue within machine learning, because algorithms can zero in on something, they're going to always find something, and what they're zeroing in on could be noise that's not necessarily reproducible.

And I wanted to get your thoughts on one example that I thought of where this might be, this idea of reproducibility was in the Israeli study, we were discussing. They found that the microbiome was a massive factor predicting postprandial glucose responses, right? And so, there's a company set up now, and you send your poop off to it. But then you look at PREDICT, and the microbiome had minimal contribution to predicting postprandial blood glucose response. I think it was like 6.5%, and that was the same for insulin and triglycerides, nothing was higher than about 7.5%. So already, I look at those two studies, and there's a disconnect in that kind of reproducibility with that element. And so, it made me wonder, is there the potential that the findings that they're getting reflect the algorithm in a way that they've built, based on the fact that the algorithm will absolutely find something, and you are always going to end up coming out with some sort of correlations. And yes, you could validate that within your group. But there's very much, I guess, a dependency on the algorithm to whatever algorithm you use, and then the assumption that any one algorithm, it produces the same results across the board. It just struck me that there was so much discourse from people in the field unrelated to nutrition about how there's a major reproducibility problem with AI, and with algorithms in particular, and I just thought

about that in relation to the PNP and the PREDICT study, and then, how they came to such really divergent conclusions in relation to the contribution of the microbiome to postprandial glucose, and it made me think, maybe that's an algorithmic issue. But I don't know what your thoughts are on that, or the kind of reproducibility of these algorithms or the fact that they're just designed to go in and find, well, anything.

NIAMH ASPELL: Yeah, I think it's really good point, and it goes back to the same issues that we have with standards and statistical kind of practice. I think one of the benefits of machine learning is they can probably have a more rigorous investigation in terms of the correlation between more variables and different factors. So it can maybe enhance some of the regression based kind of analysis that we can currently do. But as you say, if you have enough data, you're going to find associations with different things. It's really interesting that they got microbiome [inaudible 01:28:21] in the PREDICT study, and it'd be interesting. I'd assume that they were using very similar profiles of gut microbiome that they were analyzing, I'm not particularly sure if they had looked at the same kind of species, but I'd imagine that they probably did, or they'd find something quite close anyway. I think, yeah, there's lots to still, the techniques in machine learning are not perfect, and they definitely come with their flaws. And if we think of that, going back to the black box, and the decision making, a lot of these decision making tools are automated, so you've trained them, and then you keep showing them data, and they make adjustments based on new data that they see as well. So they're constantly evolving, and they'll change what decision they make, based on their most recently seen kind of data. So they're constantly learning as well.

But how do we control that, or, how do we understand that, and how do we kind of then understand, say, if you had a patient on day

one, and we're using the same system it and makes a decision, but it makes a slightly different decision a year later as this system kind of evolves? They're making a decision based on different variables because they're putting different weight on the different features in that dataset to make that decision? And at present, most of these systems, they don't understand why the model is making that decision, and what features is it picking, what kind of parts of the data is it picking to make that decision, and we can't then, you know, it's a person's right to understand why we're making a medical or a health related decision about them for them to then go and seek treatment or to intervene in some way. At the moment, we can't really tell them. So we kind of need to overcome that with these new – so Explainable AI is a machine learning technique, and it needs to be integrated into all of these systems, essentially, it's just pulling apart the black box, and we don't have this black box system anymore. And it provides the user then with information around how their decisions were made. And this has been extended, it's currently being written into European law that any of these healthcare systems or any systems that are developed or have this Explainable AI component built in, so it's kind of the total opposite of the black box, essentially.

So it will show you all of the potential biases in your data, so what your data isn't telling you and what it's not actually telling you as well, which is really, really important. I think topic could also enhance some of the statistical, traditional kind of techniques as we get datasets, and when I say we, I say the – some of the general research population, and kind of, they pull datasets, they have a look at them, they take out some variables that don't look important, then they have some that there may be really interested in, they keep those in, they take a few others. We don't get access to that when it's published. We sometimes get access to the dataset that they use to make the

decision. But we don't have this version control system, which is usually there in machine learning, where we can show the stages of the model builder, the algorithm's development, very much so in a version controlled kind of standardized way, and kind of compliance governance around, I suppose, some of how we analyze certain studies. And I think that needs to, that kind of process in data science needs to be incorporated into statistical analysis and nutrition research as well. So it's kind of focusing again on that transparency. So if a dataset with a million data points that you've pulled down, but you've published a study on it, on the dataset, I think it's seen in that study that was published a couple of years ago around everything is correlated or related to cancer. And it's like, they looked at, or, I'm thinking sorry of the redacted study, the PREDIMED study where they looked at all the secondary papers that were published on that, and they've seen that the population sizes for the same cohorts were different in so many different papers that were published. So they didn't describe in those papers, well, who did you remove my data points that you removed, because that's really important in how you came to your decision. So yeah, I think transparency around AI and machine learning needs to be incorporated, and it will be, it's going to be a requirement for any of these systems that are developed, we'll have to incorporate that, and I think that practice should be brought into nutrition research as well. So there's a little bit more kind of responsibility on the researcher to be very open about their analysis.

ALAN FLANAGAN: A 100%, I think that in many respects, with some of these conversations that we're having about ML or traditional methods. It's like actually, at a core level, at a level of principle, the potential issues are often quite similar, if not the same. So transparency, we talk of the open access revolution, currently, in science, where there's a massive impetus on researchers to make their data available, and this is all

great, it's all a very needed step. And I think about that a lot in relation to my prior life in law, where you have that phrase, justice must not only be done, but must be seen to be done. And I've always felt the same about science – science, if it's going to have an impact in the real world, it has to be seen to be done, it has to be transparent. And ultimately, whether a machine is deriving predictions or traditional methods of statistical analysis is deriving the prediction or the analysis, it's still a human endeavor, as we still need to take what that finding is as a statistic, so to speak, or as an outcome, or as a point estimate. And it's ultimately, people that need to then decide what that means for the real world, what does this mean in application. The process of causal inference and deriving causal conclusions will always be still, again, a human endeavor in terms of taking this data and synthesizing it into something that actually has meaning in the real world. So I don't think that process is going to go away necessarily, and I think that it's likely something where it's going to be still that principle of converging lines of evidence from different lines of inquiry, and perhaps for smaller datasets or smaller sample size research, a traditional multivariate regression analysis would be sufficient. But you're getting up to kind of genomic or metabolomics, and suddenly, that's really not going to be, because you could be leaving important predictors out, and machine learning provides the ability then to actually take a sophisticated approach to explaining that available data.

One question, I think that I also kind of had from some of the kind of critiques in relation to ML that I had had read, one interesting example, the point being made was that, ultimately, the algorithms are going to identify correlations, but they'll assume that – and I don't know this as I'm kind of paraphrasing the critique – they'll assume that that correlation is some form of kind of causal relationship, rather than – so the example in the article was if you put in marriages in the State of Kentucky, and

deaths from people drowning, falling over the side of a boat, they would correlate so strongly that the algorithm would pick up on this relationship, and it would tell you these two are. But it still requires an appropriate interpretation on the back end by someone to not just look, oh well, the machine gave me this, therefore getting married in Kentucky means you're going to die falling over the side of a boat and drowning. So it was just something that struck me...

NIAMH ASPELL: You can't say for sure if that's not accurate though.

ALAN FLANAGAN: Yeah. And so, is that a correct understanding of how the algorithms operate in this regard, particularly in relation to deep neural networks. One of the critiques that I came across is because they have such a plethora of parameters. If there's a lack of algorithmic knowledge, like, any kind of domain knowledge, you can kind of end up with spurious results, but because the algorithm gave them to you, you're just like, aha, publish these results, no matter, you know, without taking that kind of extra step, and is there something that we need to think about in that regard, if the algorithms are designed to be predictive and predicting is what they're good at, is there the potential that they can detect spurious correlations.

NIAMH ASPELL: Yeah, of course, it's exactly the same challenges we have in nutrition research and the traditional methods that we use. But I thought domain expertise, you need to give the tech developers the direction that you base your hypotheses on. You need to say, this is what we need to look at. If there's additional things that come out, we need you to kind of come back, show us those, we'll think about them, did they make practical sense, go off and explore that. And you kind of, again, provide kind of direction around, okay, the next steps are, but that's the really, really important thing here is that you can give – anybody can take a dataset,

if I was to take a dataset on, I don't know, financial figures, and you told me to look for a correlation. I could find loads of things, but they might mean absolutely nothing at all, because I have no understanding about anything in that area. So I think, yeah, the collaboration and the domain expertise need to be incorporated into the tech development. We can create anything if we want to create something. But is it useful, and does it make sense? And if it doesn't, then it's not – it's again, it's the whole thing of bad data in bad iPods. It's like, it's the same principle.

DANNY LENNON: Cool. So before we maybe get to any concluding remarks, is there anything that we have forgotten to address, or, is there any points that any of you guys think is particularly important to bring up or touch on that we haven't got to yet? Otherwise, we can move to maybe some kind of conclusions, I guess, and start wrapping up.

NIAMH ASPELL: I think that's everything that I kind of have, although if I, yeah, scribbling down here [inaudible 01:39:11] it's been a really enjoyable conversation.

DANNY LENNON: Yeah. So with that maybe to leave people with some kind of clear things to take away, on any of the things we've said, what are some of the maybe couple of things you would like to refresh people's memory on and leave them with going away from this conversation – and maybe I'll start with you Niamh, what are a couple of important things you would most like people to remember from this conversation?

NIAMH ASPELL: Yeah, I think it's really important just to highlight that we probably will start seeing machine learning being incorporated a lot more into nutrition research, but we need to maybe not think of machine learning as an AI, as this kind of silver bullet or this kind of magic pill. I think there's the same limitations with machine learning that we see as well in statistical and traditional statistical

approaches, but we've had a lot longer to create solutions for traditional approaches to overcome that, and to test that, and to kind of validate how we can kind of offset some of those limitations. Machine learning is very much still so driven from the technical side of things, a lot of the publications I looked at were in computational researchers or kind of computer science and departments within different universities. So I'd like to see, if you're looking at any of these papers, I think, go and look at the affiliations, and see if there's also people who are, if it's a paper on cancer and nutrition research, make sure there's some oncologists involved, and make sure there's some nutrition people involved in that paper as well. And it's not just entirely from a computational perspective, so just to kind of take that as well. And I think it will be really interesting for people to become a bit more familiar with machine learning techniques, because, not so much in the application, but just the principles of them, because I think it will start to come up more and more. So it's something that we probably can't just ignore, and I think people should maybe consider how we could maybe use some of them to support traditional techniques. I don't think that they will necessarily overtake nutrition research and traditional kind of practices.

DANNY LENNON: Fantastic. And Alan, any particular points that you'd want to most impress upon people before we wrap up?

ALAN FLANAGAN: Yeah, I agree with everything Niamh said, like, it's a really exciting area. But like any new methods, it's important that we don't kind of fall to the, I guess, tech halo and just assume, oh, it's tech, so it's all fine, and we apply the same rigor to validation and everything that we would otherwise. And then I think it's, again, for me, science is a very human endeavor, what is it we're trying to achieve here, where is this going to be applied, it could give us all the answers in the world, it could give us all the accuracy in the world. But it may not

necessarily mean that it's going to be useful for application for addressing or removing some of the major barriers that are extant in our ability to actually make a difference in the burden of disease as it relates to diet in the population. So it'll still come back to taking data and having to have that human element to its interpretation, and the consideration of what it means and how we go apply, how we go about applying it in the real world to actually make a difference with a lot of the issues that we currently face.

DANNY LENNON: Fantastic.

[01:42:40]