Sarah: Welcome to Proto, a podcast that explores the frontiers of medicine. I'm Sarah [Algier 00:00:07].

Dr. Jordan Smol...: And I'm Dr. Jordan Smoller. Since the turn of the millennium, the suicide rate in the United States has increased by nearly 30% turning this tide is a top priority for psychiatrists today.

Sarah: Can we find a solution to this very human problem with the help of machine intelligence? The government of Canada is asking one artificial intelligence company to look at how people behave online in suicide hotspots in the hope of better identifying dangerous patterns in the future.

Kenton White: We can look and see what kinds of conversations were happening prior to the suicides to help our AI learn and understand what might be precursors and how to pick up on those.

Dr. Jordan Smol...: And we'll hear about an AI project here in Massachusetts that could help us spot patients at high risk and alert their physicians before suicide attempt happens.

Sarah: Coming up on this episode of the Proto Podcast brought to you by Massachusetts General Hospital. Each episode, we explore a topic from the frontiers of medicine. To help us along, we are joined by an expert from MGH. My co-host today is Dr. Jordan Smoller. He's the director of the psychiatric and neurodevelopmental genetics unit at the Center for Genomic Medicine. Dr. Smoller is well known for his work in tracing psychiatric disorders to their roots in our DNA. It's the type of painstaking work that involves sifting through vast databases of genetic and patient data. But patient data can do more than unravel the causes of disease, he says. In the right hands, it can actually predict problems with impressive accuracy. Dr. Smoller, welcome.

Dr. Jordan Smol...: Thank you, Sarah. It's really a pleasure to be here today. I'm very excited about our show.

Sarah: So let's talk first about how clinicians traditionally have tried to figure out whether a patient is at risk.

Dr. Jordan Smol...: Well, traditionally assessing whether a patient is at risk of self harm or suicidal behavior is one of the most difficult challenges that we face, but it's a public health emergency. There are more than a million and a half people in this country every year who attempt suicide. More than 45,000 die by suicide, which is an extraordinary number. And it is the second leading cause of death among young people in this country. So risk assessment right now in a healthcare setting is largely based on clinical judgment and also asking people essentially, if they are having suicidal thoughts. We have some general risk factors, things like substance abuse, other mental health conditions, certain symptoms like anxiety, access to lethal weapons, those kinds of things. But generally we're asking people and making a judgment about whether we think they're at risk.

Sarah: So what are the weaknesses with the current clinical approach?

Dr. Jordan Smol...: Well, the weaknesses have to do with the fact that we really have a limited set of strong predictors that clinicians can take advantage of. And as I said, mostly, we're asking people to self report about how they're feeling. Of course, some people may be motivated not to be upfront about their intentions, because if they are really intending to do something, they don't want an intervention. We know that the clinical risk assessment tools that we have now are not doing the job. And in fact, a colleague of mine, Matt Nock, at Harvard University and his colleagues did a heroic systematic review of all the research on suicide risk prediction in the last 50 years. They looked at 365 different studies, and they basically said, what do we know? And the answer was sobering. There was no category of predictors really that was doing much better than chance.

Sarah: So that brings us to your work. In 2016, in a paper you proposed a new model for predicting suicide. What is that new model?

Dr. Jordan Smol...: Well, we started this research with really three things in mind. One, is that clinicians don't currently have a good way of integrating all the information that may be relevant to predicting risk. Two, is that the majority of people are actually seen in healthcare settings so it does provide an important opportunity. And three, that we now have a kind of enormous and ever-growing repository of health-related data in the form of electronic health records that we thought might be able to help advance this problem. And so we asked, can we move the field forward by using the big data of electronic health records to develop a risk prediction tool that could inform clinical care? And to do that, we turn to machine learning techniques.

Sarah: So AI tends to be this catch all name for machines looking at complex problems and coming up with their own solutions. How did AI exactly work in this case?

Dr. Jordan Smol...: Well, our method uses an approach that is called supervised machine learning, which is a fancy term. But what that really means is that we begin by identifying a set of instances or cases of the thing that we're trying to predict and a set of controls who don't have that outcome of interest, in this case, it's suicidal behavior. And then we essentially train a computer algorithm or machine to identify which patterns of predictors best distinguish those two groups of individuals. So we're giving it kind of the labels and saying now with all the data, you have, find the best patterns of predictors that will distinguish those two. And once we trained the model, we found that it was able detect 45% of all suicide attempts with 90% specificity, an average of two or three years in advance.

Sarah: So when you fed all of this information into the computer, you discovered some factors that were to be expected like substance abuse and in certain psychiatric disorders, but others were a surprise to you.

Dr. Jordan Smol...: That's right. And this is kind of the interesting potential power of machine learning approaches and the AI that it can find relationships and predictive patterns that no human being could discover or would think of through intuition or judgment, or even looking at some set of data. So it found things like, as you suggested, thought to be common risk factors, things like substance abuse, diagnoses, depression, and so on.

Sarah: Right.

Dr. Jordan Smol...: But then there were other things that it detected as being relatively strong predictors that no clinician, I think, would be thinking about. And those are things like certain kinds of infections, osteomyelitis, hepatitis C, certain kinds of fractures, rib fractures, and so on. And an important point here is what we're really after is a model that does best in prediction, not necessarily explaining the risk factors and their relationship to the outcome. What I mean by that is there might be risk factors here that don't make a whole lot of sense to us as we look at them just at face value. But what the model is telling us is that they are somehow tagging or correlated with other risk factors, maybe some that we're not seeing, and therefore they're conveying themselves some degree of risk. So it's not saying that they're all causally related to suicide risk, but they are picking up this predictive signal.

Sarah: Oh. Interesting. If one day you're able to get an alert within a patient record saying that a person is at an increased risk of an attempt in the next few years, what happens next?

Dr. Jordan Smol...: So that is kind of a part of the end game of this is to develop an app essentially that could be integrated into the electronic health record. And we're working on prototypes of that at this point. So that at the point of care, a clinician, seeing a patient could be alerted to some degree of higher risk. And it's important to recognize that we, as I say, don't see this as a substitute for clinical judgment, but as a tool to inform clinical decision-making. So it's really kind of a positive screen that alerts the physician this is somebody who really might require further evaluation, might be somebody you need to think carefully about in terms of their risk, ask the questions if you're not already asking them, and so on.

So what we're trying to do is sort of bend the curve of this rising rate of suicide by using a lever in our healthcare systems that could improve the specificity of how we assess risk in a very busy setting where people have typically limited time to assess all the things they need to assess. This is essentially a flag that might say, you know what? Pay a little bit of more attention here.

Sarah: So it sounds like you have a lot of work ahead.

Dr. Jordan Smol...: We do.

Sarah: Thank you, Dr. Smoller. Coming up, we'll hear about an AI with a political pedigree who is now working to find Canadians in need of mental health interventions.

Dr. Jordan Smol...: You're listening to the Proto Podcast, a production of Massachusetts General Hospital.

Sarah: Not all suicide counseling can happen in the physician's office. Some people may avoid seeing the doctor when they're feeling depressed or for countless other reasons. So prevention efforts have historically relied on public health outreach, publishing an ad with a suicide hotline number, for instance, in the hopes that it gets seen by someone who will use it. But if you want those efforts to have the most impact and save the most lives, where should you focus your efforts? Artificial intelligence can help on that front too, at least that is the hope driving a new project from the government of Canada. In February, 2018, it began an ambitious partnership with Advanced Symbolics, Inc, a leading AI firm in Ottawa. Together they will seek out groups of Canadians who might be dangerously depressed. Kenton White is the Chief Data Scientist of Advanced Symbolics and teaches computer science at the University of Ottawa. His company is working with publicly available data to identify groups of people who might be saved by a well-timed mental health intervention. Kenton, welcome.

Kenton White: Hi. Thanks, Sarah.

Sarah: All right. So, Kenton, can you introduce us to your AI? Your AI has a name. Right?

Kenton White: Yes. We call her Polly. And that comes from the work that we've done with her on many of the political elections that she's had some great successes in.

Sarah: So it sounds like she already has a number of serious accomplishments under her belt. Can you tell us about them?

Kenton White: So she's done a lot of work in the political forecasting area, and she's the only one that has gotten the Canadian federal election correct in 2015. And she called Brexit that that was going to be a leave in 2016. She called that 52% leave, 48% remain, which was bang on. And she also got the US election right. And I was really proud of her on that because she got the popular vote right. So she saw that Clinton would win the popular vote, but she saw also that Trump was going to come out ahead.

Sarah: No kidding.

Kenton White: And that was a really fabulous result.

Sarah: So what kind of information is she using to make these predictions?

Kenton White: Well, we create randomized samples of people online from all sorts of different sources, Twitter, Facebook, Reddit, public forums where people are discussing what is on their mind, and we're not looking at what everyone is saying. We have a sample of real people who are from all walks of life and all geographies. Typically we'll work with a sample of one to 200,000 people and we're measuring what they're saying and how they're reacting.

Sarah: And so now for this suicide project with the Canadian government, is she looking at similar kinds of data?

Kenton White: Well, she's learning from scratch on the suicide. When Polly starts, she doesn't know anything about the topic and we have to teach her. And so we give her things to read. We give her news articles. We give her clinical white papers. We give her lots and lots of reading material, which she can read very quickly, and reading it she now learns about that topic and she looks at what people are saying from all these different sources and then says, this is what I think. And we have a person, in our case on the suicide it would be a clinician, who looks at what she's saying and helps her understand if she's learning it right. Or maybe she needs to read something more or modify her approach on a particular aspect.

Sarah: Could you give an example of how a mental health professional might work with Polly?

Kenton White: Oh. Absolutely. This is something really exciting because this just happened in the past a couple of weeks. So Polly was starting to identify groups or people that she thought was at low risk. And so we consulted, we worked with a clinician and looking at what it is about these people that she's seeing. And the pattern that she was picking up was these people were active politically in their community. They were going online and talking about issues in their community that they felt we should oppose or we should support. And working with the clinicians, that was a real indicator that these people were engaged. They were engaged in their community and they were feeling empowered, that they have control over their lives, which is what was putting them from a clinical perspective at lower risk. So here was an example of where Polly was discovering something. And then the clinician was helping us really dial in what it was she was picking up on among individuals, people within the sample.

Sarah: Can you give us any kind of sense of how quickly she's doing her analysis?

Kenton White: So the bottleneck is us. We are the bottlenecks. She has to wait for us to come back and help her understand what is going on. She can read these white papers in a matter of seconds. She can build up a statistical model of the words and how they are co-located and correlated and apply this to over a trillion tweets or short social messages that we have to identify people that are talking about particular topics or issues. The bottleneck is that she then needs to wait for us to come back and review her work and to decide if we need to provide her more information to improve her accuracy and her learning.

Sarah: And so for these leading indicators that she finds, can you help emphasize why AI is key in this? Why public health folks and others, why they can't do it as effectively themselves?

Kenton White: There's a number of reasons. The first is just the volume of information that can be out there. A person can't read and see what hundreds of thousands of people are doing and talking about. An AI can. A second reason is we're human and we bring with it our own biases and our own preconceptions. With an AI, we can really focus on ensuring that she only sees what is factual, what's important, and leave those biases behind. And so she's not influenced by what's happened before, what biases she's bringing to the table. And this is, I think, really important when we're dealing with many, many hot topic, hot button issues like elections. That's part of the success of Polly on the election side or suicide.

Sarah: Oh. That's interesting. She's dispassionate.

Kenton White: Yeah. That's a good word.

Sarah: Are there any particular geographic hotspots that she's looking at?

Kenton White: One that is very recent is we had, unfortunately in Canada, a large number of suicides in our northern communities. Our northern communities tend to be indigenous peoples, first nations peoples. And last year we had a large concentration of tragedy, of suicides within these communities. While this is very tragic, it does provide us an excellent research opportunity because we can go in and we can look and see what kinds of conversations were happening prior to these suicides to help our AI learn and understand what might be precursors and how to pick up on those. So we're trying to turn a tragedy into something positive that can save more lives across Canada and the world.

Sarah: Thank you so much, Kenton.

Kenton White: Oh. Thank you, Sarah. It's been a pleasure.

Sarah: Kenton White is the Chief Data Scientist of Advanced Symbolics. I'm joined again by my co-host, Dr. Jordan Smoller of the psychiatric and neurodevelopmental genetics unit at MGH. Kenton White's effort is mining data from online forums, as we just heard. Your team has been working with data from digital health records. Are these the best assets we have for research psychiatrists? Or are there other big data sets just waiting to be mined?

Dr. Jordan Smol...: I think undoubtedly there are other opportunities down the road. For example, we know that suicide is partly heritable and the idea that sometime in the future, maybe not in the not too distant future, genomic information may be incorporated into our health care systems. That provides a potential opportunity to use genetic risk scores or genetic information to maybe augment some of these other approaches to assessing risk. And that's become an area of great interest. Another area of development and interest by many people is the notion of digital health using mobile sensors and real time data collection through things like smartphones. That includes things like having people fill out surveys on their phone, but also capturing some of the data that can perhaps even passively, without somebody having to fill something out, indicate changes in activity levels and in mood and behavior and so on. So there's a lot of interest in seeing how useful might that be.

Sarah: So to come back to medical records, are there any parts of that data you wish could be better? I guess what I'm asking is does what the clinician types into the computer always completely reflect what's happening during the visit?

Dr. Jordan Smol...: That's a very good question. And actually the big challenge in using electronic health record data is also its main strength, that is the data are collected from real-world clinical care and the notes and coding that emerges from that care, that makes them highly relevant to what's going on. But it also means they're not as clean and interpretable as what we might do in a study designed for research purposes. And we also know that clinicians have various habits of how they record things that may be different among clinicians and not everything gets recorded. And so that is clearly a limitation. One way that that may improve in the future is, for example, through the use of systematic assessments in the course of a visit. So for example, a number of healthcare systems now are using questionnaires that measure mood or suicide risk and can be sort of inputted into the electronic health record in a kind of a standardized way, but that still doesn't necessarily capture some of the nuances and the ebb and flow of a clinical visit, and that remains a challenge.

Sarah: So clearly brilliant scientists are working on this, like Kenton White, but do you find that there's a benefit that you are both a researcher looking into this and a clinician sort of seeing these patients yourself?

Dr. Jordan Smol...: I do. And actually in Kenton's description of the work he's doing, I think one of the things that jumps out is that clinicians are essential to that work as well. For those of us out there who may be worried that ultimately we're saying robots are going to be replacing all of us, that's comforting. I think. You know, the clinical judgment and the clinical familiarity with some of these nuances that really comes from interacting with patients, a long history of providing care, knowing what various things mean in a particular healthcare context has been essential to our work, I think. And I think that is true more broadly.

Sarah: I understand that there are other applications of this technology, including predicting domestic abuse. What other conditions are being looked at? And what do you hope to see?

Dr. Jordan Smol...: Well, actually the domestic abuse example was, in many ways, the motivation for what we've done since, and it came from the work of Ben Rice who's now a very close colleague who really has developed many of the methods that we've used in this work on suicide and is our close collaborator on that. And he had previously shown that you could use these kinds of approaches to predict whether somebody was more likely to be the victim of domestic abuse in advance. There are a lot of big questions in psychiatry that we don't have very good answers to. And for which this kind of approach might be something that really helps people do better. So for example, just to give examples of things that we've been working on, the example of undiagnosed psychotic illness. There are about a hundred thousand people at any given time in the country who have psychotic illness, but it has not been diagnosed, and certainly obviously hasn't been treated.

And it turns out that the longer the delay in that detection, the worse the outcome. People don't do as well. So again, having a system for helping clinicians attend to folks who may need help is potentially very valuable. We're looking now beginning to look at post-traumatic stress disorder. You know, most people in this country experience, at some point in their lives, a major traumatic event, but only about one in 10 will actually develop PTSD, which is an often disabling disorder. We have no way really of knowing who those folks might be and directing our efforts at prevention and intervention. And so we're using these kinds of approaches to see if we can make progress there.

The distinction between bipolar disorder and major depressive disorder, that is depression, is an important one for which again, clinicians have a big dilemma. It's hard to know whether somebody who's depressed has depression as we commonly recognize it or may have a bipolar depression. The reason that is that the treatment choices differ and treating somebody with bipolar disorder with simply an antidepressant might actually make things a little bit worse. So there are all kinds of these big questions that we have not had a way of approaching very effectively, in many cases, that we hope exploring these kinds of approaches might help.

Sarah: Dr. Smoller, thank you for joining us today. And, listeners, thank you for tuning in to the Proto Podcast.

Dr. Jordan Smol...: Today's podcast was produced by Emily Silver, Bradley Klein, and Jason Anthony.

Sarah: Thanks also to our technical directors, Adam Keller and Chelsea Andes. Subscribe to the Proto Podcast on iTunes and Stitcher and follow us on Facebook and Twitter. See you next time.