#312: Screening along the spectrum:  
The search for a genetic test for autism

This is Up Close, the research talk show from the University of Melbourne, Australia.

SHANE HUNTINGTON
I'm Dr Shane Huntington. Thanks for joining us. Human beings are social animals. We rely on language and the subtle social cues that accompany our words to communicate with each other. But for people with Autism Spectrum Disorder, or ASD for short, the simple acts of communicating and interacting with others in a social setting can be baffling or even terrifying. Currently ASD diagnosis is complex. Psychological assessments and interviews are combined with behavioural observations by parents and teachers and a multitude of other mental disorders need to be carefully ruled out. But we know from twin studies that there's a genetic component to ASD, so why don't we have a genetic test for this condition? Are behavioural observations really the best we can do for desperate parents seeking answers for the challenging behaviour in their children? Surely our extraordinary advances in genetics hint at effective DNA based tests. Today on Up Close we speak to a neuropsychiatrist and an electrical engineer about how we might one day test for ASD based on our genetics. Chris Pantelis is Professor of Neuropsychiatry and Scientific Director of the Melbourne Neuropsychiatry Centre at the University of Melbourne and Melbourne Health. Stan Skafidas is Professor of Neural Engineering at the Department of Electrical and Electronic Engineering; leads the Melbourne School of Engineering's research in nanoelectronics and is the Director of the Centre for Neural Engineering. Welcome to Up Close Stan and Chris.

STAN SKAFIDAS
Thank you.

CHRIS PANTELIS
Thank you.
SHANE HUNTINGTON
Chris, I might start with you. What sorts of tests are currently available to diagnose someone with Autism Spectrum Disorder?

CHRISTOS PANTELIS
So the diagnosis of Autism Spectrum Disorder relies very much on clinical observation. It requires careful considered observation of behaviour, social interaction and particularly looking at language and communication; also observations related to stereotype, the repetitive behaviours that many of these children manifest. The disorder is diagnosed early. The onset is before the age of three and it's the observation that children are not engaging, not socialising appropriately, that they're delayed in their language and that they may have stereotyped or repetitive behaviours. So very much the diagnosis is based on clinical observation at this point in time. Now as you rightly point out it is clear that there is a genetic component to this disorder. It runs in families. Those twins that are monozygotic have a high concordance, which means that if one twin has the disorder there's a high likelihood that the co-twin is also affected. This means that we should be able to examine the genetics of this disorder and see if we can come up with a test if you like that might help us in our clinical diagnosis.

SHANE HUNTINGTON
You mentioned we can look at children as young as three. It would seem difficult that you'd be able to extract the sort of behavioural anomalies that you're talking about at that age, given the wide variety of developmental speeds that we find out kids growing up with. Now some kids learn language very quick, others don't. How successful is it in terms of determining if a child is positive at age three?

CHRISTOS PANTELIS
Again a very good and I think the important thing here is that one needs to take account of the trajectory of development of any individual child. And often clinicians looking at these children will assess them over a lengthy period of time. The diagnosis might be suspected but may not be confirmed for a considerable period of time, perhaps a number of years. It depends on the severity of the presentation, the range of symptoms and how they're developing.

SHANE HUNTINGTON
You mentioned the possibility of genetic testing. It would seem that we have a genetic test for every second illness at the moment. There are a lot of new ones around, the most commonly known ones such as those for breast cancer and so forth. There is definitely a genetic component to this as you say from twin studies. Why is it that we don't have a genetics test at this point for autism?

CHRISTOS PANTELIS
A disorder like autism is not caused by one gene. It's likely to be many genes perhaps acting together, so it's a polygenic disorder and there are a number of conditions that fall into this kind of category, making it more difficult, more complex in order to work out what the genetics of such conditions are.
SHANE HUNTINGTON
Stan, with regards to the genetics it's common I suppose for people to use these genome-wide association studies. Is this the type of study that you used looking at the entire genome in order to determine whether there was a particular area of interest for autism?

STAN SKAFIDAS
What we did do is look at the whole genome or data available from approximately 5000 children and their parents. We did look at the genes or the single nucleotide polymorphisms, SNPs, and differences between cases and controls in that cohort. Now the problem with your standard GWAS is that you're trying to make an association on a SNP based level and in these individuals that have been genotyped there are many SNPs that have been measured. If you follow the traditional GWAS approach and you look at each one of these individually, because you've got multiple comparisons you quickly begin to lose power or the ability to categorically be able to state is this single SNP in some way related or relevant to a disorder.

SHANE HUNTINGTON
Now a SNP there being just a small piece of our DNA, is that right?

STAN SKAFIDAS
Right, so there are many billions of nucleotides that form up our DNA. If you're looking at one at a time you will lose power and statistical significance. The other underlying assumption is that many of these common variants are not associated with the disorder and instead what is actually happening is there are these very rare variants which contribute to the disorder. Another way that we can start interrogating that data to try to determine if it's really a case of multiple SNPs or genes interacting in order to give rise to this disorder and is it a case that some of these common variants are contributing to this disorder. So instead of following our standard GWAS approach what we decided to do is look at canonical cell signalling pathways, collections of these genes independently and try to ascertain whether these canonical pathways had some kind of association with the disorder. It allowed us to reduce our search space significantly and contend with these issues of multiple comparisons.

SHANE HUNTINGTON
That's one of the parts I'm a little unclear on, how you go from having these millions of pieces of DNA you can look at, down to a smaller group that you can interrogate. What is the process of eliminating all the ones you don't need and looking specifically at ones that are shall we say better candidates?

STAN SKAFIDAS
So Shane, essentially what we've done is looked at collections of these genes which form up canonical cell signalling pathways and then looked at the SNPs that we've measured in those genotypes associated with that particular pathway. The hypothesis that we're testing is whether a particular cell signalling pathway contributes or is in some way associated with the disorder and in doing so we're
looking at large collections of SNPs and genes simultaneously, in order to be able to say is this relevant to the disorder or not. We're looking at collections of roughly about 400 canonical pathways; it means that we're looking at doing 400 hypothesis tests instead of looking at one million SNPs independently.

SHANE HUNTINGTON
Chris, do we know once we've found some of these genetic markers in particular what they actually do? I mean I hear that they're somewhat associated with autism, but presumably they also do other things. Do we have an understanding of what they do?

CHRISTOS PANTELIS
So we in identifying the pathways and the genetic markers, the SNPs relevant to those pathways, it leads us to ask questions about how those pathways might be implicated in a disorder such as Autism Spectrum Disorder. And indeed the identification of such pathways might lead to an understanding of the pathophysiology or neurobiology of the disorder and that might then lead us to investigate how these genes are expressed, what happens in terms of that expression in these disorders. It might lead to new treatments that might be relevant.

SHANE HUNTINGTON
Have you found anything interesting to date in that regard? I mean as a neurobiologist you see these genes, they're doing certain jobs. Maybe those jobs have got nothing to do with our brain at all. Are there links you are finding that sort of suggest that yes, this is the right one?

CHRISTOS PANTELIS
Well I think there are two areas to discuss here and one is that the analysis has allowed us to look at genetic markers that not only contribute and increase the risk towards the disorder, but also we've identified SNPs that appear to be protective and reduce the risk. Understanding the balance between those two, I think, is an important way forward. Secondly the SNPs that we've identified seem to implicate certain pathways that may be relevant. Specifically, we've started to focus on the metabotropic glutamate receptor 5 which is implicated by the analyses that we've undertaken. We've undertaken now some post-mortem work and I'll ask Stan perhaps to comment further with regard to this particular pathway. I think it may lead us in new directions with regard to the disorder.

STAN SKAFIDAS
As Chris states, one of our primary candidates relates to metabotropic glutamate receptor 5, but a lot of the other SNPs that we identified reside on an important neuronal pathway, in particular the glutamatergic pathway. Glutamate neurons are important excitatory neurons in the brain. That has lead us to begin to look at can we biologically verify some of these results. Now of course with statistics and a lot of the approaches that we've undertaken, we're merely showing an association. As an engineer an association is interesting, but until you prove it experimentally it's merely an interesting hypothesis. In conjunction with Professor Ian Everall, Dr Gursharan
Chana as part of our group, they've undertaken a post-mortem study where we're looking at the expression of the metabotropic glutamate receptor 5 and genes on the glutamatergic pathway in this post-mortem tissue. And we're finding some very interesting results that in some way biologically provide some further evidence on the validity of many of the candidates that we've actually identified as part of this work.

SHANE HUNTINGTON
I'm Shane Huntington and you're listening to Up Close. We're discussing a new way of detecting autism with electrical engineer Stan Skafidas and neuropsychiatrist Chris Pantelis. Stan, how many of these genetic markers did you actually find were important in total?

STAN SKAFIDAS
In total 237 markers that were necessary to build this classifier. The reason why there are 237 and potentially there could be a much larger number was the way that we constrained our analysis. And that was by looking at the pathways, genes that reside on those pathways and known SNPs which reside on those genes. But of course as more genotypes become available, as these pathways become more elaborate and more genes and interactions are identified in these pathways, when the analysis is done in the future there might be further candidates that are identified.

SHANE HUNTINGTON
Let's just clarify exactly how this works because a lot of people would be aware with the breast cancer genes for example, you test positive for those genes, there's a pretty good chance you will get breast cancer because those genes are causing the problem. We're not talking about that here though are we? We're talking about something that gives you an identification of risk, is that right? And why is that? Why is there the difference? How is it that we're not finding the gene that's causing autism, just finding the risk genes, as it were?

STAN SKAFIDAS
There's an important difference. We believe that autism is a polygenic disorder, except for certain syndromic situations where you can clearly identify a gene or SNP that is relevant to that syndromic form of the disorder. Idiopathic autism is polygenic. You know in undertaking those statistical analysis the only thing you can actually say is well if I appropriately weight these collections of SNPs that I get a score which gives me some relevant information about the disorder. So it's merely an association at this moment in time. It's very different to breast cancer. Again just to reiterate, it's an association and we believe that a strong genetic component for autism is the highly complex interaction of common variants.

SHANE HUNTINGTON
Now, how do you actually go about testing whether these particular markers actually mean something?

STAN SKAFIDAS
So we had two cohorts. There was what I'd refer to as our discovery cohort. That was used to identify these 237 markers and we developed a classifier. Once we had done that we took a second independent cohort and we applied our analysis to that second independent cohort and tried to estimate the ability to discern unrelated individuals to cases and see what classification accuracy we could achieve. And secondly, if we looked at parents versus their affective children, what classification accuracy could we achieve? So if we were looking more specifically at parents versus controls and they have the same ancestry; And if they've got the same ancestry and a high percentage of similar genetics, the issues relating to population stratification are unlikely to be present. And then also what few SNPs are driving the difference between the parents and their affected children. There was an independent validation cohort that we actually used to validate or to test this classifier on.

SHANE HUNTINGTON
Is there a difference between doing this test on one ethnic population versus another? How does that play out given the substantial genetic differences that we can find between these cohorts?

CHRISTOS PANTELIS
The issue with regard to ethnicity is that the genetic markers differ by population. It's important when looking at classifiers based on this kind of analysis to look at ethnically homogeneous populations, or to develop classifiers that are specific to a particular ethnic group. What we seem to be demonstrating is that some of the pathways we've identified are similar across different ethnic groups, but the markers that lead to the classifier we've developed are going to be different so that's a particularly important issue. The other important issue with regard to looking at using a classifier of this kind is that the ability to identify an individual within the population as being at risk for the disorder is dependent on the prevalence and incidence of that disorder. So the positive predictive value based on the statistics we've generated, when you look at it at the population level it's actually quite low. That means that the best way to use a classifier of this kind is to use an enriched population, to use the classifier in, perhaps, a family where there are already affected members, where there is a high risk perhaps for other members to be affected. That improves the performance of a classifier of this kind.

SHANE HUNTINGTON
And how will a classifier of this type function in a place such as our city of Melbourne, where we have a diverse and multicultural population? It would seem as though from what you're saying this would not be something we could use locally, whereas if you went to a country where the population was less diverse it would be more effective and potentially even more effective where there were more people with autism in that population? Is that all correct?

CHRISTOS PANTELIS
I think that is correct. I think the way perhaps to examine that is to use genetic markers that are relevant to ancestry and you can use ancestral markers to identify
where individuals come from ethnically. And that might allow you to then use classifiers accordingly.

SHANE HUNTINGTON
If I understand this correctly, you guys are focusing very much on the sorts of genes that produce proteins here. There's a large portion of our genome that doesn't do this. Are we potentially leaving something out or missing something important by not looking at that part of the genome?

STAN SKAFIDAS
Absolutely, so Shane as you're pointing out, one of the shortcomings of the analysis is that we've focused on genes that relate to ones that produce proteins. And there's a potential that we've omitted or left out some very important genes which either don't produce proteins or we just don't know what their function is at this moment in time. But interestingly, focusing on genes that do produce proteins allows us now to try to biologically verify if we can see differences in these identified genes that do produce proteins in post-mortem tissue. If these genes don't produce a protein then it's very difficult to be able to do your immunohistochemistry in post-mortem tissue, to be able to verify whether that particular gene has some relevance. It's a very interesting and important question, but we also need to come up with candidates that we can actually somehow test biologically.

SHANE HUNTINGTON
You're listening to Up Close. I'm Shane Huntington. My guests are Chris Pantelis and Stan Skafidas. We're talking about developing a diagnostic test for autism. Stan, the study that you've put out not unsurprisingly has attracted a lot of attention and some criticism and some scrutiny. Why is that? Why is there such a response to this? I mean I can imagine the ethical concerns around this could be interesting especially if it could be done in utero, but is that the only focus of attention?

STAN SKAFIDAS
So first of all I just need to reiterate what Chris said. If we're looking at a general population, our positive predictor value is in the order of about three per cent, so the issue of using it as a general screening test or in utero, I don't think that should ever be considered. Now where it becomes interesting is in an enriched population as Chris has mentioned, but the work has received a lot of attention but also a lot of criticism. The criticism has related to whether this is a true signal that we've detected, or is it due to some other important confound and in particular people are asking is this due to population stratification. And for everyone's benefit, population stratification is really a case where the cases versus controls are from either different ancestry, or there is some other way that the data has been processed which explains the difference rather than the underlying disorder. We believe that we've addressed that in identifying our cohort with identified ancestrally informative markers, which have enabled us to be able to detect an ancestrally uniform population and then have undertaken our genetic study after we've determined this homogeneous population. Some other groups have looked at other ways of choosing their cohort. What they've done is said well we're not going to
find people of a particular ethnic background or ethnicity. What we’re going to do is pool everyone together and then what we’re going to do is try to contend with that variance by removing something called principal components. You determine the variants in that population. You assume that the variance is all due to differences in ancestry and then you remove as much of the variance as you can using that type of technique. The fundamental assumption there is that all the variance that is being removed is in no way related or relevant to the underlying signal. As a consequence we believe that by applying that kind of methodology what has happened is some of these other groups have actually removed the information bearing signal. That has meant that they are unable to achieve the same results as we have. But I think that that relates fundamentally to the way that they have processed that signal.

SHANE HUNTINGTON
If we were to cut that down to a very simple question; if I was for example to have a child that through the older techniques was determined to have autism, say 10 years of age by now and we had another child and we did this test. What would the accuracy be, given there is that enriched cohort in my family and so forth, what would that accuracy be? I mean I think that's the real question in a clinical setting that becomes relevant.

STAN SKAFIDAS
Okay, so our analysis says if you're trying to discern between parents versus their children or siblings, the accuracy's in the order of about 58 per cent.

SHANE HUNTINGTON
Which is a little better than a coin toss. Is that significant in the clinical setting? Like I guess I could ask this of yourself Chris. Would a clinician act on that kind of efficacy, or is that not enough?

CHRISTOS PANTELIS
Perhaps it's important to clarify here that separating parents or first degree relatives from the effect that is at that level. When you look at the comparison to healthy controls the ability to predict is higher. So based on our analyses the positive predictive accuracy was almost 72 per cent, which is certainly a lot better. The issue is, what populations are you comparing, how are you comparing them and what are you trying to predict?

SHANE HUNTINGTON
Now let's talk a little bit about the clinic setting because I think this is important. We've talked about the current way of determining autism and as you said, years is an excruciating experience for some parents to go through. What would this sort of test look like in the clinic? Is this come in one day, blood sample and so forth, move on a couple of weeks later, genetic test results come back? Or is this something that will have to sit side by side with the existing programs that are in place?

CHRISTOS PANTELIS
Well I think a couple of points. With any kind of clinical test that we use it acts as an aid. It acts as an aid to our clinical assessment, whether we're using blood tests to look at the risk for heart problems or high blood pressure et cetera. Using tests in the context of the clinical picture is always going to be the best way forward. The other point I think is that it's still very early days. What we really need to do is undertake prospective studies in populations that might be at high risk in order to test how well such a classifier might work. So there's still a great deal of work that needs to be done in order to understand how well this might work and how well it might work in the clinic and obviously very important, but we need to do those tests. We need to examine how well it works and how well it might work in the clinic. We've got quite a lot to do.

SHANE HUNTINGTON
Do we have a sort of feel for how long that will be before you think that will be something used commonly in clinics?

CHRISTOS PANTELIS
Good question. Often these things take years and if you think about the development of new drugs, that can take decades.

SHANE HUNTINGTON
Stan, Chris, thank you very much for being our guests on Up Close today.

CHRISTOS PANTELIS
Thank you.

STAN SKAFIDAS
Thank you.

SHANE HUNTINGTON
Chris Pantelis is Professor of Neuropsychiatry and Scientific Director of the Melbourne Neuropsychiatry Centre at the University of Melbourne and Melbourne Health. Stan Skafidas is Professor of Neural Engineering at the Department of Electrical and Electronic Engineering, leads the Melbourne School of Engineering’s research and nanoelectronics and is the Director of the Centre for Neural Engineering. If you'd like more information or a transcript of this episode, head to the Up Close website. Up Close is a production of the University of Melbourne, Australia. This episode was recorded on 11 July 2014. Producers were Kelvin Param, Eric van Bemmel and Dr Dyani Lewis. Audio engineering by Jeremy Taylor. Up Close is created by Eric van Bemmel and Kelvin Param. I'm Dr Shane Huntington. Until next time, goodbye.

VOICEOVER
You've been listening to Up Close. We're also on Twitter and Facebook. For more information, visit upclose.unimelb.edu.au. Copyright 2014, the University of Melbourne.