Personal Genome Service Providers
Jeffrey Gulcher, M.D., Ph.D.

MS. AU: Our next presenter was supposed to Kari Stefansson, but he is stuck in Iceland, not Switzerland.

MS. AU: So, Jeff Gulcher will be presenting for Dr. Stefansson. Thank you. Welcome.

DR. GULCHER: I'm a little shorter than Kari and have less hair.

Kari and I founded deCODE about 12 years ago. The goal was to try to find genes for common disease that might help in predictive diagnostics and also in targeting novel drug pathways, finding out which pathways might be most important.

We set up in Iceland so we could focus on one population. We have collected about 140,000 Icelanders with informed consent specifically for diseases that we ask them about. We also have collected about 230,000 non-Icelandic samples from around the world: Europe, the U.S., Asia.

That is very important because when we make discoveries in Iceland with this very large cohort we want to be able to rapidly replicate and determine whether or not there is validation outside Iceland. We have a whole host of Caucasian, Asian, and African cohorts to do that.

We started doing linkage studies using our genealogy database. That is how we made our discovery for TCF7L2 and for the 8Q region for prostate cancer initially. We have expanded and added genome-wide association data using the Illumina platform from 370,000 to a million SNPs for an individual, and we have genotyped about 45,000 Icelanders now with these high density systems to allow us to do a combination of linkage family-based studies or genetic association studies.

The whole goal of deCODE from a risk diagnostic point of view was to make available some of the discoveries that we and others have made for common diseases, just picking diseases where we thought the relative risk compared to the general population might be high enough to have an impact on prevention and early detection, at least in some cases and some niches.

That was the reason why we launched disease risk tests for individual diseases like myocardial infarction, type II diabetes, glaucoma, prostate cancer, and atrial fibrillation and stroke. We thought the risk ratios were high enough to perhaps have an impact in certain circumstances.

These of course are using only markers that have been validated or replicated in six to 60 different populations around the world. In some cases it is only in Caucasians, in other cases it crosses ethnic lines. We make that clear in our reports.

Physicians are already using some of this information in their practices to help risk stratify certain patients in certain circumstances. For example, the type II diabetes gene that we discovered, TCF7L2, we include that in a complement of four type II diabetes genes that decode T2. We are about ready to upgrade that to eight genes.

But already TCF7L2 has been shown in a prospectively collected sample cohort, the DPP study of pre-diabetics, to further double one's risk of converting to type II diabetes within a short period of time. The absolute risk of an overweight or obese prediabetic who is homozygous for TCF7L2

converting is about 50 to 70 percent within three to four years. That is based on the DPP study and the DPS study that was done in Europe as well.

Here is a special niche for a diabetes variant that has been widely replicated in 60 different cohorts but has a certain potential clinical utility of identifying patients who have prediabetes who are at especially high risk for converting. The baseline conversion rate is about 30 to 35 percent. This further doubles that on top of, obviously, the risk factors of prediabetes itself and obesity and being overweight.

ADA has recently addressed this issue of trying to identify prediabetics, number one, and encouraging those patients to lose weight. Then, for those who fail to lose weight who have additional risk factors for conversion, those patients might benefit from pharmacologic management with Metformin and now, recently, with Actos.

When it comes to personal genomics, is there a role for a direct access to consumers with or without their physicians. We saw a way, like the other companies, to make that accessible. Not to say that individuals work with this information in a vacuum. We encourage them to talk with genetic counselors. We do, by the way, offer genetic counseling and have done so for the last six months of our service for free. But we also encourage them more specifically to work with their physicians because we think they actually have a much bigger role to play when it comes to prevention or early detection of cancer or other diseases.

We have been able to pick and choose some of our discoveries to add into some of these diagnostic tools and risk genetic tools. We have been putting together large systems of information to allow us to do the replications.

For deCODEme, we offer about 30 diseases now. Since there was a large discussion yesterday at the HHS meeting about analytical and clinical validation, I want to convince you that analytical validation for genetics is a lot simpler than analytical validation for CRP or even LDL cholesterol measurements. It is because you can document the accuracy of your genotyping, whether it is individual SNP genotyping or an Illumina array, for example, like we use.

You actually have 15-fold sampling for Illumina array. You have 15 beads that are assessing the genotype. So there is redundancy that you can make use of in your quality control. We think we are compliant with CLIA, and we are required to do so under our certificate of registration with CLIA. The accuracy can be measured. It can be documented through repeated testing and matching the gold standard that the FDA sets for genotyping, which is sequencing, which we do for all the variants that we annotate, the 100 or so variants that we annotate.

Then we do quarterly proficiency testing, which is also a requirement of CLIA.

On top of that, we do clinical validation and document those clinical validations that we define the way FDA defines it. That is, replicating the markers and demonstrating that those markers are consistent across populations. That is not a formal requirement by CLIA, surprisingly, but that is something we voluntarily add. In addition to our analytical validation reports, we send the clinical validation reports also to CMS as part of our demonstration that these results are reliable.

When it comes to the genetic markers that we annotate at deCODEme or the individual disease tests, they are all well validated. We don't have any preliminary data or diseases where there might be one or two studies. They need to be replicated widely before we put them into our risk classification.

I should emphasize the relative risk. It is not just that these markers replicate. The important thing also when it comes to clinical validation is how you assess the appropriate relative risk to attach to a particular genotype. What is that based on. Is that based on just a few hundred patients; is that based on thousands of patients.

Most of these variants are actually based upon data sets that we use that include up to 10,000 patients and controls. In some cases it may be 5,000 patients and 30,000 controls, in other cases 12,000 or 17,000 patients and another 30,000 controls. So the bases for these assessments of relative risk are actually based on data sets that are much larger than are used for FDA approval or diagnostic tests that are approved by the FDA.

Then the question becomes what can you do with that information. Can you combine that information reliably, in a reliable and consistent manner. What we do is we convert each one of these variations, these odds ratios, into a relative risk, risk compared to the general population, to have a consistent reference population to attach that risk.

Then we simply multiply those risks together, because we and others have demonstrated that for the vast majority of these diseases, from the various studies that Teri mentioned, you see no interaction whatsoever, even with these large, large data sets. When we combine our data sets with others, we fail to find significant interaction terms to suggest that these are either redundant or synergistic interactions. Therefore it fits a multiplicative model, and we think we are justified at multiplying these relative risks together in a multiplicative model way. That can be useful to assess the risk for a particular individual.

When it comes to the clinical validation and the replication, just since Dr. Khoury had questioned that yesterday, let's take prostate cancer. These markers have been replicated in large numbers of cohorts even in our initial discovery papers, where we have large numbers of patients. You can see at the bottom there are 3,500 patients versus 14,000 controls, and there are five or six different Caucasian cohorts that we have in the United States and Europe. Also, on 17q, another 3,500 to 14,000 controls, and so on.

There is Chromosome 8X discovery. The patient population is 10,000 patients, 28,000 controls. I just put this up as an example. We are in a different era now than we talked about yesterday. Now, to even get published in Nature or New England Journal, you have to have wide replication.

That is the new standard today that was encouraged by Dr. Collins and others over the years. Now that standard exists and you can't even publish these discoveries until you have replication in your seminal study.

I should mention that many other groups have replicated these markers as well beyond our studies, for example NCI and U.K. Cancer. But when you take these relative risks together, you can multiply these individual relative risks, genotype-specific relative risks together. Just like Dr. Collins did in the fusion study and in his genome-wide association study on diabetes with his colleagues, you can multiply these and come up with a composite genetic risk for that particular individual that maps out across a general population like this.

These are the eight markers across a general Caucasian population. About 10 percent of the general Caucasian population has an average risk of two-fold, compared to the general population.

Lifetime risk for prostate cancer is 16 percent. That would translate to a lifetime risk of 32 percent in the absence of other risk factors.

What are the other risk factors for Caucasians. Nothing but family history. The common variants are independent of family history. Less than 5 percent of the general population has a family history of prostate cancer, so that is not necessarily the best screening tool when it comes to risk for prostate cancer. But if you have this, you have another 10 percent. You have doubling of risk of prostate cancer that is independent of their family history.

We try to communicate these results in a clear and consistent manner. We describe it in terms of relative risk because we think the patients and the physician can understand that much better than trying to create a table of odds ratios. They don't need even need to know what a SNP is and they don't need to know much about Mendelian genetics because these are risk factors. These are not Mendelian determinative risk factors. These are actually risk factors that are much more analogous to LDL cholesterol.

So if you can give a reliable risk score for that individual, they can incorporate that with the environmental and other risk factors that they use on their daily basis. They also convert this into a lifetime risk.

When it comes to consistency among the three different companies, you did see from David Duncan's talk that actually, on the face of it, if you look at the relative risks, we are actually fairly consistent across the three different companies. He already mentioned that the accuracy of the genotyping seems to be very high. But when it comes to combining the markers, we are doing it in a little bit different way.

We are getting together with PMC to come up with an industrial standard, so to speak, and getting feedback from academia so that we try to do it in a consistent way. But already it is actually fairly consistent.

The differences and the variation seem to be more in those rare instances where you are using a surrogate that is a little bit different for the original marker, the marker that was initially reported. Although, those are still well validated, those extra surrogate markers.

From that standpoint, I think there is a role for additional consistency, but already I think the results are quite consistent across the different platforms. Some of this depends on whether or not they have updated the latest prostate cancer genes, for example, in their profile as fast as some of the other companies.

I gave some examples, but I think David Duncan already did a nice job demonstrating the comparison.

Are these tests useful today? That is the other debate. Francis says that maybe only 10 percent of the genetic variance or less, or a few percentage points, are accounted for, and that is debatable. We agree, though, that the vast majority of genetic information has not been captured by this.

But, is it still useful to identify those at highest risk or not with these tests. We would contend that even for the heart attack gene, the MI gene, that we and others have discovered, it appears to have an effect that is independent of your Framingham score. When you talk about LDL cholesterol by itself, it has very little impact on the AUC. The AUC is still only 55 percent or

less with just LDL cholesterol. To push up the AUC you have to combine it with conventional risk factors.

This is a major risk factor that is not being accounted for by the Framingham score and has the opportunity of moving a low-risk patient, based on ATP3 criteria, up to intermediate risk, or from intermediate risk to high risk.

There was a recent prospective study done by Stephen Humphries [in the U.K.] that showed in this 15-year prospectively collected cohort that adding these genetic markers on 9P actually reclassified 15 percent of the patients that were originally classified with just Framingham according to the ATP3 criteria. About 5 percent of those patients overall were in the intermediate category and went from intermediate to high risk category. So it shows how it can actually have some utility.

What would you do differently. As a physician, you would target the LDL cholesterol to a different level based on that risk. That is what is recommended by NCEP3 guidelines.

I gave you my example on prostate cancer yesterday. For atrial fibrillation, we discovered these two markers which are very strong and by themselves double one's risk for atrial fibrillation. Twenty-five percent of us in this room have these high-risk genotypes.

What is interesting is these are by far the strongest-acting stroke genes. If you do a genome-wide association study for stroke, and our publication will come out in the next few weeks on that, it turns out these by far are the strongest-acting genes. Why is that. Only 15 percent of stroke is due to atrial fibrillation. That is what is thought to be. But it turns out that there is much more atrial fibrillation that contributes to stroke, especially in the cryptogenic stroke category -- that is stroke of unknown cause -- and even in large vessel stroke categories that was not realized until we did these studies.

Now we are doing a very large observational study to demonstrate this, but we estimate that 150,000 stroke and TIA patients are misdiagnosed as having a different type of stroke or a stroke of unknown cause where they really had intermittent atrial fibrillation that is asymptomatic and that can be picked up by extra cardiac monitoring.

If you applied our genetic test and you did the extra cardiac monitoring for a few extra weeks when the patient is discharged from the hospital, based on prevention, putting the patient on the correct drug, moving them from an anti-platelet to Warfarin if they do indeed have atrial fibrillation on monitoring, that can save the healthcare system $1 billion a year.

Finally, for those who didn't see my case study yesterday, this is me, 48 years old. I have no business testing my PSA, at least based on guidelines, even though my father had late-onset prostate cancer that was considered benign. According to the guidelines, family history only counts if your father was younger than 65. But I was compulsive enough to get it at age 42, with low normal.

When I got my prostate cancer test results back using these eight markers that I showed you, my relative risk was 1.88 compared to the general population. Calculated lifetime risk for a white male would be 30 percent based on this.

I also had extra markers that suggested that I am more likely to have the aggressive form rather than the non-aggressive form of prostate cancer by about 1.3-fold, not dramatic. That prompted my primary care physician to get another PSA test.

This time my PSA was still in the normal range. Zero to four is normal range. But my PSA was high normal at 2.5. The anagram is coming up later.

My high genetic risk prompted my primary care physician to refer me to a urologist. He was concerned enough to do a biopsy and found that I have prostate cancer with a Glisson score of 6. That is intermediate grade. I will have that taken out by Bill Catalona in two weeks.

Now, this is just anecdotal of course, but once again, can we try to improve the sensitivity and specificity of the biomarkers that we are using today, which we all agree are not perfect. Can we improve that specificity and sensitivity by adding extra genetic information, just like family history is already being used to guide these types of managements. Thank you very much.

MS. AU: Thank you, Mr. Gulcher.