ADAM BRUFSKY: So we have a lot of powerful tools, and their toxicity and benefits don't correlate often. We try to understand the benefits of intervening, and we have multiple molecular prediction models for the determination of the effects of any intervention we do. And we really have three familiar assays in the United States that a lot of us use, and that's the 21-gene assay, the 70-gene assay, and finally, the PAM50 ROR score. And again, just to familiarize yourself, the 21-gene assay's Oncotype, the 70-gene assay is MammaPrint, and the PAM50 ROR score's Prosigna.

So those are the three ones we know about. Just to show you, we know that the 21-gene assay is not 21 genes, it's really 16. Gene proliferation, estrogen receptor invasion, HER2, and a few other ones, like C68, which is a very interesting gene that really hasn't been explored and probably should be. And then there are five reference genes that allow Oncotype or Genomic Health to take RNA from anywhere in the world and normalize iit, so you know that, if score is 17 from a block in Israel, it's going to be 17 by the same assay and the same RNA expression from a block in Pittsburgh.

So that's the 21-gene assay. We know the 70-gene assay a little bit different. It's 70 genes, and again, by our proprietary mechanism, you look at the RNA expression levels of all these genes. Green as high, Red is low. And then you basically, with that statistical algorithm, classify either as low-risk or high-risk. Low-risk is someone who will not have a relapse in a 5-year period, where high-risk has about a 30% to 40% relapse. We'll get into the 70-gene assay a little bit later.

The PAM50 ROR score is a little bit different. RNA is extracted actually in our pathology labs. The other two are actually sent off. There's actually a system for quantitating the RNA, which is actually done also in our pathology lab here. You get an expression profile of 46 genes. And then the idea is, and we'll talk a little about intrinsic subtyping in a few minutes, but the idea behind the ROR score is that you basically, based on the RNA expression, find out which of those intrinsic subtypes it is, luminal A, luminal B, basal-like, or HER2-enriched. And then you use that, plus a proliferation score of a subset of 19 of those 46 genes plus the tumor size, and you come up with an ROR score. So that's also one of the three main marketed tests in the United States right now.

However, there are two genomic assays, which Barry alluded to, which we're not as familiar with. The Breast Cancer Index some of us are familiar with, but EPClin, or EndoPredict, we're not familiar with, because it's something that actually is likely to be marketed in the United States within the next couple months. So the Breast Cancer Index is very simple. It has two components. It has a molecular-grade index, which has five cell cycle genes, and that's really useful for early recurrence detection. It also has a ratio of HOXB13 to Il17BR. That ratio is very interesting. We're not sure what those genes really do, but we know that that ratio is a measure of estrogen sensitivity. And so, again, you just take the tumor block, you do an RT-PCR, or real-time RNA expression analysis, on it, and then you measure these two genes and assign a risk of distant recurrence.

BCI is important, not so much that it does better, and we're going to get into this a little bit more. Barry hinted at it in the last talk that BCI is better. BCI is interesting, not because of the early recurrence, but actually. It appears to actually in three separate cohorts, one of which came from us. That middle one, Rachel Jankowitz is actually one of the senior authors on that particular paper. These three cohorts actually, if you have a low-risk H over I score, you can actually tease out a patient with node-negative breast cancer that really will do very, very well between years 5 and 10, and obviously, as Barry went into that before, those are people who you don't need to give extended endrocrine therapy to.

So the EPClin, though, is something that really is new to all of us. It's really new to me. It's been around for quite a while. You can see this paper is from 2011. The idea is that EPClin took various genes, as you can see here, eight different genes, added this reference gene, and then they added nodal status and tumor size to come up with a risk class, and if your risk, by this EPClin score, is less than 3.3, you're low-risk. If you're higher than 3.3, you're high-risk. And they define the risk as a less than 10% distance recurrence at 10 years. That's what they define as low versus high risk. So this is yet another test about to come out, and this is important because we're going to go through, as Barry did briefly, the bake off, which I think was one of the more important abstracts, at least in this field, at San Antonio this year. This just shows you the EPClin, by low-risk, can actually, again, tease out, and this is from data from a European trial, ABCSG-8, can actually tease out a group of patients at very low-risk of recurrence at 10 years. And these are the women probably that you can actually deintensify their adjuvant systemic therapy.

So really, at San Antonio this year, the three questions that were there, at least in the molecular assay realm, and I'm focusing on the clinically relevant molecular assay realm, so actually a lot of interesting papers and more research papers where people are trying to figure out sensitivity to PARP inhibitors, sensitivity to platinums, that were kind of interesting, that may have clinical relevance in the not-too-distant future, but right now, that were really not ready for clinical primetime, and so I'm going to hold off on those for the time being, maybe talk about them next year. But the three questions really, can good pathology supplant the genomic assay, which genomic assay performs best, and in a real-world setting, what are we really doing, testing in node-positive breast cancer, given that almost all of this data is in node-negative breast cancer. So I'm going to talk about those three abstracts that were done.

The first one is this Anne Arundel Medical Center, which is a center outside of Annapolis. It's staffed by former Hopkins physicians that didn't want to be in an academic practice anymore and a lot of their fellows go there. It's a really nice group. In collaboration with M.D. Anderson, what they did is actually, they have a very simple model, using pathologic tumor-grade and hormone receptor status, to predict 10-year disease recurrence. And so what they did in this particular trial is they actually use their model, I'm going to show you in a minute, on a series of cases that were done at M.D. Anderson. There's 1,264 with which they had 10-year follow-up for distant recurrence, as well as an Oncotype DX score. And this just shows you the risk models here, the AAMC, the [INAUDIBLE] risk, is basically low-risk, because anything that's grade 1 and with positive PR expression.

They also use two cutoffs for Oncotype DX, the first one being the one that's used in TAYLORx, which is anything less than 12. The reason that less than 12 was chosen, it's the lower end of the 25% or the 5% confidence interval, the 95% confidence interval, around low-risk. Remember, low isn't irregular, as you can see here, anything less than 18, but 18 has a confidence interval of six on either side of it. That's something people don't realize when we do these tests, is that there is a statistical variability of about six. It's something we don't realize when we think of the precision of these tests it's not as precise as we think. Higher risk was grade 3 or ER less than 30%. Again, in TAYLORx, the higher risk of recurrence score was anything greater than 25. In Oncotype, anything greater than 30.

So what they did with all three of these is, they saw how well, in this node-negative subtype, they were able to predict distant recurrence by all three of these assays. And you can see here, the first ones are in the low-risk, actually are pretty close to each other. First of all, the AAMC picked out a third of the patients by this metric, and they had a 10-year distant recurrence of 2.7%. In TAYLORx, it was 4%, and that was about 250 patients. But when you use it less than 18, it's actually 3.4%. There's not a lot of difference. But apparently. If you use a recurrence score cutoff of less than 18, you get a lot more patients, actually, that have a low recurrence score. If you look at the high-risk and the intermediate-risk, it also was fairly close. Again, the intermediates were 8.4%, same ones from the TAYLORx cutoff. Although Oncotype DX seemed to do a little bit better, there's a much greater separation between people who had recurrence score of 18 to 29 at 15.2% recurrence risk. And finally, the bottom ones, so leave the intermediates out. If you really look at high and low, which I think is a more important thing here, if you had patients with grade 3, or an ER less than 20%, their 10-year distant recurrence risk was about 22%, which is very close to either the TAYLORx or Oncotype dx cutoff.

So the bottom line here again, looking at Kaplan-Meier curves, they're fairly reproducible. That if you use an AAMC high or AAMC low, you really can tease out the groups that really are at high-risk of recurrence and at low-risk of recurrence, very simply, with a tumor grade and a PR and an ER. Three very simple things that a lot of people can do and almost replicate the Oncotype DX score here. So really, that was their bottom line is that, with two simple things, you can actually replicate the score and perhaps use that if you don't have access to the score, but if you have really good tumor grade.

Now this is not a new concept. A group at John Hopkins just published an algorithm. We've published our Magee equations many years ago. Actually, we first started publishing these about 2007. Our latest update is in 2013, and this is available on the internet. You just go to McGee equations, you put in the Nottingham score, ER/PR, HER2, tumor size, and Ki. It can actually come up with a number, and the numbers are very close. As you can see on the left hand panel, the actual Oncotype DX score in this series of about, I think this is 255 cases. You can see, it's very close to the algorithm, especially the McGee equations 2 and 3, some of the older equations. But what's really important, and if you look at the right-hand panel here, is that, if you have an estimated recurrence score that's intermediate or low, you almost never have an Oncotype DX score that was high. And so that's the bottom line. Low recurrence score is very predictive of a low actual Oncotype recurrence score. This is something that potentially, as we'll talk about in a few minutes, can be used as a prescreen for whether one wants to do a genomic test.

So again, good immunocytochemistry can supplant gnomic profiling, as we can see from San Antonio. It should be, though, with an experienced lab. The problem here is that you've got to be someone who sees a lot of breast cancer and knows what a grade 1 is. That's the whole point of why Oncotype was developed, because again, people weren't sure that their pathologist could tell a grade 1 from a grade 2 from a grade 3. But a more important question is, looking at these intrinsic subtypes, which test actually does better, not only at predicting, but rather at doing the intrinsic subtype? And so for that, let's just talk briefly about MINDACT. MINDACT was a 70-gene signature, as we know, the first real prospective trial of one of these. TAYLORx was the other, which is reported only its low-risk patients. It was, basically, 6,700 women with node-negative and 1 to 3 node-positive. The patient enrolled over a 5-year year period ending in 2011. It was about a 50 million euro investment by the EORTC and six European cooperative groups. The idea behind this trial was to do a clinical risk based on Adjuvant! Online, which had clinical pathologic parameters in it, to do a genomic risk using the 70-gene signature.

The most important part of this trial was actually this part here that I'm using with the arrow. Patients that were clinically high, that you ordinarily would have thought to be high-risk and giving chemo to, but were genomically low, were randomized to get chemotherapy or no chemotherapy. And that was about a quarter of all the patients in this. And just to show you this group, when you treat these patients, and in this case it was 750 of those patients with hormone therapy alone, basing it on the genomic and not the clinical risk, you see you have a 5-year distant disease-free survival of 95%, which is pretty good. And the other thing is that these were patients who you ordinarily would have thought to give chemotherapy to. About 60%, or 2 centimeters or greater, were T2. 93% were grade 2 or 3. A half had lymph node-positive disease. Obviously, most of them were hormone receptor positive. So clearly what this actually showed in a prospective fashion is that there are groups of high-risk patients that, if you are 70-gene assay, low-risk, and treat with hormonal therapy alone, you get year at least 5-year, not 10-year because there isn't 10-year data, 5-year distant disease-free survival. And again, chemotherapy, just to go forward with this, didn't really have any benefit in these women, either they got chemo or not. The problem is that the trial was not powered really to answer this question, and there still could have been a 1.5% absolute benefit to chemotherapy that was not picked up by this data.

This really wasn't presented, this is more of the background. What was presented at San Antonio was you could actually do the intrinsic subtype of these patients. So there are three basic intrinsic subtypes, four, really. There's luminal A, which is a slow-growing ER-positive subgroup. Luminal B, which is an ER-positive subgroup that's higher growth, higher proliferation rate, potentially patients that would need chemotherapy. There's HER2, which is responsible and responsive to HER2-based therapies. Finally, basal, which is none of them. Those have very specific surrogate definitions by immunocytochemistry, as you can see here. Luminal A is generally ER/PR-positive, HER2-negative and a Ki-67 less than 20. Luminal B is the same ER low, and/or Ki high. HER2-enriched is HER2-positive. And finally, triple-negative is something that's all three are negative.

And so what they did in this particular trial is they took the pathologic patients and reclassified them by their molecular subtyping. They had an 80-gene assay that could classify patients' molecularly into luminal A, luminal B, HER2-enriched, or basal-like. And what you can really see here is fascinating, that when you look at the luminal As, the pathological luminal As, about 10% are luminal B. When you look at the pathological luminal Bs, it turns out 54% are luminal A. So really, almost half of the patients, if not more, potentially are being overtreated based on the pathology, not the molecular subtyping. The HER2 is even more interesting, that you find only 60% of the pathologic HER2-enriched patients, pathologic HER2 patients. The patients that you think will respond to trastuzumab, actually, only 57% of them are HER2-enriched. And there was another study, I don't know if they're going to talk about this in the HER2 section PAMELA that was presented, did the same thing. And in PAMELA, it turns out, that also about 60% of the patients who you thought were HER2-positive by immunocytochemistry and/or fluorescence in situ hybridisation, were actually HER2-enriched.

So this data set here has been replicated in other data sets. So what that says is, only 60% of the patients who you think should respond to trastuzumab, potentially, are really going to. The other 40 may not. And actually, as you can see here, 17% of those patients are luminal A. And those are patients that probably will just respond to hormonal therapy alone, a very slow-growing disease. I think that's something that will have implications going forward as we try to figure out what to do with these tests. Finally, about 10% of the triple-negatives were actually other subtypes, which also is interesting, because if you give someone chemotherapy potentially, and in the new adjuvant setting, they don't respond, maybe they're HER2-positive. And you can see here, using pathologic subtyping actually does a very good job of predicting the 10-year or the 5-year distant metastasis survival, so does molecular. But the difference is the molecular identifies another 16% of patients that were luminal A and probably didn't need further really intensive adjuvant therapy.

This actually is another piece from this, showing that in the triple-negative patients, seven of them, not a lot, were HER2-positive. But they had the worst prognosis. Why? Because they didn't get anti-HER2 therapy. So this is something that, really, we have to start putting in the back of our mind that, maybe the pathologic subtyping is not as precise as we would like it in trying to define what therapies that we do. So again, intrinsic subtype assay does better, and I think it really depends on the assay. More importantly, I think we use IHC, or we should, in clinical parameters to determine the need for gene expression assay.

So finally, which test performs best? And I'll end my 20-minute talk with this. So Barry had talked about this already. This was a bake off of 818 patients with five years of tamoxifen and an [INAUDIBLE] result. This was the ATEC study, and they had 818 tissue blocks. They had 10 years of follow-up, Distant recurrence was the primary endpoint. They use a Cox regression model, this LR chi squared model, to determine the additional benefit. Commercial cutoffs were used to determine 10-year disease risk. And again, they did node-negative and node-positive. These are the six tests that they did. They all compare them to a clinical treatment score, which is basically Adjuvant! Online. You see whether all of these other tests worked. And Barry went through the data little bit, but I was going to give you the years zero to 10 data. What you can see here is that all of the patients, all of these major assays, did pretty well. They all had about half to 2/3 of the patients that they identified as low-risk with a 10-year disease-free recurrence of about anywhere from 4% to 6%. These are people who really need very little therapy, endocrine therapy probably. How long, we don't know, but all these guys had five years. So there's a certain group of patients that are node-negative that, with five years of endocrine therapy, do absolutely fine in any one of these tests, even though they all have different genes and different parameters, seem to do the same.

When you look at node-positive patients, however, BCI, and most importantly, recurrence score, tend to break down. So they identify a group, about half of the patients, but their 10-year distant disease-free survival is fairly similar to the patients in the intermediate and higher-risk groups. You can see here, that's pretty clear, which is kind of interesting. Whereas risk of recurrence score in EPClin take fewer patients. In risk of recurrence score, it's 6.6% of the patients that have no recurrent relapses at all, were node-positive disease. And the same thing with EPClin. That's about 5.6%, or about 20% of the patients, have no relapses. And you can see the separation of those versus the intermediate, and in the case of EPClin, the EPClin high.

So there are clear differences in the performance between years five and 10. Barry hinted at this alread in node-negative. All of them do very well in years five to 10. But clearly, only ROR and EPClin are actually showing you separation between years five and 10. Which is interesting, because a lot of us use recurrence score for node-positive breast cancer. Figuring that it's going to work just as well, as it does in node-negative, and we do use BCI as well for node-negative. Some of us use it for node-positive, thinking it will work as well as node-negative, and that may not necessarily be the case. It's something that we really, really, really have to think very carefully about when we do these tests.

So again, all are reasonable, I think, in the node-negative populations and selected populations. When Barry had said, just to reiterate, between years 5 to 10, EPClin seems to do best. ROR does very well. And just as Barry said before, who do you use this in? Potentially, if a woman comes in, as you said, is miserable, not wanting to do anything, you can actually give her reassurance that when you stop her therapy, her risk of distant recurrence between years 5 and 10 is actually going to not be that high. What do we do with node-positive? Node-negative, that's clearly fine. Node-positive's a really interesting question. Do we use ROR? Do we use EPClin when it comes out? Do we use something else? At this point in time, I think the jury is out.

Now, that's interesting, because I just showed you that data for node-positive used for the recurrence score. Well, these guys actually at the Mayo Clinic went back and looked at a huge insurance database to see how often patients were given recurrence scores for node-positive disease. And basically, these guys looked at what's called the National Cancer Database, and they tried to assess the use of recurrence score in node-positive, ER-positive, HER2-negative breast cancer, to try to evaluate the factors associated with recurrence score testing, and they try to evaluate how it impacted the chemo recommendation. This National Cancer Database is an insurance database. It's maintained and has about 72,000, 73,000 patients that were identified that were ER-positive node-positive breast cancer between 2010 and 2013. Interestingly enough, Oncotype was performed in 21%, 15,000 of those patients. One in five actually had a score. And it increased. It used to be 15%. Toward the last year, it was 25%. I bet if we did a point estimate now, it'd probably be 33% or 40%.

So use in node-positive breast cancer is increasing. The patients who were tested were between 50 and 80 usually, lower T, lower N, lower grade. So these are women who you really are trying to figure out. You don't want to give them chemotherapy. Interestingly enough, blacks seem to be less tested than whites. Patients with no insurance weren't tested as much, and community cancer programs versus academic centers were not tested as much. So chemo was recommended in 81% of the patients who weren't tested and 50% of the patients who were tested. So it did decrease the use by about 30%. Now, looking at the chemotherapy recommendations, you can see here, in the low-risk, about 36% were recommended for chemo, intermediate, between 66% and 85%. Interesting, the high intermediates had more of a chemo recommendation. High-risk, just about everybody got chemotherapy. And just again, looking at that, looking at the low-risk and high-risk groups, intermediate-risk groups here. But the interesting thing is, and I want to point this out, the N2 and N3 patients, even with the low-risk, 70% of those patients were recommending chemotherapy.

So why do the tests? And this gives our payers even more impetus not to pay for these tests, but you want to do a test where you're going to use the result. Interestingly though, one must say, that 71% who got chemo, but 29% did not. So actually, a third of the patients, or 30% of the patients, didn't get chemo with node-positive breast cancer, based on this, even with high. And again, looking at MammaPrint for the 1-to-3 positive group for the 70-gene assay, you can see here again, there is really no chemotherapy benefit either for big tumors or actually tumors with one to three nodes positive. There's no benefit to chemotherapy with the patients that are clinically high, genomically low, looking at the right-hand panel here.

So really, again, I want to summarize this. What do we do now? I think the most important thing is that we do some sort of prescreen. And I know most of you in this room, who actually do in our UPMC system, and use the health plan. We all have had meetings with a health plan reasonable, and basically, they ask for some sort of clinical pathologic variable to do, either Adjuvant! Online, which is not unfortunately available right now, because it's being retooled. NHS PREDICT, which is like Adjuvant! Online, they require an age predict, or Magee Equations, prior to ordering some sort of multiparametric test. I don't know if we're ready to announce it yet, but we've worked out a much more simpler algorithm for you to actually put to tell the health plan. And I think the health plan has agreed it. We're not going to talk about that yet, because we want them to announce it, but the thing is we're going to simplify that. But there has to be some sort of prescreen.

The health plan allows one test, you get one. And so, basically, some of the recommendations, reasonable ones would be where the data is. If the patient's ER-positive or node-negative, the 70-gene assay, the 21-gene assay, or PAM50. I use the Magee Equation, but it could be any one of those prescreens, indicates that, again, it could be the Magee Equation, could be clinical risk, NHS PREDICT. And I actually am one of the lone people who do this. If you have triple-positive breast cancer, that's strong ER-positive, maybe low-KI, HER2-positive, I will do a MammaPrint, because the MammaPrint clearly shows that 23% of those patients will be luminal A, clinical low-risk, and will not need Herceptin and chemotherapy. And so again, I'm in the minority on this, but I have a feeling that this is going to be more important as time goes on. For ER-positive, lymph node-positive disease., I think the data really shows, at least for one to three nodes, and you're post-menopausal, the 70-gene assay or PAM50, that's where the data is right now. For pre-menopausal, the only real prospective data we have is for the 70-gene assay.

So that's kind of where our recommendations would be. Clearly, these are in flux as time goes on. Barry, unfortunately, got taken out to a phone call, so we can't have perfect timing when we take the question and answer session, but I'll end there.