Morgan is an assistant professor of pathology and Epidemiology at the school of Madison. She’s a member of the combined program in computational biology, as well as the Center for research on Aging and her work. Her multidisciplinary work has really been integrating new methods of statistical genetics, mathematical demography to develop, sort of a new high dimensional mix approach to aging in both humans and animal models and applying those.
Efforts to a variety of major chronic disease, most notably cancer, and so Morgan really pleased to hear about your work and looking forward to talk. Thank you so much.

OK, maybe we can see that yes. And let me make it bigger on my screen. OK, um, so today I’m going to talk about some of my work on in developing biomarkers using DNA methylation data to study aging and diseases like cancer. Why isn’t it? I’m so I usually like to kind of remind people what the biggest risk factor for most major cancers is,
and I like to illustrate this often using something like lung cancer. So a lot of times when asking students what the biggest risk factor for lung cancer is, they’ll say something like cigarette smoking, which we know increases the risk. The incidence and death from lung cancer by about 15 to 30 fold. But in reality, aging itself is actually much bigger risk factor for developing lung cancer, so for individuals who are 25 to 29 years old. About one in 200,000, you have about one in 200,000 chance of Belton lung cancer, however. Nearly 400 and 100K.
so it UH-80 full increase risk for the OR 800 fold increases for those 75 to 79. And this is the case across a wide variety of cancers. We see, UM, in general, an exponential increase with age in both incidents in mortality risks from cancer. And you know, some people have thought that this is just commit probability with time. So at the longer you live, the more time and the more likely they will develop cancer. But really, what we think is that it’s actually
Another changes that accompanied the aging process that are actually playing a causal role. In the ideology of major diseases like cancer, so I like this kind of New Yorker Cartoon, which says you’re deliberately putting yourself at risk avail help by being over 65. So one thing that my lab is really interested in is can we actually try and quantify some of these aging changes that might underlie risk for things like cancer or other major chronic diseases? And so this is where kind of
biomarkers of aging come in.

Uh, so aging is. Not an observable, it’s a latent concept. So it’s hard to define. Biomarkers can actually serve as useful proxies that we can estimate the agent Ness of a cell or tissue, or on the whole organism level. They serve a variety of purposes. They can be used as clinical trial endpoints for interventions to try and slow the rate of aging there. You can also be used for basic biology to understand aging.
And also for risk stratification and the goals in developing some of these biomarkers is that you should have a biomarker that differentiates between a 20 year old and an 8 year old, which is pretty easy. You can even use facial image to do that, but probably the harder thing is, can you actually differentiate risks among individuals of the same chronological age? So can you identify who might be aging faster or slower and then in turn, does that have implications for the risk of a future morbidity mortality? So most of the biomarkers in my lab
works on a more epigenetic biomarkers and specifically involved in DNA methylation, so I like to think of the meth alone as kind of the molecular operating system it instructs else how they should behave and respond is involved in a number of different cellular processes, but a really interesting thing that was pointed out more than I think 30 years ago is that there does seem to be genome wide patterns. Um that emerge in terms of changes in Maculation with aging. So you gotta change net in the maculation landscape as a function of age.
And based on this, uh, a number of labs, including ours who developed what we call these epigenetic clocks. So because they have been very precise, age changes that have been observed. We actually use machine learning to predict the age of a sample based on the DNA methylation level. So you can take a sample from whole blood from tissue in a cell culture, and we often measure methylation at 10s of thousands to now up to 850,000 different CPG sites across the genome. And then what people have done is applied supervised machine learning methods.
to actually develop age predictors.

So most of the early clocks were trained to predict things like chronological age, the first Clock being published in 2011.

However, more recent clocks have actually, which we call the second generation at.

The generic clocks were developed to predict age coral.

It’s so not chronological age,

but things like mortality or physiological processes that change with aging.

So mostly that was.

Our Clock is one of the second generation clocks inside the John Clock.

And the second generation clocks actually
tend to be much better predictors of future disease and mortality risk.

Uhm, but first I just want to show kind of how these clocks look across different tissues.

So this is an example of five different epigenetic clocks in a variety of different tissue are fluid samples.

On the X axis, I'm showing chronological age on the Y axis.

These two clocks by Horvath were actually trained using multiple different issues simultaneously pulled together,

so that's why you get much better agreement across the tissues in
terms of their predicted ages, whereas the other three clocks are actually all trained in whole blood, but still do predict still do show. Very heists age correlations. In other tissues, and actually, a lot of these age correlations are above .8 two point 9. But the interesting thing is you also, if you actually took the time to map these colors out is kind of these divergent issues tend to be samples from brain or these tend to be non bring samples and we actually
00:07:10.876 --> 00:07:12.970 think that it’s important to have
NOTE Confidence: 0.90098894
00:07:13.032 --> 00:07:14.897 differences in Appleton at age
NOTE Confidence: 0.90098894
00:07:14.897 --> 00:07:17.113 between tissues because we all know
NOTE Confidence: 0.90098894
00:07:17.113 --> 00:07:19.393 to choose don’t age at the same rate.
NOTE Confidence: 0.90098894
00:07:19.400 --> 00:07:21.392 So we actually shouldn’t be forcing
NOTE Confidence: 0.90098894
00:07:21.392 --> 00:07:23.160 similar epigenetic gauges across tissues.
NOTE Confidence: 0.91833216
00:07:26.170 --> 00:07:28.072 And then we can actually also
NOTE Confidence: 0.91833216
00:07:28.072 --> 00:07:30.205 show that epigenetic age is also
differentiates normal tissue from tumor.
NOTE Confidence: 0.91833216
00:07:30.205 --> 00:07:32.215 but that is not the case
NOTE Confidence: 0.91833216
00:07:32.220 --> 00:07:34.356 across all the clocks.
NOTE Confidence: 0.91833216
00:07:34.356 --> 00:07:35.780 It tends to be the case across
NOTE Confidence: 0.91833216
00:07:35.780 --> 00:07:38.167 these second generation clocks,
NOTE Confidence: 0.91833216
00:07:38.167 --> 00:07:41.996 where we can see that in the normal
NOTE Confidence: 0.91833216
00:07:41.996 --> 00:07:43.914 tissue you get significantly lower
NOTE Confidence: 0.91833216
00:07:43.914 --> 00:07:46.458 epigenetic age compared to the tumor,
and these are all adjusted for chronological age. Um, so on our Clock and also the Clock by Yang Show these differences across a variety of different tissue types. So one question that we’ve really been dealing with is, you know all these clocks for developed to predict the same thing. To capture this kind of epigenetic or metalation based change with aging. Yet they seem to be perhaps capturing different parts of this epigenetic aging signals. So basically,
00:08:17.076 --> 00:08:19.091 can we identify the individual
NOTE Confidence: 0.9050141
00:08:19.091 --> 00:08:20.670 components and decompose the
NOTE Confidence: 0.9050141
00:08:20.670 --> 00:08:22.494 signal to adapt to figure out
NOTE Confidence: 0.9050141
00:08:22.494 --> 00:08:24.343 what the different parts are and
NOTE Confidence: 0.9050141
00:08:24.343 --> 00:08:26.095 how they map onto disease risk?
NOTE Confidence: 0.9050141
00:08:26.100 --> 00:08:28.354 So this is kind of an illustration
NOTE Confidence: 0.9050141
00:08:28.354 --> 00:08:30.270 of taking the clocks apart.
NOTE Confidence: 0.9050141
00:08:30.270 --> 00:08:31.770 And then figuring out which each
NOTE Confidence: 0.9050141
00:08:31.770 --> 00:08:33.280 part of the Clock is doing.
NOTE Confidence: 0.855812
00:08:35.310 --> 00:08:38.496 So the way that we did this is we
NOTE Confidence: 0.855812
00:08:38.496 --> 00:08:41.020 applied something called WG CNA,
NOTE Confidence: 0.855812
00:08:41.020 --> 00:08:43.300 so it’s a weighted network analysis
NOTE Confidence: 0.855812
00:08:43.300 --> 00:08:46.615 and we did this a cross using six
NOTE Confidence: 0.855812
00:08:46.615 --> 00:08:48.765 different issue in fluid datasets.
NOTE Confidence: 0.855812
00:08:48.770 --> 00:08:51.626 So we had uh samples from dermis,
NOTE Confidence: 0.855812
00:08:51.630 --> 00:08:52.854 epidermis, breast dorsolateral
00:08:52.854 --> 00:08:55.295 prefrontal Cortex Colon, an full blood.

00:08:55.295 --> 00:08:58.730 And the goal here was to identify Co

00:08:58.730 --> 00:09:01.300 maculation modules that are shared

00:09:01.300 --> 00:09:04.799 across all these tissue or sample types,

00:09:04.800 --> 00:09:07.560 and from this we were able to identify

00:09:07.560 --> 00:09:10.286 16 of these Co maculation modules

00:09:10.286 --> 00:09:13.208 using Skeggs from the clocks which

00:09:13.291 --> 00:09:15.646 word starting with about 1600.

00:09:19.070 --> 00:09:21.930 I’m so the next thing we did is we actually

00:09:21.997 --> 00:09:24.511 looked at how these different modules

00:09:24.511 --> 00:09:27.119 are impacting the overall Clock scores.

00:09:27.120 --> 00:09:29.920 So in this I’ve color coded all the

00:09:29.920 --> 00:09:32.624 16 modules and you can see that in

00:09:32.624 --> 00:09:35.626 our Clock and this Clock by Hannum a

00:09:35.626 --> 00:09:38.092 large proportion of this is actually

NOTE Confidence: 0.9095152
00:09:38.100 --> 00:09:39.930 driven by this yellow module, 
NOTE Confidence: 0.9095152
00:09:39.930 --> 00:09:42.506 whereas the two clocks by Corvette seem 
NOTE Confidence: 0.9095152
00:09:42.506 --> 00:09:45.010 to have relatively similar proportions in 
NOTE Confidence: 0.9095152
00:09:45.010 --> 00:09:47.674 contributing to the overall Clock score. 
NOTE Confidence: 0.9095152
00:09:47.680 --> 00:09:49.420 But the interesting module that 
NOTE Confidence: 0.9095152
00:09:49.420 --> 00:09:51.550 I’m actually going to talk about 
NOTE Confidence: 0.9095152
00:09:51.550 --> 00:09:53.200 today is this Brown module, 
NOTE Confidence: 0.9095152
00:09:53.200 --> 00:09:55.378 which actually is shown in most 
NOTE Confidence: 0.9095152
00:09:55.378 --> 00:09:58.113 of these clocks and has a pretty 
NOTE Confidence: 0.9095152
00:09:58.113 --> 00:10:00.118 similar proportion of about uhm. 
NOTE Confidence: 0.9095152
00:10:00.120 --> 00:10:01.723 10 to 15% in each of the 
NOTE Confidence: 0.9095152
00:10:01.723 --> 00:10:03.380 clocks to the overall signal. 
NOTE Confidence: 0.9095152
00:10:06.230 --> 00:10:08.651 So the other thing we can do is not 
NOTE Confidence: 0.90328705
00:10:08.651 --> 00:10:11.139 just look at what proportion of the 
NOTE Confidence: 0.90328705
00:10:11.139 --> 00:10:13.450 clocks is explained by each module, 
NOTE Confidence: 0.90328705
00:10:13.450 --> 00:10:15.090 but whether what their capturing
00:10:15.090 --> 00:10:16.730 is actually the same signal.

00:10:16.730 --> 00:10:18.686 So this is all the modules,

00:10:18.690 --> 00:10:20.568 but I’m going to really focus

00:10:20.568 --> 00:10:22.630 just on 2 for right now,

00:10:22.630 --> 00:10:24.639 so basically this is the part of

00:10:24.639 --> 00:10:26.297 each Clock that that’s represented

00:10:26.297 --> 00:10:28.859 by Stevie jobs in this Brown module.

00:10:28.860 --> 00:10:31.803 And what you can see is that for these,

00:10:31.810 --> 00:10:33.635 epigenetic clocks have really similar

00:10:33.635 --> 00:10:35.861 or high agreements in terms of

00:10:35.861 --> 00:10:37.546 their epigenetic age signal here.

00:10:37.550 --> 00:10:39.325 However, just a contrast this

00:10:39.325 --> 00:10:40.745 on this purple module,

00:10:40.750 --> 00:10:43.340 you can see that in in two of the clocks

00:10:43.406 --> 00:10:46.010 what the proper module is contributing

NOTE Confidence: 0.90328705
is considered accelerated aging, whereas in the other two clocks or three clocks, it's considered decelerated aging. So this is an example of a module is differentially waited and might be contributing to differences in the performance by the various clocks. But for the rest of the talk, I'm going to focus on this Brown module, which seems to be the one that's most important in terms of cancer. So now what we can do is we can look at instead of looking at the entire Clock score, look at the individual modules.
Actually driving this kind of these associations that we’re seeing?

So in this case I’m looking at just the part of our Clock that’s captured by CP GS in this Brown module.

So this is just 21 CP GS over all, and what we can see is we can kind of recapitulate the finding with the tumor versus normal across these different issues.

However, in this case it’s actually more significant when we’re just considering this Brown module.

We can also look up this is in normal breast tissue and we do see that this module is significantly correlated with...
00:11:58.379 --> 00:12:02.170 age in normal breast, suggesting that.
NOTE Confidence: 0.9159804
00:12:02.170 --> 00:12:03.574 Perhaps as women age,
NOTE Confidence: 0.9159804
00:12:03.574 --> 00:12:05.329 their breasts as she develops.
NOTE Confidence: 0.9159804
00:12:05.330 --> 00:12:08.991 The more of this accelerated aging phenotype
NOTE Confidence: 0.9159804
00:12:08.991 --> 00:12:11.820 which could predispose them to cancer.
NOTE Confidence: 0.9159804
00:12:11.820 --> 00:12:13.750 And this is actually, uhm,
NOTE Confidence: 0.9159804
00:12:13.750 --> 00:12:15.736 what we can observe when we
NOTE Confidence: 0.9159804
00:12:15.736 --> 00:12:18.464 look at this is all data from
NOTE Confidence: 0.9159804
00:12:18.464 --> 00:12:20.674 normal breast tissue from women,
NOTE Confidence: 0.9159804
00:12:20.680 --> 00:12:22.600 either with or without breast
NOTE Confidence: 0.9159804
00:12:22.600 --> 00:12:24.136 cancer prior to treatment.
NOTE Confidence: 0.9159804
00:12:24.140 --> 00:12:26.625 This is a collaboration with others at
NOTE Confidence: 0.9159804
00:12:26.625 --> 00:12:29.655 Yale and we validated this in the original
NOTE Confidence: 0.9159804
00:12:29.655 --> 00:12:32.610 study and then also in another study.
NOTE Confidence: 0.9159804
00:12:32.610 --> 00:12:36.075 Or you can see that women with breast cancer,
NOTE Confidence: 0.9159804
00:12:36.080 --> 00:12:38.150 their normal tissues seems to
be epigenetically older when we look at this Brown module. And women without breast cancer. And this is all age matched our age adjusted and adjusted for things like BMI, smoking another potential confounders. Uh, we also had a really small data set where we had, uhm, this Brown module measured in tumors and we had information on survival, so this is a data set with only 51 samples an over. I totale I are over 3471 person Montes or 20 deaths. And what you can see,
we need to validate this given...those small sample where we do...see that this Brown module 1...standard deviation increase in...this module it’s associated with...about 2.25 fold increased risk of...mortality over this time period,...and that’s adjusting for things like age,...race, ethnicity, tumor grade,...ER and also chemotherapy tree....So I went looking more specifically...that’s in this Brown module....Um, these are the individual CP...GS in the Brown module and we...can actually relate each CVG to...some of the outcomes I discussed.
So this first column is whether it differentiates in normal breast tissue, women with breast cancer versus controls.

The second column is whether it can differentiate breast tumors from normal breast tissue and the third column is the survival.

I'm finding and basically what we can see is. There's about a group of 12 CP GS for which hypermethylation so increased maculation in these 12 CP GS is associated with either cancer and normal tissue or or tumor versus normal or lower survival rate. And from the these are the jeans that
these DVD’s are in an there actually almost all in promoter regions in these jeans and we can use just ease 12 to estimate an overall score. So we use PCA across these three samples and we can take PC one of those 12 jeans and follow up with that. So the other thing is that we also find that these genius seemed to have specific characteristics, so they seem to be associated with polycomb group targets and also HT K27 trimethylation occupancy and see, and they tend to be ensued. 12 pound jeans. So this is these 12 selected jeans.
These were all the jeans that were in the original ground module and these are all the CP GS that we have measured in all of our samples. So about 20,000 CP GS over also. This is kind of the background. About 65 to 70% of them are orange juice 12 pound jeans, about 50% are Co locating with H2K27 trying Appalachian and similarly 50% with Polycom group targets. And Interestingly, this Association is actually not news, so there’s some dating back about 13 years of evidence that these
polycomb mediated methylations does seem to be important in cancer and. Basically, Polycom group proteins are involved in repression of jeans that are required for salt. A stem cell differentiation. Um, so finally we also looked at these in non breast cancers again, so this is in colorectal cancer and again we find using this 12 PPG DNA methylations score that we can significantly differentiate normal tissue from cancerous tissue. And Lastly, probably to me,
the most interesting thing is we can look at this. A trustee PG score in completely normal tissue across a bunch of different tissue types. And basically we see really strong correlations with chronological age across all of these. So in brain whole glide colon, dermis, an epidermis which to me suggests that these might be changes that are naturally occurring with aging and that might be predisposing.
I’m so something that we’re really interested now is in terms of kind of a primary or secondary prevention approach. Can you identify people who are scoring higher for their age then we would expect an are those boots are? Are people who seems to be aging faster in blood also aging? And then last, um, basically,
we also looked at this using a cultured fiberglass and basically we have, the early passage controls. We haven’t immortalized transform fiberglass where you can see an acceleration of this epigenetic score immortalized, and we also looked in cellular senescence. So on pigeon induced, in essence, an replicative senescence, and these are near near senescence that were passage together so prohibitive. But they, uh, show high snacks and story, associated beta gal. And basically what you can see is compared to the early passes cells.
We can recapitulate this.
Indies cultured fiberglass.
So In conclusion, uhm,
there are different kinds of DNA
methylation changes in aging that are
captured in the different epigenetic
clocks and by deconstructing then
we can start to understand the
functionality of the signals that
are captured in these clocks.
And specifically,
the Brown module seems particularly
interesting in terms of cancer.
Is one of the biggest shared signals
across all the epigenetic clocks and
a distinguishes tumor versus normal
in a variety of different issues.

Uh, differences to normal breasts are also observed for women with cancer versus those without, and the signal from these from the model and tumors associated with survival. We can that also narrow it down to $12.00 that are really driving the signal in this Brown module there mainly capturing promoters, TPG island hypermethylation that tend to be marked by Polycom extricate 27 trimethylation and sues 12. We can observe acceleration in culture,
fiberless, appan,

immortalization transformation

and also so there’s no sense.

But to me out again,

really interesting thing is that we

actually see linear changes in this

signal across the adult range in

a bunch of different issues which

actually might be informative.

So overall,

I think this may represent an opinion

about genetic aging change that

explains the increase cancer risk.

With that I want to acknowledge people

in my lab and also my collaborators,

both at Yale.
And elsewhere, as well as my funding.

Working, thank you.

That’s a terrific presentation in a really interesting work.

And we actually have a number of questions that have been put forth on the chat or let me just run through a few Dan Demayo ask you make see that people have recently described methylation of RNA M RNA. Specifically, does that change as well in the context of what you’ve been describing?

So we haven’t looked at that here.

I know people are looking at that, um,

there’s a group at Harvard who is actually
looking at that in terms of aging, but it for now what I’m discussing here is just CG metalation in DNA.

Um, one another question sort of. Have you looked at this in the context of progeria patients, which is sort of a really interesting question as it relates to aging, is curious if if you are folks she worked with it worked in that space and so we've looked at the overall Clock scores.

In progeria and not all of them, but some of them do show acceleration in fridge area.

We haven't looked at this specific modules
for the Brown module or the 12 PPG.

Part of the Brown module in progeria, but that is actually an interesting thing and progeria something we have plans to look at all the different modules to see if there are certain parts that are picking that up because again some clocks seem to pick up the progeria acceleration whereas others don’t.

Thank you Marcus has a question which as you can see, he said for the for the 12 CP GS that you’ve identified their individual basis as opposed to islands in
any variation of those sites.

Uhm, I actually haven't looked at whether there snips um at those sites,

so they are individual CP GS, so 12 individuals seeking.

Geez, what we're interested now is actually looking at the whole region and looking at it like variation across the regions,

but we haven't done that yet.

But yeah, I should go back and actually look at whether they're adjacent snips that would be.

One question I have is, uhm, you know. Looking at your data and realizing that beyond aging there are,
you know many sort of behaviors, environmental exposures for lack of a better phrase that drive cancer. Breast colon, certainly. And should have. Is there an opportunity to study sort of, the behavior of these individuals overtime that would drive the signature in a way that you know they are sort of. They have a greater component of that. At Methylations signature that not only is reflective of promoted aging, but increase risk of cancer. Yeah, so we can see we have UM shown in the overall Clock scores that you do get accelerated at genetic age.
in Association with things that we think of as normal risk factors, so cigarette smoking obesity I need in some socioeconomic factors seem to map onto differences in these aging rates. We haven’t looked again specifically at this module, although I will say from some of our other work, it seems like the Brown module is not particularly picking up smoking. But that might just be when measured in blood, whether it is in long or or some other samples that might be different, whereas it seems more like that purple module that it didn’t really go into.
It’s actually picking up more of those smoking, and the influence was smoking in when measured in blood.

Another question is that the methyl lation that of the 12 jeans in breast and with regarding the breast in memory you can obviously the questions you can see is that breast tissue is. A combination of various cell types and have you narrowed down sort of the epithelial, fibroblast, other cell types with regard to what you’re finding. Yeah, so unfortunately we just have bulk samples
so we can actually narrow it down to which cell type this is coming from, but I think because breast is so heterogeneous we actually the age correlation with our measures actually much weaker and breast, I think because it’s a little bit confounded by the cell composition. However, you know, part of the reason we did to follow up in the culture fiberglass was to make sure we weren’t just capturing something about cell composition changes with aging. And the other interesting thing is that at least the Brown module
00:24:57.653 --> 00:24:59.417 seems to be pretty conserved across cell and tissue types,

00:25:00.730 --> 00:25:02.906 so I don’t think it is picking up something from a specific tissue type.

00:25:02.906 --> 00:25:04.807 It would be interesting to look at epithelial versus fiberglass and see if one of those is driving the signal more than the other,

00:25:06.910 --> 00:25:10.274 but right now we don’t have that data.

00:25:15.020 --> 00:25:16.976 And then the last question before we break is if you looked at expression of of the individual jeans potentially classic tumor suppressor genes or other typical mechanisms.
I’m so that is the follow up that we’re actually doing right now, so everything I showed today is either on the first part of the talk is impressed. The second part is in progress, so it’s kind of early days still on this. But yeah, our goal is then to move to expression. We have looked at human protein at listen. Do see some associations in terms of. Answer and expression in the jeans in our 12 CG set so we are optimistic that we’ll see differential expression. Well thank you were or just now at the top of the hour and I want to thank Morgan and Marcus for two superb
talks that it really elucidated.

Gray science being conducted at our Cancer Center.

Thank you all for joining us again for virtual grand rounds and look forward again to seeing you all again next week.

Great. Thanks, thank you.