WEBVTT

NOTE duration: ”00:18:17.557”

NOTE Confidence: 0.8804754

00:00:00.240 --> 00:00:02.000 Our next speaker is a

NOTE Confidence: 0.8804754

00:00:02.000 --> 00:00:04.180 true yearly. It really epitomizes

NOTE Confidence: 0.9965292

00:00:04.640 --> 00:00:06.339 the kind of rising star,

NOTE Confidence: 0.9354704

00:00:07.120 --> 00:00:08.960 that the Adam Center is

NOTE Confidence: 0.9354704

00:00:08.960 --> 00:00:09.460 looking

NOTE Confidence: 0.96305406

00:00:09.920 --> 00:00:10.740 forward to,

NOTE Confidence: 0.98782927

00:00:11.519 --> 00:00:12.019 recruit,

NOTE Confidence: 0.94892514

00:00:12.400 --> 00:00:13.920 in in the future. So

NOTE Confidence: 0.94892514

00:00:13.920 --> 00:00:15.125 if you are interested,

NOTE Confidence: 0.9993924

00:00:15.905 --> 00:00:17.045 shoot me an email.

NOTE Confidence: 0.9515666

00:00:17.745 --> 00:00:18.725 Jason Liu,

NOTE Confidence: 0.97137845

00:00:19.585 --> 00:00:22.225 has an a bachelor's degree

NOTE Confidence: 0.97137845

00:00:22.225 --> 00:00:24.145 in applied mathematics from Yale

NOTE Confidence: 0.97137845

00:00:24.145 --> 00:00:25.904 and then received his PhD

NOTE Confidence: 0.97137845

00:00:25.904 --> 00:00:27.365 in computational biology

NOTE Confidence: 0.99627495

00:00:27.910 --> 00:00:29.370 and biomedical informatics,

NOTE Confidence: 0.9216788

00:00:29.910 --> 00:00:31.590 again from Yale, and is

NOTE Confidence: 0.9216788

00:00:31.590 --> 00:00:33.530 now a postdoctoral associate

NOTE Confidence: 0.7828878

00:00:34.070 --> 00:00:35.770 in Marc Gerstein's lab.

NOTE Confidence: 0.99193656

00:00:36.550 --> 00:00:37.050 And

NOTE Confidence: 0.9203885

00:00:37.510 --> 00:00:39.050 looking forward to your talk.

NOTE Confidence: 0.9849324

00:00:39.670 --> 00:00:40.809 Jason is

NOTE Confidence: 0.9557589

00:00:41.175 --> 00:00:42.135 connecting genomics,

NOTE Confidence: 0.9787655

00:00:43.575 --> 00:00:44.055 to,

NOTE Confidence: 0.97609895

00:00:44.455 --> 00:00:44.955 biosensors

NOTE Confidence: 0.97509223

00:00:45.415 --> 00:00:47.595 and really, opening a new

NOTE Confidence: 0.97509223

00:00:47.735 --> 00:00:49.735 new area of genetics that

NOTE Confidence: 0.97509223

00:00:49.735 --> 00:00:51.675 I think of as digital

NOTE Confidence: 0.97509223

00:00:51.815 --> 00:00:54.455 genetics or biosensor genetics, and

NOTE Confidence: 0.97509223

00:00:54.455 --> 00:00:55.835 that sort of allows

NOTE Confidence: 0.9370492

00:00:56.390 --> 00:00:57.450 to link what,

NOTE Confidence: 0.9932746

00:00:58.230 --> 00:00:59.930 this digital twin model,

NOTE Confidence: 0.96411

00:01:00.550 --> 00:01:02.950 back to biosensors in live

NOTE Confidence: 0.96411

00:01:02.950 --> 00:01:05.370 patients, which has enormous applications.

NOTE Confidence: 0.9220617

00:01:06.790 --> 00:01:08.069 Great. Thank you so much,

NOTE Confidence: 0.9220617

00:01:08.069 --> 00:01:09.530 doctor, for the introduction.

NOTE Confidence: 0.9871527

00:01:10.645 --> 00:01:11.145 And,

NOTE Confidence: 0.99278224

00:01:11.525 --> 00:01:12.485 I'm very honored to be

NOTE Confidence: 0.99278224

00:01:12.485 --> 00:01:13.365 here today to speak to

NOTE Confidence: 0.99278224

00:01:13.365 --> 00:01:14.725 you all, about some of

NOTE Confidence: 0.99278224

00:01:14.725 --> 00:01:16.565 our exciting work using AI

NOTE Confidence: 0.99278224

00:01:16.565 --> 00:01:17.305 and biosensors

NOTE Confidence: 0.99945456

00:01:17.845 --> 00:01:19.305 to drive precision medicine.

NOTE Confidence: 0.9994187

00:01:19.685 --> 00:01:21.205 So I'll just kind of

NOTE Confidence: 0.9994187

00:01:21.205 --> 00:01:22.565 give a quick introduction in

NOTE Confidence: 0.9994187

00:01:22.565 --> 00:01:23.845 terms of the terminology that

NOTE Confidence: 0.9994187

00:01:23.845 --> 00:01:24.985 I'll be using today.

NOTE Confidence: 0.98444843

00:01:25.530 --> 00:01:26.410 On the left here, you'll

NOTE Confidence: 0.98444843

00:01:26.410 --> 00:01:26.910 see,

NOTE Confidence: 0.9896021

00:01:27.210 --> 00:01:28.970 we'll refer to the clinical

NOTE Confidence: 0.9896021

00:01:28.970 --> 00:01:30.430 diagnosis as macrophenotype,

NOTE Confidence: 0.9877249

00:01:31.290 --> 00:01:32.830 this brown m here.

NOTE Confidence: 0.95930946

00:01:33.210 --> 00:01:34.410 And then, of course, we

NOTE Confidence: 0.95930946

00:01:34.410 --> 00:01:35.770 have the genotype here in

NOTE Confidence: 0.95930946

00:01:35.770 --> 00:01:37.370 the gray g. And a

NOTE Confidence: 0.95930946

00:01:37.370 --> 00:01:38.595 lot of the research here

NOTE Confidence: 0.95930946

00:01:38.595 --> 00:01:39.955 at the Adam Center and,

NOTE Confidence: 0.97855353

00:01:40.355 --> 00:01:41.555 many other people are interested

NOTE Confidence: 0.97855353

00:01:41.555 --> 00:01:42.775 in is linking together

NOTE Confidence: 0.90958047

00:01:43.315 --> 00:01:44.135 the macrophenotype

NOTE Confidence: 0.99652445

00:01:44.675 --> 00:01:46.535 to the genotype to understand

NOTE Confidence: 0.99652445

00:01:46.595 --> 00:01:48.375 the genetic architecture of disease.

NOTE Confidence: 0.97407204

00:01:49.075 --> 00:01:49.975 And, of course,

NOTE Confidence: 0.9998004

00:01:50.355 --> 00:01:51.795 this has been done quite

NOTE Confidence: 0.9998004

00:01:51.795 --> 00:01:52.295 successfully

NOTE Confidence: 0.9357612

00:01:52.920 --> 00:01:54.920 through many different case control

NOTE Confidence: 0.9357612

00:01:54.920 --> 00:01:55.420 GWAS.

NOTE Confidence: 0.9874565

00:01:55.880 --> 00:01:56.920 So just here as a

NOTE Confidence: 0.9874565

00:01:56.920 --> 00:01:57.820 couple of examples,

NOTE Confidence: 0.98692155

00:01:58.120 --> 00:01:59.260 Parkinson's GWAS,

NOTE Confidence: 0.99313116

00:01:59.800 --> 00:02:01.180 over fifty thousand cases,

NOTE Confidence: 0.8915104

00:02:01.640 --> 00:02:03.320 recently published in Nature Genetics

NOTE Confidence: 0.8915104

00:02:03.320 --> 00:02:04.460 and and also

NOTE Confidence: 0.8003236

00:02:05.000 --> 00:02:05.500 neuropsychiatric,

NOTE Confidence: 0.9738565

00:02:06.280 --> 00:02:08.065 conditions here at ADHD has

NOTE Confidence: 0.9738565

00:02:08.065 --> 00:02:09.264 been widely studied by the

NOTE Confidence: 0.9738565

00:02:09.264 --> 00:02:09.764 PGC.

NOTE Confidence: 0.9857807

00:02:10.465 --> 00:02:11.905 So today, most of my

NOTE Confidence: 0.9857807

00:02:11.905 --> 00:02:13.525 talk is gonna focus on

NOTE Confidence: 0.91490936

00:02:13.905 --> 00:02:14.405 ADHD,

NOTE Confidence: 0.9608477

00:02:15.105 --> 00:02:16.944 not directly Parkinson's disease, but

NOTE Confidence: 0.9608477

00:02:16.944 --> 00:02:17.264 I,

NOTE Confidence: 0.98085314

00:02:17.985 --> 00:02:19.264 hope to kind of show

NOTE Confidence: 0.98085314

00:02:19.264 --> 00:02:20.465 that a lot of this

NOTE Confidence: 0.98085314

00:02:20.465 --> 00:02:21.285 work is,

NOTE Confidence: 0.9992805

00:02:21.760 --> 00:02:23.060 adaptable and applicable

NOTE Confidence: 0.99937546

00:02:23.440 --> 00:02:24.720 to the research going on

NOTE Confidence: 0.99937546

00:02:24.720 --> 00:02:25.940 with Parkinson's disease.

NOTE Confidence: 0.99978715

00:02:26.480 --> 00:02:26.980 So

NOTE Confidence: 0.9734172

00:02:27.440 --> 00:02:29.200 despite all this, GWAS that

NOTE Confidence: 0.9734172

00:02:29.200 --> 00:02:30.639 has been occurring, you know,

NOTE Confidence: 0.9734172

00:02:30.639 --> 00:02:32.400 there there still is this,

NOTE Confidence: 0.9734172

00:02:32.639 --> 00:02:34.560 question of missing heritability, which

NOTE Confidence: 0.9734172

00:02:34.560 --> 00:02:35.985 is that if we look

NOTE Confidence: 0.9734172

00:02:35.985 --> 00:02:37.504 at the heritability given by

NOTE Confidence: 0.9734172

00:02:37.504 --> 00:02:39.025 a twin study, it's very

NOTE Confidence: 0.9734172

00:02:39.025 --> 00:02:40.544 high, especially for things like

NOTE Confidence: 0.9734172

00:02:40.544 --> 00:02:42.325 Parkinson's, Alzheimer's, ADHD.

NOTE Confidence: 0.99932873

00:02:42.944 --> 00:02:43.444 But

NOTE Confidence: 0.97800934

00:02:43.745 --> 00:02:45.424 what we actually capture using

NOTE Confidence: 0.97800934

00:02:45.424 --> 00:02:47.105 these case and control GWAS

NOTE Confidence: 0.97800934

00:02:47.105 --> 00:02:48.224 is only a fraction of

NOTE Confidence: 0.97800934

00:02:48.224 --> 00:02:49.709 that. And that difference, the

NOTE Confidence: 0.97800934

00:02:49.709 --> 00:02:51.150 missing heritability, is kind of

NOTE Confidence: 0.97800934

00:02:51.150 --> 00:02:52.049 what we're after.

NOTE Confidence: 0.9921632

00:02:52.349 --> 00:02:53.650 And so the question is,

NOTE Confidence: 0.9921632

00:02:53.870 --> 00:02:54.829 are there ways that we

NOTE Confidence: 0.9921632

00:02:54.829 --> 00:02:56.669 can improve this, capture more

NOTE Confidence: 0.9921632

00:02:56.669 --> 00:02:57.889 of that missing heritability?

NOTE Confidence: 0.99504864

00:02:59.470 --> 00:03:00.349 And so one of the

NOTE Confidence: 0.99504864

00:03:00.349 --> 00:03:01.915 ways, that we wanna do

NOTE Confidence: 0.99504864

00:03:01.915 --> 00:03:03.435 so is through this concept

NOTE Confidence: 0.99504864

00:03:03.435 --> 00:03:04.735 of precision phenotyping

NOTE Confidence: 0.9982287

00:03:05.035 --> 00:03:06.655 or intermediate phenotypes.

NOTE Confidence: 0.9904325

00:03:07.194 --> 00:03:08.315 So in the previous speakers,

NOTE Confidence: 0.9904325

00:03:08.315 --> 00:03:09.275 we've kind of heard already

NOTE Confidence: 0.9904325

00:03:09.275 --> 00:03:10.495 a little bit how genomics

NOTE Confidence: 0.9904325

00:03:10.555 --> 00:03:12.255 is used using, for example,

NOTE Confidence: 0.9904325

00:03:12.474 --> 00:03:13.355 single cell,

NOTE Confidence: 0.971001

00:03:13.755 --> 00:03:15.455 RNA. You can do EQTLs,

NOTE Confidence: 0.9981656

00:03:16.459 --> 00:03:17.680 to link to the genetics.

NOTE Confidence: 0.9300515

00:03:18.299 --> 00:03:19.419 But today, what I'm gonna

NOTE Confidence: 0.9300515

00:03:19.419 --> 00:03:20.780 mainly speak about is how

NOTE Confidence: 0.9300515

00:03:20.780 --> 00:03:22.639 we can use digital technology

NOTE Confidence: 0.9300515

00:03:22.939 --> 00:03:24.540 and in particular, wearable and

NOTE Confidence: 0.9300515

00:03:24.540 --> 00:03:25.040 smartwatches.

NOTE Confidence: 0.9885727

00:03:25.980 --> 00:03:26.939 There's been kind of an

NOTE Confidence: 0.9885727

00:03:26.939 --> 00:03:28.159 emergence of this,

NOTE Confidence: 0.9981148

00:03:28.565 --> 00:03:29.865 the popularity of smartwatches.

NOTE Confidence: 0.9896242

00:03:30.245 --> 00:03:31.445 Many, many people have it.

NOTE Confidence: 0.9896242

00:03:31.445 --> 00:03:32.585 It's more accessible

NOTE Confidence: 0.99286604

00:03:32.965 --> 00:03:34.025 in terms of cost.

NOTE Confidence: 0.97975636

00:03:34.325 --> 00:03:35.945 And so the question is,

NOTE Confidence: 0.97975636

00:03:36.085 --> 00:03:37.705 if we use the information

NOTE Confidence: 0.99419606

00:03:38.085 --> 00:03:39.705 captured by the smartwatch,

NOTE Confidence: 0.9535028

00:03:40.085 --> 00:03:41.625 can we first, number one,

NOTE Confidence: 0.9535028

00:03:41.685 --> 00:03:42.725 link it to the macro

NOTE Confidence: 0.9535028

00:03:42.725 --> 00:03:44.519 phenotype or the disease, gain

NOTE Confidence: 0.9535028

00:03:44.519 --> 00:03:46.060 some sort of clinical insight?

NOTE Confidence: 0.97588634

00:03:46.439 --> 00:03:47.640 And then second of all,

NOTE Confidence: 0.97588634

00:03:47.640 --> 00:03:48.519 if we can do that

NOTE Confidence: 0.97588634

00:03:48.519 --> 00:03:49.019 successfully,

NOTE Confidence: 0.99559075

00:03:49.319 --> 00:03:50.920 can we then relink it

NOTE Confidence: 0.99559075

00:03:50.920 --> 00:03:51.980 back to the genotype

NOTE Confidence: 0.9899207

00:03:52.359 --> 00:03:54.299 to hopefully gain more statistical

NOTE Confidence: 0.9899207

00:03:54.439 --> 00:03:55.719 power in terms of our

NOTE Confidence: 0.9899207

00:03:55.719 --> 00:03:57.739 ability for genetic discovery?

NOTE Confidence: 0.9970253

00:04:00.815 --> 00:04:01.615 So for the rest of

NOTE Confidence: 0.9970253

00:04:01.615 --> 00:04:02.975 the talk, I'm gonna, kind

NOTE Confidence: 0.9970253

00:04:02.975 --> 00:04:04.035 of give an overview

NOTE Confidence: 0.98737293

00:04:04.335 --> 00:04:05.695 into one of our recent,

NOTE Confidence: 0.98737293

00:04:05.935 --> 00:04:06.435 projects.

NOTE Confidence: 0.96852595

00:04:07.055 --> 00:04:09.215 This is, digital phenotyping with

NOTE Confidence: 0.96852595

00:04:09.215 --> 00:04:09.715 wearables,

NOTE Confidence: 0.9754817

00:04:10.095 --> 00:04:10.995 for a cohort,

NOTE Confidence: 0.99309355

00:04:11.295 --> 00:04:12.355 known as the ABCD,

NOTE Confidence: 0.9269087

00:04:12.990 --> 00:04:15.410 adolescent brain cognitive development study.

NOTE Confidence: 0.9445888

00:04:15.950 --> 00:04:17.390 And, it was recently published

NOTE Confidence: 0.9445888

00:04:17.390 --> 00:04:18.690 in Cell. And,

NOTE Confidence: 0.9723137

00:04:19.310 --> 00:04:20.430 so just the kind of

NOTE Confidence: 0.9723137

00:04:20.430 --> 00:04:21.550 a high level overview of

NOTE Confidence: 0.9723137

00:04:21.550 --> 00:04:22.910 the data, it's a few

NOTE Confidence: 0.9723137

00:04:22.910 --> 00:04:25.490 thousand individuals with psychiatric diagnosis.

NOTE Confidence: 0.9857246

00:04:26.029 --> 00:04:27.470 They also have digital data.

NOTE Confidence: 0.9857246

00:04:27.470 --> 00:04:28.705 In this case, it's Fitbit

NOTE Confidence: 0.9857246

00:04:28.765 --> 00:04:29.265 smartwatches.

NOTE Confidence: 0.9674689

00:04:30.045 --> 00:04:31.404 And then finally, we have,

NOTE Confidence: 0.9674689

00:04:31.725 --> 00:04:33.085 genetic information for all of

NOTE Confidence: 0.9674689

00:04:33.085 --> 00:04:33.904 these individuals.

NOTE Confidence: 0.98427284

00:04:37.085 --> 00:04:39.245 And so oftentimes, the question

NOTE Confidence: 0.98427284

00:04:39.245 --> 00:04:40.525 that people ask is, well,

NOTE Confidence: 0.98427284

00:04:40.525 --> 00:04:41.565 what does that data look

NOTE Confidence: 0.98427284

00:04:41.565 --> 00:04:43.005 like? What does Fitbit data

NOTE Confidence: 0.98427284

00:04:43.005 --> 00:04:44.500 look like? So just on

NOTE Confidence: 0.98427284

00:04:44.500 --> 00:04:45.560 the left here, I'm

NOTE Confidence: 0.97260904

00:04:45.940 --> 00:04:47.240 showing, three individuals,

NOTE Confidence: 0.99738365

00:04:47.860 --> 00:04:50.020 their signal tracks across a

NOTE Confidence: 0.99738365

00:04:50.020 --> 00:04:51.400 variety of different modalities.

NOTE Confidence: 0.9930727

00:04:51.779 --> 00:04:53.060 So for these three individuals,

NOTE Confidence: 0.9930727

00:04:53.060 --> 00:04:53.940 we can see their heart

NOTE Confidence: 0.9930727

00:04:53.940 --> 00:04:56.339 rate, their calories, activity, steps,

NOTE Confidence: 0.9930727

00:04:56.339 --> 00:04:57.460 so on and so forth.

NOTE Confidence: 0.9930727

00:04:57.460 --> 00:04:58.180 These are things that,

NOTE Confidence: 0.97203016

00:04:59.224 --> 00:05:00.025 probably if you have a

NOTE Confidence: 0.97203016

00:05:00.025 --> 00:05:01.625 smartwatch, then you can also,

NOTE Confidence: 0.97203016

00:05:01.944 --> 00:05:02.925 look at as well.

NOTE Confidence: 0.9677412

00:05:03.384 --> 00:05:04.824 But, you know, this data

NOTE Confidence: 0.9677412

00:05:04.824 --> 00:05:06.104 is very noisy, and it

NOTE Confidence: 0.9677412

00:05:06.104 --> 00:05:06.604 has,

NOTE Confidence: 0.99106467

00:05:07.464 --> 00:05:08.504 problems with it in terms

NOTE Confidence: 0.99106467

00:05:08.504 --> 00:05:09.865 of downstream analysis. So how

NOTE Confidence: 0.99106467

00:05:09.865 --> 00:05:11.065 how can we prepare or

NOTE Confidence: 0.99106467

00:05:11.065 --> 00:05:12.820 process this data in a

NOTE Confidence: 0.99106467

00:05:12.820 --> 00:05:14.260 way that makes it suitable

NOTE Confidence: 0.99106467

00:05:14.260 --> 00:05:15.480 for downstream analysis?

NOTE Confidence: 0.9908884

00:05:16.020 --> 00:05:17.700 Well, one of the, very

NOTE Confidence: 0.9908884

00:05:17.700 --> 00:05:19.320 straightforward ways to do so

NOTE Confidence: 0.9908884

00:05:19.460 --> 00:05:21.140 is to summarize the data.

NOTE Confidence: 0.9908884

00:05:21.140 --> 00:05:22.340 So if I have somebody's

NOTE Confidence: 0.9908884

00:05:22.340 --> 00:05:23.620 heart rate, it changes across

NOTE Confidence: 0.9908884

00:05:23.620 --> 00:05:24.660 the day, I can do

NOTE Confidence: 0.9908884

00:05:24.660 --> 00:05:25.940 something very simple, which is

NOTE Confidence: 0.9908884

00:05:25.940 --> 00:05:27.285 to say, just take take

NOTE Confidence: 0.9908884

00:05:27.285 --> 00:05:28.404 the mean heart rate during

NOTE Confidence: 0.9908884

00:05:28.404 --> 00:05:29.365 the day or the mean

NOTE Confidence: 0.9908884

00:05:29.365 --> 00:05:30.805 heart rate at night. And

NOTE Confidence: 0.9908884

00:05:30.805 --> 00:05:31.605 we can look at a

NOTE Confidence: 0.9908884

00:05:31.605 --> 00:05:33.545 variety of different statistical measures

NOTE Confidence: 0.9908884

00:05:33.765 --> 00:05:35.125 and then summarize them into

NOTE Confidence: 0.9908884

00:05:35.125 --> 00:05:36.725 what we're calling the static

NOTE Confidence: 0.9908884

00:05:36.725 --> 00:05:37.225 features.

NOTE Confidence: 0.98353285

00:05:37.845 --> 00:05:39.285 We call them static features

NOTE Confidence: 0.98353285

00:05:39.285 --> 00:05:39.785 because,

NOTE Confidence: 0.98884106

00:05:40.509 --> 00:05:41.710 by nature of being a

NOTE Confidence: 0.98884106

00:05:41.710 --> 00:05:43.089 time series and then summarized,

NOTE Confidence: 0.9742051

00:05:43.469 --> 00:05:44.669 we lose some of that

NOTE Confidence: 0.9742051

00:05:44.669 --> 00:05:45.729 temporal resolution.

NOTE Confidence: 0.9853584

00:05:46.430 --> 00:05:47.550 But the good thing about

NOTE Confidence: 0.9853584

00:05:47.550 --> 00:05:49.389 the static features is they're

NOTE Confidence: 0.9853584

00:05:49.389 --> 00:05:50.589 very easy to understand. They're

NOTE Confidence: 0.9853584

00:05:50.589 --> 00:05:51.550 easy to work with. We

NOTE Confidence: 0.9853584

00:05:51.550 --> 00:05:53.409 can make this nice matrix

NOTE Confidence: 0.9853584

00:05:53.469 --> 00:05:55.009 of individuals by features,

NOTE Confidence: 0.99066126

00:05:55.585 --> 00:05:56.705 and then we can attach

NOTE Confidence: 0.99066126

00:05:56.705 --> 00:05:58.004 on a set of covariates

NOTE Confidence: 0.99066126

00:05:58.065 --> 00:05:59.444 that we would typically use

NOTE Confidence: 0.99066126

00:05:59.504 --> 00:06:00.724 for these types of studies.

NOTE Confidence: 0.99353755

00:06:01.345 --> 00:06:03.104 And so the static features

NOTE Confidence: 0.99353755

00:06:03.104 --> 00:06:03.904 are are one of the

NOTE Confidence: 0.99353755

00:06:03.904 --> 00:06:05.104 ways that we're gonna process

NOTE Confidence: 0.99353755

00:06:05.104 --> 00:06:05.764 the data.

NOTE Confidence: 0.9924671

00:06:06.464 --> 00:06:08.004 But as I just mentioned,

NOTE Confidence: 0.97504646

00:06:08.650 --> 00:06:10.250 the static features, we lose

NOTE Confidence: 0.97504646

00:06:10.250 --> 00:06:11.310 some of the temporal

NOTE Confidence: 0.98734283

00:06:11.690 --> 00:06:12.570 dynamics, meaning,

NOTE Confidence: 0.99920475

00:06:13.529 --> 00:06:15.529 how somebody is changing behaviorally

NOTE Confidence: 0.99920475

00:06:15.529 --> 00:06:16.270 or physiologically

NOTE Confidence: 0.9749379

00:06:16.890 --> 00:06:18.170 on a seconds or minute

NOTE Confidence: 0.9749379

00:06:18.170 --> 00:06:20.089 level. And so in order

NOTE Confidence: 0.9749379

00:06:20.089 --> 00:06:21.435 to preserve some of that,

NOTE Confidence: 0.9749379

00:06:21.675 --> 00:06:22.955 we still want to kind

NOTE Confidence: 0.9749379

00:06:22.955 --> 00:06:23.435 of,

NOTE Confidence: 0.9868336

00:06:23.995 --> 00:06:25.275 move to some sort of

NOTE Confidence: 0.9868336

00:06:25.275 --> 00:06:27.195 feature set that is temporally

NOTE Confidence: 0.9868336

00:06:27.195 --> 00:06:29.355 resolved. And, we're calling that

NOTE Confidence: 0.9868336

00:06:29.355 --> 00:06:31.195 temporally resolved feature set the

NOTE Confidence: 0.9868336

00:06:31.195 --> 00:06:32.175 dynamic features.

NOTE Confidence: 0.9840396

00:06:32.794 --> 00:06:33.755 And there's a variety of

NOTE Confidence: 0.9840396

00:06:33.755 --> 00:06:34.680 steps that we,

NOTE Confidence: 0.97372335

00:06:35.240 --> 00:06:36.600 perform on the raw data

NOTE Confidence: 0.97372335

00:06:36.600 --> 00:06:37.800 to achieve that,

NOTE Confidence: 0.99457175

00:06:38.199 --> 00:06:39.720 dynamic features, and I'll just

NOTE Confidence: 0.99457175

00:06:39.720 --> 00:06:40.680 briefly go over some of

NOTE Confidence: 0.99457175

00:06:40.680 --> 00:06:41.740 those steps now.

NOTE Confidence: 0.98826116

00:06:42.279 --> 00:06:44.199 So first of all, one

NOTE Confidence: 0.98826116

00:06:44.199 --> 00:06:45.560 of the challenges of this,

NOTE Confidence: 0.98826116

00:06:45.720 --> 00:06:47.400 wearable data is everyone's data

NOTE Confidence: 0.98826116

00:06:47.400 --> 00:06:48.919 looks very different. It's collected

NOTE Confidence: 0.98826116

00:06:48.919 --> 00:06:50.445 at different times. And so

NOTE Confidence: 0.98826116

00:06:50.445 --> 00:06:51.885 in the first step, what

NOTE Confidence: 0.98826116

00:06:51.885 --> 00:06:53.645 we're really interested is in

NOTE Confidence: 0.98826116

00:06:53.645 --> 00:06:54.785 aligning individuals

NOTE Confidence: 0.9817915

00:06:55.485 --> 00:06:57.425 across different days, seasonalities.

NOTE Confidence: 0.99296045

00:06:58.604 --> 00:07:00.445 And after we've aligned it,

NOTE Confidence: 0.99296045

00:07:00.445 --> 00:07:01.565 then we want to take,

NOTE Confidence: 0.99296045

00:07:01.885 --> 00:07:03.885 optimal window selection or a

NOTE Confidence: 0.99296045

00:07:03.885 --> 00:07:05.040 slice of that data.

NOTE Confidence: 0.9711807

00:07:05.440 --> 00:07:07.060 Some individuals, they may have

NOTE Confidence: 0.9711807

00:07:07.279 --> 00:07:08.720 three weeks of data. Others

NOTE Confidence: 0.9711807

00:07:08.720 --> 00:07:10.160 might just have a few

NOTE Confidence: 0.9711807

00:07:10.160 --> 00:07:11.360 days of data. And it's

NOTE Confidence: 0.9711807

00:07:11.440 --> 00:07:12.640 so it's important that we're

NOTE Confidence: 0.9711807

00:07:12.640 --> 00:07:13.680 kind of making a fair

NOTE Confidence: 0.9711807

00:07:13.680 --> 00:07:15.600 comparison, and and we have

NOTE Confidence: 0.9711807

00:07:15.600 --> 00:07:17.700 this empirical optimal window selection.

NOTE Confidence: 0.9711807

00:07:17.884 --> 00:07:19.085 And that results, on the

NOTE Confidence: 0.9711807

00:07:19.085 --> 00:07:20.125 right hand side, you can

NOTE Confidence: 0.9711807

00:07:20.125 --> 00:07:22.365 see, with about sixty seven

NOTE Confidence: 0.9711807

00:07:22.365 --> 00:07:24.544 percent of the data remaining,

NOTE Confidence: 0.99474907

00:07:25.245 --> 00:07:26.365 and for kind of over

NOTE Confidence: 0.99474907

00:07:26.365 --> 00:07:27.805 two thousand individuals. You can

NOTE Confidence: 0.99474907

00:07:27.805 --> 00:07:28.525 see if we if we

NOTE Confidence: 0.99474907

00:07:28.525 --> 00:07:29.745 go to a higher threshold

NOTE Confidence: 0.99474907

00:07:29.805 --> 00:07:30.544 of inclusion,

NOTE Confidence: 0.9898975

00:07:31.005 --> 00:07:32.125 the sample size on the

NOTE Confidence: 0.9898975

00:07:32.125 --> 00:07:33.490 y axis drops quite significantly.

NOTE Confidence: 0.9898975

00:07:33.490 --> 00:07:35.250 So this was, kind of

NOTE Confidence: 0.9898975

00:07:35.250 --> 00:07:36.550 an empirically derived,

NOTE Confidence: 0.99971133

00:07:37.010 --> 00:07:38.230 optimal window selection.

NOTE Confidence: 0.9984687

00:07:39.970 --> 00:07:41.090 And then after we've done

NOTE Confidence: 0.9984687

00:07:41.090 --> 00:07:42.470 the optimal window selection,

NOTE Confidence: 0.9693828

00:07:43.010 --> 00:07:44.370 of course, there's still some

NOTE Confidence: 0.9693828

00:07:44.370 --> 00:07:45.729 small amounts of missing data,

NOTE Confidence: 0.9693828

00:07:45.729 --> 00:07:46.930 and we do a very

NOTE Confidence: 0.9693828

00:07:46.930 --> 00:07:49.325 simple linear imputation here. But

NOTE Confidence: 0.9693828

00:07:49.325 --> 00:07:50.445 in addition to the linear

NOTE Confidence: 0.9693828

00:07:50.445 --> 00:07:50.945 imputation,

NOTE Confidence: 0.9955071

00:07:51.405 --> 00:07:53.665 we're also tracking which variables

NOTE Confidence: 0.9955071

00:07:53.725 --> 00:07:55.165 at what time points were

NOTE Confidence: 0.9955071

00:07:55.165 --> 00:07:57.265 actually imputed and which ones

NOTE Confidence: 0.9955071

00:07:57.485 --> 00:07:58.625 are actually observed.

NOTE Confidence: 0.9998726

00:07:58.925 --> 00:07:59.965 And the point of that

NOTE Confidence: 0.9998726

00:07:59.965 --> 00:08:01.425 is to let downstream

NOTE Confidence: 0.99549925

00:08:01.965 --> 00:08:03.425 AI and machine learning models

NOTE Confidence: 0.99549925

00:08:03.565 --> 00:08:05.660 actually decide how reliable these

NOTE Confidence: 0.99549925

00:08:05.660 --> 00:08:06.960 imputed values are.

NOTE Confidence: 0.98749375

00:08:07.660 --> 00:08:09.340 And this kind of process

NOTE Confidence: 0.98749375

00:08:09.340 --> 00:08:10.699 in aggregate gives us what's

NOTE Confidence: 0.98749375

00:08:10.699 --> 00:08:12.320 known as the dynamic features.

NOTE Confidence: 0.9900744

00:08:12.780 --> 00:08:13.979 So we have static features

NOTE Confidence: 0.9900744

00:08:13.979 --> 00:08:15.260 and dynamic features. This is

NOTE Confidence: 0.9900744

00:08:15.260 --> 00:08:17.419 how we've processed smartwatch data

NOTE Confidence: 0.9900744

00:08:17.419 --> 00:08:18.220 to build kind of a

NOTE Confidence: 0.9900744

00:08:18.220 --> 00:08:19.120 digital phenotype.

NOTE Confidence: 0.97065324

00:08:19.525 --> 00:08:20.565 And the next question is,

NOTE Confidence: 0.97065324

00:08:20.565 --> 00:08:21.365 how how do we use

NOTE Confidence: 0.97065324

00:08:21.365 --> 00:08:22.664 it in in modeling?

NOTE Confidence: 0.97557724

00:08:24.005 --> 00:08:25.365 So for the static features,

NOTE Confidence: 0.97557724

00:08:25.365 --> 00:08:26.724 it's it's quite straightforward. You

NOTE Confidence: 0.97557724

00:08:26.724 --> 00:08:28.324 have a a matrix of

NOTE Confidence: 0.97557724

00:08:28.324 --> 00:08:29.705 individuals by features,

NOTE Confidence: 0.96769714

00:08:30.164 --> 00:08:31.865 and we can use traditional

NOTE Confidence: 0.96769714

00:08:31.925 --> 00:08:32.745 machine learning,

NOTE Confidence: 0.96837276

00:08:33.370 --> 00:08:34.809 models here. In this case,

NOTE Confidence: 0.96837276

00:08:34.809 --> 00:08:37.309 we've, tested XGBoost and RandomForce,

NOTE Confidence: 0.96837276

00:08:37.370 --> 00:08:38.750 which are quite popular,

NOTE Confidence: 0.9986525

00:08:39.370 --> 00:08:39.870 machine

NOTE Confidence: 0.934721

00:08:40.250 --> 00:08:41.309 machine learning models.

NOTE Confidence: 0.97139865

00:08:41.769 --> 00:08:43.050 And the idea is, can

NOTE Confidence: 0.97139865

00:08:43.050 --> 00:08:44.750 we use those static features

NOTE Confidence: 0.97139865

00:08:44.889 --> 00:08:45.550 to then,

NOTE Confidence: 0.9512733

00:08:45.929 --> 00:08:48.029 predict individuals with a particular,

NOTE Confidence: 0.9874702

00:08:48.795 --> 00:08:49.995 disease or if they're a

NOTE Confidence: 0.9874702

00:08:49.995 --> 00:08:50.815 healthy control.

NOTE Confidence: 0.99726456

00:08:51.515 --> 00:08:52.815 And one of the byproducts

NOTE Confidence: 0.99726456

00:08:52.875 --> 00:08:54.235 of doing machine learning like

NOTE Confidence: 0.99726456

00:08:54.235 --> 00:08:54.975 this is,

NOTE Confidence: 0.9625291

00:08:55.275 --> 00:08:56.155 you can see I've I've

NOTE Confidence: 0.9625291

00:08:56.155 --> 00:08:57.755 kind of marked here the,

NOTE Confidence: 0.9625291

00:08:58.235 --> 00:09:00.735 green d score, digital phenotyping

NOTE Confidence: 0.9625291

00:09:00.875 --> 00:09:02.830 score, or an AI generated

NOTE Confidence: 0.9625291

00:09:02.890 --> 00:09:04.410 risk score, it is to

NOTE Confidence: 0.9625291

00:09:04.410 --> 00:09:05.870 move away from binary

NOTE Confidence: 0.9996779

00:09:06.250 --> 00:09:06.750 definitions

NOTE Confidence: 0.99871475

00:09:07.210 --> 00:09:07.950 of disease.

NOTE Confidence: 0.978602

00:09:08.490 --> 00:09:09.690 We're very used to using

NOTE Confidence: 0.978602

00:09:09.690 --> 00:09:11.630 zeros and ones to identify

NOTE Confidence: 0.978602

00:09:11.690 --> 00:09:12.730 people who have or do

NOTE Confidence: 0.978602

00:09:12.730 --> 00:09:14.105 not have a disease, But

NOTE Confidence: 0.978602

00:09:14.105 --> 00:09:15.385 this is, kind of one

NOTE Confidence: 0.978602

00:09:15.385 --> 00:09:16.425 of the byproducts of using

NOTE Confidence: 0.978602

00:09:16.425 --> 00:09:17.704 machine learning models is we

NOTE Confidence: 0.978602

00:09:17.704 --> 00:09:18.665 can move towards kind of

NOTE Confidence: 0.978602

00:09:18.665 --> 00:09:20.125 a continuum or a spectrum,

NOTE Confidence: 0.9960369

00:09:20.745 --> 00:09:21.865 and that can be more

NOTE Confidence: 0.9960369

00:09:21.865 --> 00:09:24.105 inclusive and, more precise in

NOTE Confidence: 0.9960369

00:09:24.105 --> 00:09:25.485 terms of defining individuals.

NOTE Confidence: 0.9866208

00:09:27.660 --> 00:09:28.620 Now in terms of the

NOTE Confidence: 0.9866208

00:09:28.620 --> 00:09:30.399 dynamic features, those are

NOTE Confidence: 0.98836994

00:09:30.779 --> 00:09:31.760 a time series,

NOTE Confidence: 0.9786421

00:09:32.300 --> 00:09:33.500 and they're a little bit

NOTE Confidence: 0.9786421

00:09:33.500 --> 00:09:34.620 different. We we can't use

NOTE Confidence: 0.9786421

00:09:34.620 --> 00:09:36.220 a traditional machine learning model

NOTE Confidence: 0.9786421

00:09:36.220 --> 00:09:37.660 to deal with them. And

NOTE Confidence: 0.9786421

00:09:37.660 --> 00:09:39.579 so instead, what we've adopted

NOTE Confidence: 0.9786421

00:09:39.579 --> 00:09:41.179 here is a convolutional neural

NOTE Confidence: 0.9786421

00:09:41.179 --> 00:09:42.139 net. So this is a

NOTE Confidence: 0.9786421

00:09:42.139 --> 00:09:43.245 a deep learning model.

NOTE Confidence: 0.99402

00:09:43.804 --> 00:09:44.684 And on the left side

NOTE Confidence: 0.99402

00:09:44.684 --> 00:09:46.125 here, you'll see those dynamic

NOTE Confidence: 0.99402

00:09:46.125 --> 00:09:47.245 features. We we treat them

NOTE Confidence: 0.99402

00:09:47.245 --> 00:09:49.165 actually very similar to an

NOTE Confidence: 0.99402

00:09:49.165 --> 00:09:49.665 image.

NOTE Confidence: 0.9514907

00:09:50.125 --> 00:09:51.585 You have many different channels,

NOTE Confidence: 0.9607034

00:09:52.045 --> 00:09:53.985 across different time points.

NOTE Confidence: 0.96956635

00:09:54.525 --> 00:09:56.525 And, very similar to, like,

NOTE Confidence: 0.96956635

00:09:56.525 --> 00:09:58.065 a image classification task,

NOTE Confidence: 0.9761034

00:09:58.429 --> 00:09:59.949 we would want to use

NOTE Confidence: 0.9761034

00:09:59.949 --> 00:10:01.970 this stack of digital phenotypes

NOTE Confidence: 0.9761034

00:10:02.269 --> 00:10:03.630 to, again, predict the macro

NOTE Confidence: 0.9761034

00:10:03.630 --> 00:10:04.130 phenotype.

NOTE Confidence: 0.99964064

00:10:04.589 --> 00:10:06.269 Again, we're able to generate

NOTE Confidence: 0.99964064

00:10:06.269 --> 00:10:06.769 this

NOTE Confidence: 0.969058

00:10:07.149 --> 00:10:09.069 continuous based risk score to

NOTE Confidence: 0.969058

00:10:09.069 --> 00:10:11.630 hopefully more precisely characterize an

NOTE Confidence: 0.969058

00:10:11.630 --> 00:10:12.130 individual.

NOTE Confidence: 0.98428065

00:10:12.995 --> 00:10:14.755 And and one technical detail

NOTE Confidence: 0.98428065

00:10:14.755 --> 00:10:16.215 that I'll just highlight here

NOTE Confidence: 0.98428065

00:10:16.434 --> 00:10:17.795 is the use of a

NOTE Confidence: 0.98428065

00:10:17.795 --> 00:10:20.434 variable size convolutional filter here

NOTE Confidence: 0.98428065

00:10:20.434 --> 00:10:21.795 in the bottom. And and

NOTE Confidence: 0.98428065

00:10:21.795 --> 00:10:22.915 the idea of the variable

NOTE Confidence: 0.98428065

00:10:22.915 --> 00:10:25.795 size convolutional filter is behavioral

NOTE Confidence: 0.98428065

00:10:25.795 --> 00:10:27.255 and physiological changes

NOTE Confidence: 0.9758035

00:10:27.740 --> 00:10:29.519 may occur on a minute

NOTE Confidence: 0.9758035

00:10:29.660 --> 00:10:30.940 or a very small time

NOTE Confidence: 0.9758035

00:10:30.940 --> 00:10:32.779 scale, and we're definitely interested

NOTE Confidence: 0.9758035

00:10:32.779 --> 00:10:33.600 in those changes.

NOTE Confidence: 0.9980628

00:10:33.980 --> 00:10:35.339 But we're also interested in

NOTE Confidence: 0.9980628

00:10:35.339 --> 00:10:37.100 more global changes that perhaps

NOTE Confidence: 0.9980628

00:10:37.100 --> 00:10:38.220 occur on the day or

NOTE Confidence: 0.9980628

00:10:38.220 --> 00:10:39.279 the weekly level.

NOTE Confidence: 0.981792

00:10:39.595 --> 00:10:40.875 And so by using this

NOTE Confidence: 0.981792

00:10:40.875 --> 00:10:43.035 variable size convolutional filter, we're

NOTE Confidence: 0.981792

00:10:43.035 --> 00:10:44.715 actually able to capture more

NOTE Confidence: 0.981792

00:10:44.715 --> 00:10:46.335 of those behavioral and physiological

NOTE Confidence: 0.981792

00:10:46.554 --> 00:10:48.575 changes at different time scale.

NOTE Confidence: 0.9646418

00:10:50.395 --> 00:10:51.755 So what what are the

NOTE Confidence: 0.9646418

00:10:51.755 --> 00:10:53.530 main results or kind of,

NOTE Confidence: 0.9646418

00:10:53.809 --> 00:10:55.330 findings of using these models

NOTE Confidence: 0.9646418

00:10:55.330 --> 00:10:56.150 in this data?

NOTE Confidence: 0.964383

00:10:56.610 --> 00:10:58.130 So, again, we're here look

NOTE Confidence: 0.964383

00:10:58.210 --> 00:10:59.750 not looking exactly at Parkinson's.

NOTE Confidence: 0.964383

00:10:59.890 --> 00:11:01.170 We're looking at ADHD and

NOTE Confidence: 0.964383

00:11:01.170 --> 00:11:03.250 anxiety disorder, but, again, many,

NOTE Confidence: 0.99632114

00:11:03.650 --> 00:11:05.190 applications and extensions

NOTE Confidence: 0.9567792

00:11:05.654 --> 00:11:06.394 to Parkinson's.

NOTE Confidence: 0.995418

00:11:07.255 --> 00:11:08.054 So in in the case

NOTE Confidence: 0.995418

00:11:08.054 --> 00:11:09.495 of ADHD, the top row

NOTE Confidence: 0.995418

00:11:09.495 --> 00:11:10.154 in blue,

NOTE Confidence: 0.98373634

00:11:10.855 --> 00:11:12.695 and and anxiety disorder, in

NOTE Confidence: 0.98373634

00:11:12.695 --> 00:11:13.735 the first column here, we

NOTE Confidence: 0.98373634

00:11:13.735 --> 00:11:15.415 have three different box plots.

NOTE Confidence: 0.98373634

00:11:15.415 --> 00:11:16.375 And and these are showing

NOTE Confidence: 0.98373634

00:11:16.375 --> 00:11:18.394 the accuracy of identifying individuals

NOTE Confidence: 0.98373634

00:11:18.615 --> 00:11:19.915 with or without the disease.

NOTE Confidence: 0.98373634

00:11:20.214 --> 00:11:21.390 And we can see that

NOTE Confidence: 0.98190963

00:11:22.429 --> 00:11:23.570 and we can see that

NOTE Confidence: 0.93687516

00:11:23.950 --> 00:11:25.470 the baseline model at the

NOTE Confidence: 0.93687516

00:11:25.470 --> 00:11:26.830 first model is without any

NOTE Confidence: 0.93687516

00:11:26.830 --> 00:11:27.809 wearable information.

NOTE Confidence: 0.96866435

00:11:28.190 --> 00:11:29.630 The second model is using

NOTE Confidence: 0.96866435

00:11:29.630 --> 00:11:31.309 the static features, and then

NOTE Confidence: 0.96866435

00:11:31.309 --> 00:11:32.429 the third model is using

NOTE Confidence: 0.96866435

00:11:32.429 --> 00:11:34.040 actually those deep learning based,

NOTE Confidence: 0.95947593

00:11:35.065 --> 00:11:36.824 features, the time series. And

NOTE Confidence: 0.95947593

00:11:36.824 --> 00:11:37.625 we can see the a

NOTE Confidence: 0.95947593

00:11:37.625 --> 00:11:38.824 kind of a increase in

NOTE Confidence: 0.95947593

00:11:38.824 --> 00:11:40.105 performance and accuracy of the

NOTE Confidence: 0.95947593

00:11:40.105 --> 00:11:41.625 model as we incorporate more

NOTE Confidence: 0.95947593

00:11:41.625 --> 00:11:43.804 of those temporally resolved features.

NOTE Confidence: 0.9995104

00:11:44.745 --> 00:11:46.024 But in addition to just

NOTE Confidence: 0.9995104

00:11:46.024 --> 00:11:47.165 improving the accuracy,

NOTE Confidence: 0.9955177

00:11:47.570 --> 00:11:49.190 we're also able to identify

NOTE Confidence: 0.9955177

00:11:49.330 --> 00:11:50.709 what are the the actual

NOTE Confidence: 0.9955177

00:11:50.770 --> 00:11:52.610 physiological features that drive that

NOTE Confidence: 0.9955177

00:11:52.610 --> 00:11:53.110 prediction.

NOTE Confidence: 0.9991849

00:11:53.649 --> 00:11:54.149 So

NOTE Confidence: 0.86389184

00:11:54.929 --> 00:11:56.149 for example, in ADHD,

NOTE Confidence: 0.98884994

00:11:56.529 --> 00:11:57.570 we find that heart rate

NOTE Confidence: 0.98884994

00:11:57.570 --> 00:11:58.850 is really kind of the

NOTE Confidence: 0.98884994

00:11:58.850 --> 00:11:59.750 the key driver,

NOTE Confidence: 0.97123176

00:12:00.195 --> 00:12:02.035 but in anxiety disorder, sleep

NOTE Confidence: 0.97123176

00:12:02.035 --> 00:12:02.855 quality is.

NOTE Confidence: 0.95720387

00:12:03.315 --> 00:12:04.675 And not only that, we're

NOTE Confidence: 0.95720387

00:12:04.675 --> 00:12:06.915 able to temporarily resolve that

NOTE Confidence: 0.95720387

00:12:06.915 --> 00:12:07.415 importance.

NOTE Confidence: 0.9828557

00:12:07.795 --> 00:12:09.635 Meaning, in in this curve

NOTE Confidence: 0.9828557

00:12:09.635 --> 00:12:11.554 here, we're showing where was

NOTE Confidence: 0.9828557

00:12:11.554 --> 00:12:13.040 heart rate important. Is heart

NOTE Confidence: 0.9828557

00:12:13.040 --> 00:12:14.079 rate important all the time

NOTE Confidence: 0.9828557

00:12:14.079 --> 00:12:14.720 or just some of the

NOTE Confidence: 0.9828557

00:12:14.720 --> 00:12:15.760 time? And and we're able

NOTE Confidence: 0.9828557

00:12:15.760 --> 00:12:17.040 to temporarily resolve the heart

NOTE Confidence: 0.9828557

00:12:17.040 --> 00:12:18.079 rate importance to kind of

NOTE Confidence: 0.9828557

00:12:18.079 --> 00:12:20.500 the early to, late afternoon.

NOTE Confidence: 0.96739477

00:12:20.880 --> 00:12:22.160 And then, of course, for

NOTE Confidence: 0.96739477

00:12:22.160 --> 00:12:23.440 sleep, it's during the night.

NOTE Confidence: 0.96739477

00:12:23.440 --> 00:12:24.880 But in particular, you can

NOTE Confidence: 0.96739477

00:12:24.880 --> 00:12:26.559 see a peak there around

NOTE Confidence: 0.96739477

00:12:26.559 --> 00:12:27.600 five AM. And and, again,

NOTE Confidence: 0.96739477

00:12:27.600 --> 00:12:29.745 these are adolescents with anxiety

NOTE Confidence: 0.96739477

00:12:29.804 --> 00:12:31.165 disorder, and so we're kind

NOTE Confidence: 0.96739477

00:12:31.165 --> 00:12:32.285 of showing that perhaps there's

NOTE Confidence: 0.96739477

00:12:32.285 --> 00:12:33.105 sleep disturbance

NOTE Confidence: 0.96193033

00:12:33.405 --> 00:12:34.605 as they're waking up going

NOTE Confidence: 0.96193033

00:12:34.605 --> 00:12:35.885 to school, and that's really

NOTE Confidence: 0.96193033

00:12:35.885 --> 00:12:36.385 driving

NOTE Confidence: 0.9587805

00:12:36.765 --> 00:12:37.265 potential,

NOTE Confidence: 0.9984649

00:12:37.725 --> 00:12:38.865 phenotypic traits.

NOTE Confidence: 0.99146163

00:12:40.540 --> 00:12:41.980 Okay. So that was kind

NOTE Confidence: 0.99146163

00:12:41.980 --> 00:12:42.380 of the,

NOTE Confidence: 0.9975522

00:12:42.940 --> 00:12:44.220 using this wearable data to

NOTE Confidence: 0.9975522

00:12:44.220 --> 00:12:44.960 make predictions,

NOTE Confidence: 0.94379675

00:12:45.660 --> 00:12:46.559 about phenotype.

NOTE Confidence: 0.99218404

00:12:47.179 --> 00:12:48.300 But then, of course, we're

NOTE Confidence: 0.99218404

00:12:48.300 --> 00:12:50.160 very interested in genetic discovery.

NOTE Confidence: 0.99218404

00:12:50.300 --> 00:12:51.679 So the question then becomes,

NOTE Confidence: 0.9955167

00:12:52.059 --> 00:12:53.100 if we move away from

NOTE Confidence: 0.9955167

00:12:53.100 --> 00:12:54.895 binary traits and go to

NOTE Confidence: 0.9955167

00:12:54.895 --> 00:12:56.915 these continuous based digital phenotypes,

NOTE Confidence: 0.9993992

00:12:57.295 --> 00:12:58.415 are we actually able to

NOTE Confidence: 0.9993992

00:12:58.415 --> 00:13:00.255 gain statistical power in terms

NOTE Confidence: 0.9993992

00:13:00.255 --> 00:13:01.395 of genetic discovery?

NOTE Confidence: 0.98857343

00:13:02.815 --> 00:13:04.434 So in order to evaluate

NOTE Confidence: 0.98857343

00:13:04.575 --> 00:13:06.015 that, the first thing we

NOTE Confidence: 0.98857343

00:13:06.015 --> 00:13:07.455 wanna do is establish kind

NOTE Confidence: 0.98857343

00:13:07.455 --> 00:13:09.554 of a baseline comparison, meaning

NOTE Confidence: 0.9939822

00:13:09.910 --> 00:13:11.450 let's use just the individuals

NOTE Confidence: 0.9939822

00:13:11.510 --> 00:13:12.710 we have and perform a

NOTE Confidence: 0.9939822

00:13:12.710 --> 00:13:14.550 traditional GWAS, meaning a case

NOTE Confidence: 0.9939822

00:13:14.550 --> 00:13:15.450 control study.

NOTE Confidence: 0.9699644

00:13:15.830 --> 00:13:17.190 And this would be using

NOTE Confidence: 0.9699644

00:13:17.190 --> 00:13:18.630 a zero or one oops.

NOTE Confidence: 0.9699644

00:13:18.630 --> 00:13:20.410 Sorry. A zero or one,

NOTE Confidence: 0.9929315

00:13:20.870 --> 00:13:22.550 for the disease and then

NOTE Confidence: 0.9929315

00:13:22.550 --> 00:13:23.929 the genotype here.

NOTE Confidence: 0.9807165

00:13:24.464 --> 00:13:25.584 And when we do such,

NOTE Confidence: 0.9807165

00:13:25.904 --> 00:13:27.824 analysis, this is using twelve

NOTE Confidence: 0.9807165

00:13:27.824 --> 00:13:28.725 hundred individuals.

NOTE Confidence: 0.9422883

00:13:29.345 --> 00:13:30.404 Perhaps unsurprisingly,

NOTE Confidence: 0.9915664

00:13:30.944 --> 00:13:32.144 we we don't find any

NOTE Confidence: 0.9915664

00:13:32.144 --> 00:13:34.144 genetic loci at genome wide

NOTE Confidence: 0.9915664

00:13:34.144 --> 00:13:36.005 significance above this blue line.

NOTE Confidence: 0.9915664

00:13:36.225 --> 00:13:38.065 And twelve hundred individuals is

NOTE Confidence: 0.9915664

00:13:38.065 --> 00:13:39.400 is really not that many

NOTE Confidence: 0.9915664

00:13:39.400 --> 00:13:41.240 individuals. Most GWAS studies might

NOTE Confidence: 0.9915664

00:13:41.240 --> 00:13:42.760 have hundreds or millions of

NOTE Confidence: 0.9915664

00:13:42.760 --> 00:13:43.260 individuals.

NOTE Confidence: 0.95772576

00:13:44.360 --> 00:13:44.860 So,

NOTE Confidence: 0.9997714

00:13:45.160 --> 00:13:46.600 again, this is not using

NOTE Confidence: 0.9997714

00:13:46.600 --> 00:13:47.900 any of the wearable data.

NOTE Confidence: 0.9459895

00:13:49.080 --> 00:13:49.900 Now instead,

NOTE Confidence: 0.9983546

00:13:50.600 --> 00:13:51.800 let's imagine we have this

NOTE Confidence: 0.9983546

00:13:51.800 --> 00:13:53.204 wearable data as kind of

NOTE Confidence: 0.9983546

00:13:53.204 --> 00:13:53.944 a multidimensional

NOTE Confidence: 0.87763083

00:13:54.485 --> 00:13:56.264 array, a digital phenotype,

NOTE Confidence: 0.8799044

00:13:56.804 --> 00:13:58.345 and this vector d.

NOTE Confidence: 0.97041595

00:13:58.804 --> 00:14:00.245 And instead, now let's model

NOTE Confidence: 0.97041595

00:14:00.245 --> 00:14:02.024 that, against the genotype.

NOTE Confidence: 0.96899086

00:14:02.964 --> 00:14:05.144 And we also include this,

NOTE Confidence: 0.96899086

00:14:05.365 --> 00:14:06.990 genotype by macrophenotype

NOTE Confidence: 0.9956238

00:14:07.290 --> 00:14:08.890 interaction term to ensure that

NOTE Confidence: 0.9956238

00:14:08.890 --> 00:14:10.170 any changes with the digital

NOTE Confidence: 0.9956238

00:14:10.170 --> 00:14:11.850 phenotype are actually tied to

NOTE Confidence: 0.9956238

00:14:11.850 --> 00:14:12.990 the disease itself.

NOTE Confidence: 0.98449385

00:14:13.690 --> 00:14:15.050 And what we find here

NOTE Confidence: 0.98449385

00:14:15.050 --> 00:14:16.250 in the same exact set

NOTE Confidence: 0.98449385

00:14:16.250 --> 00:14:17.530 of individuals that you previously

NOTE Confidence: 0.98449385

00:14:17.530 --> 00:14:18.590 saw, the twelve hundred,

NOTE Confidence: 0.95653415

00:14:19.295 --> 00:14:20.815 individuals, now we're able to

NOTE Confidence: 0.95653415

00:14:20.815 --> 00:14:22.015 see much more enrichment in

NOTE Confidence: 0.95653415

00:14:22.015 --> 00:14:24.255 statistical power. We identified two

NOTE Confidence: 0.95653415

00:14:24.255 --> 00:14:25.635 significant genetic loci.

NOTE Confidence: 0.99802494

00:14:26.175 --> 00:14:27.535 And just to highlight that

NOTE Confidence: 0.99802494

00:14:27.535 --> 00:14:28.975 a little bit, this this

NOTE Confidence: 0.99802494

00:14:28.975 --> 00:14:30.035 particular loci

NOTE Confidence: 0.9536345

00:14:30.415 --> 00:14:30.915 is,

NOTE Confidence: 0.9925504

00:14:31.295 --> 00:14:33.295 related to sedentary time, and

NOTE Confidence: 0.9925504

00:14:33.295 --> 00:14:34.415 you can see that there's

NOTE Confidence: 0.9925504

00:14:34.415 --> 00:14:36.490 a clear difference between individuals

NOTE Confidence: 0.9925504

00:14:36.550 --> 00:14:38.070 that are healthy controls and

NOTE Confidence: 0.9925504

00:14:38.070 --> 00:14:39.690 those individuals that have ADHD.

NOTE Confidence: 0.9312769

00:14:40.310 --> 00:14:41.050 But, additionally,

NOTE Confidence: 0.9966742

00:14:41.510 --> 00:14:42.470 in the group that,

NOTE Confidence: 0.9506235

00:14:43.030 --> 00:14:44.630 individuals of ADHD, you can

NOTE Confidence: 0.9506235

00:14:44.630 --> 00:14:46.710 see that sedentary time drop

NOTE Confidence: 0.9506235

00:14:46.710 --> 00:14:47.210 significantly

NOTE Confidence: 0.99856365

00:14:47.935 --> 00:14:49.215 as we go through, the

NOTE Confidence: 0.99856365

00:14:49.215 --> 00:14:50.115 different genotypes.

NOTE Confidence: 0.98907125

00:14:50.575 --> 00:14:51.855 And so, really, the the

NOTE Confidence: 0.98907125

00:14:51.855 --> 00:14:52.975 result here is kind of

NOTE Confidence: 0.98907125

00:14:52.975 --> 00:14:54.755 establishing this relationship

NOTE Confidence: 0.99955124

00:14:55.375 --> 00:14:57.455 between the disease itself, the

NOTE Confidence: 0.99955124

00:14:57.455 --> 00:14:57.955 genetics,

NOTE Confidence: 0.98109645

00:14:58.495 --> 00:15:00.435 but also these digital phenotypes.

NOTE Confidence: 0.9989552

00:15:03.380 --> 00:15:05.060 And I'll present kind of

NOTE Confidence: 0.9989552

00:15:05.060 --> 00:15:06.180 one other way that we've

NOTE Confidence: 0.9989552

00:15:06.180 --> 00:15:07.240 tackled this problem,

NOTE Confidence: 0.870576

00:15:07.700 --> 00:15:08.440 which is,

NOTE Confidence: 0.99013126

00:15:08.900 --> 00:15:10.760 if we recall from before,

NOTE Confidence: 0.9962292

00:15:11.220 --> 00:15:12.500 using those machine learning and

NOTE Confidence: 0.9962292

00:15:12.500 --> 00:15:13.780 deep learning models, we're able

NOTE Confidence: 0.9962292

00:15:13.780 --> 00:15:15.480 to generate these digital phenotype

NOTE Confidence: 0.9687743

00:15:15.860 --> 00:15:17.720 or AI generated risk scores.

NOTE Confidence: 0.9653911

00:15:18.365 --> 00:15:19.565 And and those scores really

NOTE Confidence: 0.9653911

00:15:19.565 --> 00:15:20.764 are aggregating all of the

NOTE Confidence: 0.9653911

00:15:20.764 --> 00:15:21.985 information from the smartwatch.

NOTE Confidence: 0.94994533

00:15:22.365 --> 00:15:24.065 Instead, now let's use those

NOTE Confidence: 0.94994533

00:15:24.365 --> 00:15:25.324 as a,

NOTE Confidence: 0.9660576

00:15:26.045 --> 00:15:28.144 target for the genetic discovery.

NOTE Confidence: 0.9660576

00:15:28.285 --> 00:15:29.324 So here, the d, the

NOTE Confidence: 0.9660576

00:15:29.324 --> 00:15:30.845 digital phenotyping score as a

NOTE Confidence: 0.9660576

00:15:30.845 --> 00:15:31.985 function of the genetics.

NOTE Confidence: 0.9992586

00:15:32.764 --> 00:15:33.264 And

NOTE Confidence: 0.95138574

00:15:33.710 --> 00:15:35.070 using this method, again, we

NOTE Confidence: 0.95138574

00:15:35.070 --> 00:15:36.830 we increase the statistical power

NOTE Confidence: 0.95138574

00:15:36.830 --> 00:15:37.870 even more. So now ten

NOTE Confidence: 0.95138574

00:15:37.870 --> 00:15:38.770 genetic loci.

NOTE Confidence: 0.9853341

00:15:39.790 --> 00:15:41.150 And many of these, the

NOTE Confidence: 0.9853341

00:15:41.150 --> 00:15:42.510 pink and the blue represent

NOTE Confidence: 0.9853341

00:15:42.510 --> 00:15:44.270 psychiatric associated genes that are

NOTE Confidence: 0.9853341

00:15:44.270 --> 00:15:44.770 known.

NOTE Confidence: 0.9530967

00:15:45.310 --> 00:15:47.090 Some are ADHD associated,

NOTE Confidence: 0.73629475

00:15:48.074 --> 00:15:49.134 but also we identify,

NOTE Confidence: 0.99984103

00:15:49.834 --> 00:15:51.675 an array of novel targets

NOTE Confidence: 0.99984103

00:15:51.675 --> 00:15:52.334 as well.

NOTE Confidence: 0.97842073

00:15:53.035 --> 00:15:54.394 And so just kind of

NOTE Confidence: 0.97842073

00:15:54.394 --> 00:15:55.855 in in in summary here,

NOTE Confidence: 0.97842073

00:15:55.915 --> 00:15:57.915 really, the idea for the

NOTE Confidence: 0.97842073

00:15:57.915 --> 00:15:58.894 digital phenotypes

NOTE Confidence: 0.9805574

00:15:59.435 --> 00:16:01.290 is, number one, we can

NOTE Confidence: 0.9805574

00:16:01.449 --> 00:16:02.889 use it to drive clinical

NOTE Confidence: 0.9805574

00:16:02.889 --> 00:16:03.389 discovery,

NOTE Confidence: 0.98867744

00:16:03.850 --> 00:16:06.009 clinical characterization of individuals and

NOTE Confidence: 0.98867744

00:16:06.009 --> 00:16:06.670 their subtypes,

NOTE Confidence: 0.9744498

00:16:07.209 --> 00:16:08.829 but also to really drive

NOTE Confidence: 0.90415806

00:16:09.209 --> 00:16:10.190 forward the,

NOTE Confidence: 0.9966195

00:16:11.449 --> 00:16:12.670 the genetics discovery.

NOTE Confidence: 0.9949515

00:16:13.385 --> 00:16:14.825 And I would just wanna

NOTE Confidence: 0.9949515

00:16:14.825 --> 00:16:16.505 emphasize that this,

NOTE Confidence: 0.9992236

00:16:16.985 --> 00:16:17.485 this

NOTE Confidence: 0.9959159

00:16:17.865 --> 00:16:19.464 result here of course, there's

NOTE Confidence: 0.9959159

00:16:19.464 --> 00:16:21.325 there's many questions about causality,

NOTE Confidence: 0.96891236

00:16:22.265 --> 00:16:23.625 and that's kind of ongoing

NOTE Confidence: 0.96891236

00:16:23.625 --> 00:16:24.745 work that we we wanna

NOTE Confidence: 0.96891236

00:16:24.745 --> 00:16:25.949 address and, of course, that

NOTE Confidence: 0.96891236

00:16:25.949 --> 00:16:27.230 this is just the starting

NOTE Confidence: 0.96891236

00:16:27.230 --> 00:16:29.089 point of using digital phenotypes.

NOTE Confidence: 0.96891236

00:16:29.389 --> 00:16:30.589 We there's many much more

NOTE Confidence: 0.96891236

00:16:30.589 --> 00:16:31.790 translational work that needs to

NOTE Confidence: 0.96891236

00:16:31.790 --> 00:16:33.870 be done, and validation as

NOTE Confidence: 0.96891236

00:16:33.870 --> 00:16:34.370 well.

NOTE Confidence: 0.9927921

00:16:36.269 --> 00:16:36.769 So,

NOTE Confidence: 0.9963923

00:16:37.389 --> 00:16:38.430 to kind of end and

NOTE Confidence: 0.9963923

00:16:38.430 --> 00:16:39.885 wrap up here, the the

NOTE Confidence: 0.9963923

00:16:39.885 --> 00:16:40.685 future work,

NOTE Confidence: 0.9958564

00:16:41.405 --> 00:16:42.845 of course, is expanding to

NOTE Confidence: 0.9958564

00:16:42.845 --> 00:16:44.385 other modalities and diseases.

NOTE Confidence: 0.9723864

00:16:45.405 --> 00:16:46.625 The hope here is that

NOTE Confidence: 0.9723864

00:16:46.685 --> 00:16:48.545 this framework, this triangle here,

NOTE Confidence: 0.9723864

00:16:48.605 --> 00:16:49.885 can be used for things

NOTE Confidence: 0.9723864

00:16:49.885 --> 00:16:51.645 like Parkinson's disease and for

NOTE Confidence: 0.9723864

00:16:51.645 --> 00:16:53.325 other kind of, related movement

NOTE Confidence: 0.9723864

00:16:53.325 --> 00:16:54.705 disorders and other neurodegeneration.

NOTE Confidence: 0.9416473

00:16:55.940 --> 00:16:57.240 And the idea is to

NOTE Confidence: 0.9416473

00:16:57.300 --> 00:16:59.080 move away from directly associating

NOTE Confidence: 0.9416473

00:16:59.220 --> 00:17:00.500 the genotype to the macro

NOTE Confidence: 0.9416473

00:17:00.500 --> 00:17:02.040 phenotype, but instead

NOTE Confidence: 0.98115474

00:17:02.580 --> 00:17:04.740 using this intermediate phenotype to

NOTE Confidence: 0.98115474

00:17:04.740 --> 00:17:05.400 to characterize

NOTE Confidence: 0.95052624

00:17:05.780 --> 00:17:08.280 both molecularly, physiologically, and behaviorally

NOTE Confidence: 0.9987735

00:17:08.775 --> 00:17:09.275 individuals

NOTE Confidence: 0.9855703

00:17:09.815 --> 00:17:11.675 better, provide those clinical sorry.

NOTE Confidence: 0.9331482

00:17:12.055 --> 00:17:13.515 Provide those clinical insights,

NOTE Confidence: 0.9643158

00:17:13.895 --> 00:17:15.195 and, of course, most importantly,

NOTE Confidence: 0.95643806

00:17:15.575 --> 00:17:16.475 is to retrace

NOTE Confidence: 0.9997581

00:17:16.935 --> 00:17:18.395 back to the genetic information

NOTE Confidence: 0.9447879

00:17:19.095 --> 00:17:21.994 and, link to potential molecular

NOTE Confidence: 0.9447879

00:17:22.055 --> 00:17:22.555 mechanisms.

NOTE Confidence: 0.9648231

00:17:23.050 --> 00:17:24.330 And and there's a lot

NOTE Confidence: 0.9648231

00:17:24.330 --> 00:17:25.790 of ongoing work right now.

NOTE Confidence: 0.9977339

00:17:26.170 --> 00:17:27.130 Of course, we heard a

NOTE Confidence: 0.9977339

00:17:27.130 --> 00:17:28.490 little bit about the spatial

NOTE Confidence: 0.9977339

00:17:28.490 --> 00:17:28.990 data.

NOTE Confidence: 0.9877596

00:17:29.690 --> 00:17:31.130 I'm very excited about the

NOTE Confidence: 0.9877596

00:17:31.130 --> 00:17:33.230 digital health sphere using wearables,

NOTE Confidence: 0.98478234

00:17:33.609 --> 00:17:35.530 smartwatches, but also video data,

NOTE Confidence: 0.98478234

00:17:35.530 --> 00:17:36.750 video capture data,

NOTE Confidence: 0.99750584

00:17:37.050 --> 00:17:38.170 as well as brain imaging

NOTE Confidence: 0.99750584

00:17:38.170 --> 00:17:38.670 data.

NOTE Confidence: 0.9921805

00:17:38.984 --> 00:17:39.725 And so,

NOTE Confidence: 0.939606

00:17:40.585 --> 00:17:42.984 yes, happy to, collaborate and

NOTE Confidence: 0.939606

00:17:42.984 --> 00:17:43.385 and,

NOTE Confidence: 0.9721721

00:17:43.785 --> 00:17:45.065 work on this with others

NOTE Confidence: 0.9721721

00:17:45.065 --> 00:17:46.265 more. So I'll I'll end

NOTE Confidence: 0.9721721

00:17:46.265 --> 00:17:48.205 here just with the acknowledgment

NOTE Confidence: 0.9721721

00:17:48.425 --> 00:17:49.945 slide. In particular, I just

NOTE Confidence: 0.9721721

00:17:49.945 --> 00:17:51.565 wanna highlight Mark Gerstein,

NOTE Confidence: 0.9245277

00:17:51.945 --> 00:17:53.145 whose lab I'm in, and

NOTE Confidence: 0.9245277

00:17:53.145 --> 00:17:55.119 also Walter Roberts from Yale

NOTE Confidence: 0.9245277

00:17:55.119 --> 00:17:56.260 Psychiatry who

NOTE Confidence: 0.90234005

00:17:56.640 --> 00:17:56.960 we,

NOTE Confidence: 0.93134123

00:17:58.080 --> 00:17:59.680 they they're the co co

NOTE Confidence: 0.93134123

00:17:59.680 --> 00:18:01.119 senior authors on the study.

NOTE Confidence: 0.93134123

00:18:01.119 --> 00:18:02.000 And then also I just

NOTE Confidence: 0.93134123

00:18:02.000 --> 00:18:03.600 wanna highlight Beatrice who co

NOTE Confidence: 0.93134123

00:18:03.600 --> 00:18:04.560 led this study with me

NOTE Confidence: 0.93134123

00:18:04.560 --> 00:18:06.160 and, Yun Yang, a very

NOTE Confidence: 0.93134123

00:18:06.160 --> 00:18:07.680 talented graduate student who's been

NOTE Confidence: 0.93134123

00:18:07.680 --> 00:18:08.932 working with me, on this

NOTE Confidence: 0.93134123

00:18:08.932 --> 00:18:10.452 project as well, as well

NOTE Confidence: 0.93134123

00:18:10.452 --> 00:18:11.512 as some of the collaborators,

NOTE Confidence: 0.98459685

00:18:11.972 --> 00:18:14.132 from Barcelona and California. So

NOTE Confidence: 0.98459685

00:18:14.132 --> 00:18:15.252 with that, thank you very

NOTE Confidence: 0.98459685

00:18:15.252 --> 00:18:15.752 much.

NOTE Confidence: 0.7274739

00:18:16.132 --> 00:18:16.632 Yep.